

إدارة الشؤون الاقتصادية والاجتماعية  
الشعبة الإحصائية

السلسلة واو، العدد ٩٨

دراسات في الطرق

# تصميم عيّنات استقصاءات الأسر المعيشية: مبادئ توجيهية عملية



الأمم المتحدة  
نيويورك، ٢٠١٠

## إدارة الشؤون الاقتصادية والاجتماعية

إدارة الشؤون الاقتصادية والاجتماعية التابعة للأمانة العامة للأمم المتحدة حلقة وصل حيوية بين السياسات العالمية في المجالات الاقتصادي والاجتماعي والبيئي والعمل الوطني. تعمل الإدارة في ثلاثة مجالات رئيسية مترابطة، هي: '١' تُجمّع وتولّد وتحلّل سلسلة واسعة من البيانات والمعلومات الاقتصادية والاجتماعية والبيئية، تستفيد منها الدول الأعضاء في الأمم المتحدة لاستعراض المشاكل المشتركة وتقييم الخيارات السياسية؛ '٢' تيسّر مفاوضات الدول الأعضاء في هيئات حكومية - دولية كثيرة حول مسارات عمل مشتركة للتصدّي لتحديات عالمية مستمرة أو ناشئة؛ '٣' تسدي المشورة للحكومات المهتمة بشأن طرق ووسائل ترجمة أطر السياسة التي توضع في مؤتمرات الأمم المتحدة ومؤتمرات القمة إلى برامج على الصعيد القطري، بواسطة المساعدة التقنية، وتساعد على بناء القدرات الوطنية.

### ملاحظة

ليس في التسميات المستخدمة في هذا المنشور ولا في طريقة عرض مادته ما يتضمّن التعبير عن أيّ رأي للأمانة العامة للأمم المتحدة بشأن المركز القانوني لأيّ بلد أو إقليم أو مدينة أو منطقة، أو لسلطات أيّ منها، أو بشأن تعيين حدودها أو تخومها.

يشير مصطلح "البلد" كما هو مستخدم في هذا المنشور أيضاً، حسب الاقتضاء، إلى أقاليم ومناطق.

يُقصد بتسميات "البلدان أو المناطق المتقدمة النمو" أو "النامية" أو المناطق "الأكثر تقدماً" أو "الأقل تقدماً" أو "أقل البلدان تقدماً" أن تُستخدم تيسيراً للإحصائيات، ولا تعبر بالضرورة عن أيّ حكم في المرحلة التي بلغها أيّ بلد أو منطقة من مراحل عملية التنمية.

تتألف رموز وثائق الأمم المتحدة من حروف كبيرة وأرقام. ويعني ذكر الرمز إشارة إلى وثيقة من وثائق الأمم المتحدة.

ST/ESA/STAT/SER.F/98

ISBN 978-92-1-661025-8

منشورات الأمم المتحدة

رقم المبيع: A.06.XVII.13

حقوق المؤلف © مسجلة للأمم المتحدة، ٢٠٠٨

جميع الحقوق محفوظة

## تصدير

الغرض الرئيسي من تصميم عيّنات استقصاءات الأسر المعيشية: مبادئ توجيهية عملية هو أن تكون بمثابة دليل يشمل، في منشور واحد، المسائل الرئيسية لتصميم عيّنات للاستقصاءات يمكن أن يرجع إليها الإحصائيون الوطنيون الممارسون، والباحثون والمحلّون المعنيون بإجراء أعمال وأنشطة في مجال الاستقصاء بالعيّنة في البلدان. وتُستخدم في هذا الدليل تقنيات سليمة منهجياً، راسخة في النظرية الإحصائية، وهذا ينطوي على استخدام عيّنات احتمالية في كل مرحلة من مراحل عملية اختيار العيّنّة. ويمكن لاستقصاءات الأسر المعيشية المصمّمة جيداً، والمنفذة تنفيذاً صحيحاً، أن تولّد المعلومات اللازمة بقدر كافٍ من الجودة، والدقة بسرعة وبتكلفة منخفضة نسبياً.

يمكن أيضاً استخدام محتويات هذا المنشور، جزئياً، كدليل تدريبي لدورات تمهيدية في تصميم عيّنات الاستقصاءات في مؤسسات مختلفة للتدريب الإحصائي تقدم دورات في الإحصاءات التطبيقية، لا سيما في منهجية الاستقصاء بالعيّنة.

بالإضافة إلى ذلك، أعدّ هذا المنشور ليكون تكملة لمنشورات أخرى تتناول منهجية الاستقصاءات بالعيّنة في البلدان النامية والبلدان التي تمر بمرحلة انتقالية<sup>1</sup> والسلسلة المنشورة تحت برنامج القدرات الوطنية على إجراء استقصاءات الأسر المعيشية (NHSCP).

أهداف هذا الدليل على وجه أكثر تحديداً هي:

(أ) أن يقدم، في منشور واحد، مفاهيم أساسية وإجراءات سليمة منهجياً لتصميم عيّنات استقصاءات للأسر المعيشية على الصعيد الوطني، على وجه التحديد، مع توكيد الجوانب التطبيقية لتصميم عيّنات استقصاءات الأسر المعيشية؛

(ب) أن يكون بمثابة دليل عملي للقائمين بإجراء الاستقصاءات في تصميم وتنفيذ عيّنات ذات كفاءة لدراسات استقصائية للأسر المعيشية؛

(ج) أن يوضح الترابط بين تصميم العيّنات، وجمع البيانات، والتقدير، والإعداد، والتحليل؛

(د) أن يبرز أهمية التحكم بالأخطاء غير أخطاء المعاينة في عيّنات استقصاءات الأسر المعيشية بالعيّنة.

مع أن وجود خلفية في المعاينة (أخذ العيّنات) سيُساعد مستخدمي هذا الدليل، سوف يتمكن الأشخاص الآخرون، الذين لديهم معرفة عامة بالمفاهيم الإحصائية والرياضية، من استخدامه وتطبيق محتوياته بشيء قليل من المساعدة أو حتى من دون مساعدة. والسبب في ذلك هو أن واحداً من المقاصد الرئيسية للدليل هو تقديم المواد في صورة عملية جاهزة، خلافاً لتوكيد الجوانب النظرية للمعاينة. غير أن توكيدات نظرية تقدّم عندما يكون ذلك

<sup>1</sup> دراسات في الطرق؛ العدد ٩٦ (منشورات الأمم المتحدة، رقم المبيع: E.05.XVII.6).

ضرورياً. ويتوقّع أن تكون معرفة مبادئ علم الجبر هي كل ما يحتاج إليه المرء لمتابعة المواد المعروضة بسهولة وتطبيق التقنيات. بناءً على ذلك، تُقدّم أمثلة عديدة لتوضيح المفاهيم والطرق والتقنيات.

ساهم عدد من الخبراء في إعداد هذا الدليل. فقد قام السيد أنتوني تيرنر، خبير استشاري في المعاينة، بكتابة الفصول الثالث والرابع والخامس، وراجع الوثيقة بصيغتها النهائية الموحّدة؛ وقام السيد إبراهيم يانسانة، نائب رئيس شعبة تكاليف المعيشة، التابعة للجنة الخدمة المدنية الدولية، بكتابة الفصلين السادس والسابع؛ وقام السيد مافيون جامبوا، الخبير الإحصائي في أمانة الجماعة الإنمائية للجنوب الإفريقي، بكتابة الفصل التاسع.

وقام السيد جيريميا باندا، الموظف في الشعبة الإحصائية بالأمم المتحدة، الذي كان المحرّر الرئيسي والمنسق الفني لهذا الدليل بكتابة الفصول الأول والثاني والثامن، بما في ذلك المرفق الأول. وساعدت السيدة كليز مينوزي في تحرير المسودة الأولى لفصول مختلفة؛ وقدمت السيدة بيزوجينيت كاسا مساعدة قيمة في الأعمال السكرتارية.

قام بمراجعة مسودات الفصول فريق من الخبراء في اجتماع نظّمته الشعبة الإحصائية، وعقد في نيويورك في الفترة من ٣ إلى ٥ كانون الأول/ديسمبر ٢٠٠٣. ترد في المرفق الثاني قائمة بالمشاركين. بالإضافة إلى ذلك، قام بمراجعة الدليل، على مستوى الأقران، الدكتور ألفريدو بوستوس، والسيدة أنا ماريا لانديروس، والسيد إدواردو ريوس، من المعهد الوطني المكسيكي للإحصاءات والجغرافيا والمعلومات الإلكترونية، وقدموا تعليقات قيمة جداً عليه.

بول شيونغ

المدير

الشعبة الإحصائية بالأمم المتحدة

إدارة الشؤون الاقتصادية والاجتماعية

# المحتويات

الصفحة

ج..... تصدير

## الفصل الأول

### مصادر البيانات للإحصاءات الاجتماعية والديمغرافية

١ - ١	مقدمة	١
٢ - ١	مصادر البيانات	١
١ - ٢ - ١	استقصاءات الأسر المعيشية	٢
٢ - ٢ - ١	عمليات تعداد السكان والمساكن	٤
٣ - ٢ - ١	السجلات الإدارية	٥
٤ - ٢ - ١	أوجه التكامل بين المصادر الثلاثة للبيانات	٦
٥ - ٢ - ١	ملاحظات ختامية	٨
٨	المراجع ومواد أخرى للقراءة	٨

## الفصل الثاني

### تخطيط الاستقصاءات وتنفيذها

١ - ٢	تخطيط الاستقصاءات	١١
١ - ١ - ٢	أهداف الاستقصاء	١١
٢ - ١ - ٢	عالم الاستقصاء	١٣
٣ - ١ - ٢	المعلومات المراد جمعها	١٤
٤ - ١ - ٢	ميزانية الاستقصاء	١٤
٢ - ٢	تنفيذ الاستقصاءات	١٤
١ - ٢ - ٢	طرق جمع البيانات	١٤
٢ - ٢ - ٢	تصميم الاستبيان	٢٠
٣ - ٢ - ٢	خطة الجدولة والتحليل	٢٣
٤ - ٢ - ٢	تنفيذ العمل الميداني	٢٣
٢٧	المراجع ومواد أخرى للقراءة	٢٧

## الفصل الثالث

### استراتيجيات المعاينة

١ - ٣	مقدمة	٢٩
١ - ١ - ٣	نظرة عامة	٢٩

## الصفحة

٣١.....	٢ - ١ - ٣	قائمة مصطلحات المعاينة ومصطلحات ذات صلة بالموضوع.....
٣٣.....	٣ - ١ - ٣	ملاحظات.....
٣٤.....	٢ - ٣	العيّنات الاحتمالية مقابل والطرق الأخرى للمعاينة في استقصاءات الأسر المعيشية.....
٣٤.....	١ - ٢ - ٣	المعاينة الاحتمالية.....
٣٦.....	٢ - ٢ - ٣	طرق المعاينة غير الاحتمالية.....
٣٨.....	٣ - ٣	تحديد حجم العيّنة لاستقصاءات الأسر المعيشية.....
٣٩.....	١ - ٣ - ٣	أحجام تقديرات الاستقصاء.....
٤٠.....	٢ - ٣ - ٣	المجموعة السكانية المستهدفة.....
٤٠.....	٣ - ٣ - ٣	الدقة والثقة الإحصائية.....
٤١.....	٤ - ٣ - ٣	مجموعات التحليل: المجالات.....
٤٤.....	٥ - ٣ - ٣	آثار التقسيم إلى مجموعات.....
٤٥.....	٦ - ٣ - ٣	تعديل حجم العيّنة لعدم الاستجابة المتوقع.....
٤٥.....	٧ - ٣ - ٣	حجم العيّنة للعيّنات الرئيسية.....
٤٦.....	٨ - ٣ - ٣	تقدير التغيّر أو المستوى.....
٤٦.....	٩ - ٣ - ٣	ميزانية الاستقصاء.....
٤٦.....	١٠ - ٣ - ٣	حساب حجم العيّنة.....
٤٩.....	٤ - ٣	التقسيم إلى طبقات.....
٤٩.....	١ - ٤ - ٣	التقسيم إلى طبقات وتخصيص العيّنات.....
٥٠.....	٢ - ٤ - ٣	قواعد التقسيم إلى طبقات.....
٥١.....	٣ - ٤ - ٣	التقسيم الضمني إلى طبقات.....
٥٢.....	٥ - ٣	معاينة المجموعات.....
٥٤.....	١ - ٥ - ٣	خصائص معاينة المجموعات.....
٥٤.....	٢ - ٥ - ٣	أثر تصميم المجموعة.....
٥٥.....	٣ - ٥ - ٣	حجم المجموعة.....
٥٧.....	٤ - ٥ - ٣	حساب أثر التصميم (deff).....
٥٧.....	٥ - ٥ - ٣	عدد المجموعات.....
٥٧.....	٦ - ٣	المعاينة على مراحل.....
٥٧.....	١ - ٦ - ٣	فوائد المعاينة على مراحل.....
٥٨.....	٢ - ٦ - ٣	استخدام المراحل الصورية.....
٦١.....	٣ - ٦ - ٣	التصميم على مرحلتين.....
٦١.....	٧ - ٣	المعاينة باحتمال متناسب مع الحجم و باحتمال متناسب مع الحجم المقدّر.....
٦٢.....	١ - ٧ - ٣	المعاينة باحتمال متناسب مع الحجم.....
٦٥.....	٢ - ٧ - ٣	المعاينة باحتمال متناسب مع الحجم المقدّر.....
٦٦.....	٨ - ٣	الخيارات في المعاينة.....
	١ - ٨ - ٣	المعاينة مع تساوي الاحتمالات، والمعاينة باحتمال متناسب مع الحجم،
٦٦.....		والمعاينة مع الحجم الثابت ومع النسبة الثابتة.....

## الصفحة

٧٠	الاستقصاء الديمغرافي والصحي	٢ - ٨ - ٣
٧١	التصميم المعدل للمجموعة: استقصاءات المجموعات المتعددة المؤشرات	٣ - ٨ - ٣
٧٣	مواضيع خاصة: العينات ذات المرحلتين والمعينة لتحديد الاتجاهات	٩ - ٣
٧٣	المعينة على مرحلتين	١ - ٩ - ٣
٧٥	المعينة لتقدير التغيُّر أو الاتجاه	٢ - ٩ - ٣
٧٧	عندما يحدث خطأ في التنفيذ	١٠ - ٣
٧٨	تعريف السكان المستهدفين وتغطيتهم	١ - ١٠ - ٣
٧٩	حجم العينة أكبر مما تتحمّله ميزانية الاستقصاء	٢ - ١٠ - ٣
٧٩	حجم المجموعة أكبر أو أصغر مما كان متوقَّعاً	٣ - ١٠ - ٣
٧٩	تناؤل حالات عدم الاستجابة	٤ - ١٠ - ٣
٨٠	موجز المبادئ التوجيهية	١١ - ٣
٨١	المراجع ومواد أخرى للقراءة	

## الفصل الرابع

## أطر المعاينة والعيّنات الرئيسية

٨٣	أطر المعاينة في استقصاءات الأسر المعيشية	١ - ٤
٨٣	تعريف إطار المعاينة	١ - ١ - ٤
٨٤	خصائص أطر المعاينة	٢ - ١ - ٤
٨٦	أطر المناطق	٣ - ١ - ٤
٨٧	أطر القوائم	٤ - ١ - ٤
٨٨	الأطر المتعددة	٥ - ١ - ٤
٩٠	الإطار النمطي (الأطر النمطية) في تصاميم ذات مرحلتين	٦ - ١ - ٤
٩٠	أطر المعاينة الرئيسية	٧ - ١ - ٤
٩١	المشاكل المشتركة بين الأطر والإصلاحات المقترحة	٨ - ١ - ٤
٩٤	أطر المعاينة الرئيسية	٢ - ٤
٩٤	تعريف العينة الرئيسية واستخداماتها	١ - ٢ - ٤
٩٥	الخصائص المثالية لوحدات المعاينة الأولية المستخدمة في إطار العينة الرئيسية	٢ - ٢ - ٤
٩٦	استخدام العينات الرئيسية لدعم الاستقصاءات	٣ - ٢ - ٤
٩٧	المخصّصات للمجالات المختلفة (المناطق الإدارية، إلخ.)	٤ - ٢ - ٤
٩٨	صيانة العينات الرئيسية وتحديثها	٥ - ٢ - ٤
٩٩	تناوب وحدات المعاينة الأولية في العينات الرئيسية	٦ - ٢ - ٤
٩٩	أمثلة قطرية للعينات الرئيسية	٧ - ٢ - ٤
١٠٦	موجز المبادئ التوجيهية	٣ - ٤
١٠٧	المراجع ومواد أخرى للقراءة	

## الفصل الخامس

## توثيق تصاميم العيّنات وتقييمها

١٠٩.....	١ - ٥	مقدمة.....
١٠٩.....	٢ - ٥	الحاجة إلى توثيق العيّنات وتقييمها وأنواع هذا التوثيق والتقييم.....
١١٠.....	٣ - ٥	وضع بطاقات عناوين لمتغيرات التصميم.....
١١٢.....	٤ - ٥	احتمالات الاختيار.....
١١٢.....	٥ - ٥	نسب الاستجابة ونسب التغطية في المراحل المختلفة من مراحل اختيار العيّنات.....
١١٤.....	٦ - ٥	الترجيح: مُعاملات الترجيح الأساسية، وعدم الاستجابة وتعديلات أخرى.....
١١٤.....	٧ - ٥	معلومات عن تكاليف المعاينة وتنفيذ الاستقصاء.....
١١٥.....	٨ - ٥	التقييم: أوجه القصور في بيانات الاستقصاء.....
١١٧.....	٩ - ٥	موجز المبادئ التوجيهية.....
١١٨.....		المراجع ومواد أخرى للقراءة.....

## الفصل السادس

## بناء مُعاملات ترجيح العيّنات واستخدامها

١١٩.....	١ - ٦	مقدمة.....
١١٩.....	٢ - ٦	الحاجة إلى مُعاملات ترجيح العيّنات.....
١٢٠.....	١ - ٢ - ٦	نظرة عامة.....
١٢٠.....	٣ - ٦	تطوير مُعاملات ترجيح العيّنات.....
١٢١.....	١ - ٣ - ٦	تعديلات مُعاملات ترجيح العيّنة للأهلية غير المعلومة.....
١٢٢.....	٢ - ٣ - ٦	تعديلات مُعاملات ترجيح العيّنة لمراعاة النسخ المكررة.....
١٢٢.....	٤ - ٦	الترجيح لاحتمالات الاختيار غير المتساوية.....
١٢٦.....	١ - ٤ - ٦	دراسة حالة إفرادية في مُعاملات الترجيح: استقصاء فييت نام الصحي الوطني، ٢٠٠١.....
١٢٧.....	٢ - ٤ - ٦	العيّنات ذاتية الترجيح.....
١٢٨.....	٥ - ٦	تعديل مُعاملات ترجيح العيّنات لعدم الاستجابة.....
١٢٨.....	١ - ٥ - ٦	تقليل التحيز الناتج عن عدم الاستجابة في استقصاءات الأسر المعيشية.....
١٢٩.....	٢ - ٥ - ٦	التعويض عن عدم الاستجابة.....
١٣٠.....	٣ - ٥ - ٦	تعديل مُعاملات ترجيح العيّنة بسبب عدم الاستجابة.....
١٣١.....	٦ - ٦	تعديل مُعاملات ترجيح العيّنة لتعويض عدم التغطية.....
١٣٢.....	١ - ٦ - ٦	مصادر عدم التغطية في استقصاءات الأسر المعيشية.....
١٣٣.....	٢ - ٦ - ٦	التعويض عن عدم التغطية في استقصاءات الأسر المعيشية.....
١٣٤.....	٧ - ٦	زيادة التباين في العيّنة نتيجة للترجيح.....
١٣٦.....	٨ - ٦	تخفيض مُعاملات الترجيح.....
١٣٧.....	٩ - ٦	ملاحظات ختامية.....
١٣٨.....		المراجع ومواد أخرى للقراءة.....

## الفصل السابع

## تقدير أخطاء المعاينة لبيانات الاستقصاء

١٤١	١ - ٧	مقدمة
١٤١	١ - ١ - ٧	تقدير أخطاء المعاينة لبيانات الاستقصاء المركب
١٤٢	١ - ٢ - ٧	نظرة عامة
١٤٣	٢ - ٧	تباين العيّنات بموجب المعاينة العشوائية البسيطة
١٤٧	٣ - ٧	قياسات أخرى لخطأ المعاينة
١٤٨	١ - ٣ - ٧	الخطأ المعياري
١٤٨	٢ - ٣ - ٧	مكافئ التباين
١٤٨	٣ - ٣ - ٧	أثر التصميم
١٤٨	٤ - ٧	حساب تباين العيّنات لتصاميم معيارية أخرى
١٤٩	١ - ٤ - ٧	المعاينة التطبيقية
١٥٢	٥ - ٧	المعلومات المشتركة لتصاميم وبيانات عيّنة استقصاء الأسر المعيشية
١٥٢	١ - ٥ - ٧	انحرافات تصاميم استقصاء الأسر المعيشية عن المعاينة العشوائية البسيطة
١٥٢	٢ - ٥ - ٧	إعداد ملفات البيانات لتحليلها
١٥٣	٣ - ٥ - ٧	أنواع تقديرات الاستقصاء
١٥٤	٦ - ٧	مبادئ توجيهية لتقديم معلومات عن أخطاء المعاينة
١٥٤	١ - ٦ - ٧	اتخاذ قرار بشأن ما يوضع في التقرير
١٥٥	٢ - ٦ - ٧	كيف تقدم تقارير عن معلومات أخطاء المعاينة
١٥٥	٣ - ٦ - ٧	قاعدة تقريبية في الإبلاغ عن الأخطاء المعيارية
١٥٦	٧ - ٧	طرق تقدير التباين في استقصاءات الأسر المعيشية
١٥٦	١ - ٧ - ٧	الطرق المضبوطة
١٥٦	٢ - ٧ - ٧	طريقة المجموعة النهائية
١٦٠	٣ - ٧ - ٧	تقريبات ضبط الخطية
١٦٢	٤ - ٧ - ٧	التكرار
١٦٤	٥ - ٧ - ٧	بعض تقنيات التكرار
١٦٩	٨ - ٧	مزالق استخدام البرامج الحاسوبية الإحصائية القياسية لتحليل بيانات استقصاءات الأسر المعيشية
١٧٠	٩ - ٧	برامج حاسوبية لتقدير أخطاء المعاينة
١٧٣	١٠ - ٧	مقارنة عامة للبرامج الحاسوبية
١٧٤	١١ - ٧	ملاحظات ختامية
١٧٥		المراجع ومواد أخرى للقراءة

## الفصل الثامن

## أخطاء غير أخطاء المعاينة في استقصاءات الأسر المعيشية

١٧٧.....	١ - ٨	مقدمة.....
١٧٨.....	٢ - ٨	التحيّز والخطأ المتغير.....
١٨٠.....	١ - ٢ - ٨	العنصر المتغير.....
١٨٠.....	٢ - ٢ - ٨	الخطأ المنهجي (التحيّز).....
١٨١.....	٣ - ٢ - ٨	تحيّز المعاينة.....
١٨١.....	٤ - ٢ - ٨	مزيد من المقارنة بين التحيّز والخطأ المتغير.....
١٨١.....	٣ - ٨	مصادر الخطأ غير أخطاء المعاينة.....
١٨٢.....	٤ - ٨	العناصر الني تتكوّن منها الأخطاء غير أخطاء المعاينة.....
١٨٢.....	١ - ٤ - ٨	خطأ التحديد.....
١٨٢.....	٢ - ٤ - ٨	خطأ التغطية أو الإطار.....
١٨٤.....	٣ - ٤ - ٨	عدم الاستجابة.....
١٨٦.....	٤ - ٤ - ٨	أخطاء القياس.....
١٨٧.....	٥ - ٤ - ٨	أخطاء الإعداد.....
١٨٧.....	٦ - ٤ - ٨	أخطاء التقدير.....
١٨٧.....	٥ - ٨	تقدير الأخطاء غير أخطاء المعاينة.....
١٨٧.....	١ - ٥ - ٨	مراجعات الاتساق.....
١٨٨.....	٢ - ٥ - ٨	فحص العيّنة/التحقّق منها.....
١٨٨.....	٣ - ٥ - ٨	فحوص ما بعد الاستقصاء أو فحوص لإعادة المقابلات.....
١٨٩.....	٤ - ٥ - ٨	تقنيات مراقبة الجودة.....
١٨٩.....	٥ - ٥ - ٨	دراسة أخطاء الذاكرة.....
١٩٠.....	٦ - ٥ - ٨	العيّنات الفرعية المتداخلة.....
١٩٠.....	٦ - ٨	ملاحظات ختامية.....
١٩٠.....		المراجع ومواد أخرى للقراءة.....

## الفصل التاسع

## إعداد البيانات لاستقصاءات الأسر المعيشية

١٩٣.....	١ - ٩	مقدمة.....
١٩٣.....	٢ - ٩	دورة استقصاء الأسر المعيشية.....
١٩٥.....	٣ - ٩	تخطيط الاستقصاء ونظام إعداد البيانات.....
١٩٥.....	١ - ٣ - ٩	أهداف الاستقصاء ومضمونه.....
١٩٥.....	٢ - ٣ - ٩	إجراءات الاستقصاء وأدواته.....
١٩٨.....	٣ - ٣ - ٩	تصميم نُظم إعداد البيانات في استقصاءات الأسر المعيشية.....
٢٠٢.....	٤ - ٩	عمليات الاستقصاء وإعداد البيانات.....

## الصفحة

٢٠٢.....	إنشاء الأطر وتصميم العيّنات	١ - ٤ - ٩
٢٠٤.....	جمع البيانات وإدارة البيانات	٢ - ٤ - ٩
٢٠٥.....	إعداد البيانات	٣ - ٤ - ٩
٢٢١.....	تذييل	٥ - ٩
٢٢١.....	مكتب مايكروسوفت ( <i>The Microsoft Office</i> )	١ - ٥ - ٩
٢٢٢.....	فيجيوال بيسك ( <i>Visual Basic</i> )	٢ - ٥ - ٩
٢٢٢.....	سنفار ( <i>CENVAR</i> )	٣ - ٥ - ٩
٢٢٢.....	بي سي كارب ( <i>PC CARP</i> )	٤ - ٥ - ٩
٢٢٣.....	نظام إعداد تعداد السكان والاستقصاء ( <i>CSPPro</i> )	٥ - ٥ - ٩
٢٢٣.....	حساب إحصاءات مفيدة لأخطاء المعاينة ووضع قائمة بها ( <i>CLUSTERS</i> )	٦ - ٥ - ٩
٢٢٣.....	النظام المتكامل لتحليل الاستقصاءات ( <i>ISSA</i> )	٧ - ٥ - ٩
٢٢٣.....	نظام التحليل الإحصائي ( <i>SAS</i> )	٨ - ٥ - ٩
٢٢٣.....	البرنامج الحاسوبي الإحصائي للعلوم الاجتماعية ( <i>SPSS</i> )	٩ - ٥ - ٩
٢٢٤.....	تحليل بيانات الاستقصاء ( <i>SUDAAN</i> )	١٠ - ٥ - ٩
٢٢٤.....	المراجع ومواد أخرى للقراءة	

## المرفق الأول

## أساسيات تصميم عيّنات الاستقصاء

٢٢٧.....	مقدمة	١ - ميم
٢٢٧.....	وحدات الاستقصاء ومفاهيمه	٢ - ميم
٢٢٩.....	تصميم العيّنة	٣ - ميم
٢٢٩.....	١ - ٢ - المتطلبات الأساسية لتصميم عيّنة احتمالية	ميم
٢٢٩.....	٢ - ٣ - أهمية المعاينة الاحتمالية في الاستقصاءات الكبيرة للأسر المعيشية	ميم
٢٢٩.....	٣ - ٣ - إجراءات الاختيار والتنفيذ والتقدير	ميم
٢٣٠.....	٤ - أساسيات استراتيجيات المعاينة الاحتمالية	ميم
٢٣٠.....	١ - ٤ - المعاينة العشوائية البسيطة	ميم
٢٣٣.....	٢ - ٤ - المعاينة المنهجية	ميم
٢٣٧.....	٣ - ٤ - المعاينة الطباقية	ميم
٢٤٣.....	٤ - ٤ - معاينة المجموعات	ميم

## المرفق الثاني

## قائمة بأسماء الخبراء

٢٤٧.....	الخبراء المشاركون في اجتماع فريق الخبراء التابع للأمم المتحدة لاستعراض مسودة الدليل المتعلق بتصميم استقصاءات الأسر المعيشية، نيويورك، ٣ - ٥ كانون الأول/ديسمبر ٢٠٠٥
----------	---

## الجدول

قائمة مصطلحات المعاينة ومصطلحات ذات صلة بالموضوع.....	٣١	١ - ٣
ملاحظات مختارة تُستخدَمُ قيماً للسكان وخصائص للعيّنات .....	٣٣	٢ - ٣
مقارنة عناصر التجميع لأثر التصميم بغية تنويع مُعاملات الارتباط داخل الطبقة $\delta$ وأحجام المجموعات $n$ .....	٥٦	٣ - ٣
الخطط البديلة للمعاينة: المرحلتان الأخيرتان من مراحل الاختيار.....	٦٧	٤ - ٣
فئات الاستجابة في الاستقصاء .....	١٢١	١ - ٦
مُعاملات الترجيح بموجب احتمالات الاختيار غير المتساوية.....	١٢٣	٢ - ٦
تعديل مُعاملات الترجيح تعويضاً عن عدم الاستجابة.....	١٣١	٣ - ٦
الترجيح فيما بعد التقسيم الطبقي لتعديل التغطية.....	١٣٤	٤ - ٦
مُعَلَمَات الطبقة للتباين.....	١٣٥	٥ - ٦
تخفيض مُعاملات الترجيح.....	١٣٧	٦ - ٦
النفقات الشهرية بالدولار على الغذاء لكل أسرة معيشية.....	١٤٣	١ - ٧
حساب تباين المعاينة الحقيقي لـ $\hat{\lambda}$ ، مَعَلَمَة المعدّل.....	١٤٤	٢ - ٧
التقديرات وتبايناتها لخصائص سكانية مختارة.....	١٤٦	٣ - ٧
نفقات الأسرة الأسبوعية على الغذاء وملكية جهاز تلفزيون للأسرة المدرجة في العيّنة.....	١٤٦	٤ - ٧
مثال بيانات لتصميم عيّنة طبقية.....	١٥٠	٥ - ٧
نسب الأطفال المحصنين الذين هم في سن المدرسة في ١٠ مناطق تعداد، باعتبارها المتغير الذي هو موضع اهتمام.....	١٥١	٦ - ٧
النفقات الأسبوعية للأسرة المعيشية على الغذاء، بحسب الطبقة.....	١٥٩	٧ - ٧
تنفيذ الخطوات في طريقة المجموعة النهائية لتقدير التباين.....	١٦٠	٨ - ٧
هيكل ملف البيانات لنهج التكرار.....	١٦٣	٩ - ٧
قيم العامل الثابت في معادلة التباين لمختلف تقنيات التكرار.....	١٦٥	١٠ - ٧
تطبيق طريقة المطواة لتقدير التباين على عيّنة صغيرة وعيّناتها الفرعية.....	١٦٦	١١ - ٧
العيّنة الكاملة: النفقات بحسب الطبقة.....	١٦٧	١٢ - ٧
طريقة المطواة (أسقط وحدة المعاينة الأولى ٢ من الطبقة ١).....	١٦٧	١٣ - ٧
التقديرات بناءً على أساس المكررات.....	١٦٨	١٤ - ٧
طريقة التكرار المعاد المتوازن (أسقط وحدة المعاينة الأولى ٢ من الطبقة ١ والطبقة ٢؛ ووحدة المعاينة الأولى ١ من الطبقة ٢).....	١٦٨	١٥ - ٧
استخدام برامج حاسوبية مختلفة لتقدير تباينات تقديرات الاستقصاء، مع نسبة النساء اللائي أبدين إيجابية مصل من بين النساء اللائي وضعن مواليدهن مؤخراً، بوروندي، ١٩٨٨ - ١٩٨٩.....	١٧٠	١٦ - ٧
تصنيف أخطاء الاستقصاءات.....	١٧٨	١ - ٨
مثال على مواضيع/ وحدات تحليل في استقصاء أُخذ من الاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي سنة ١٩٨٧.....	٢٠٠	١ - ٩
ملفات الأسر المعيشية والأفراد المستخدمة في الاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي في سنة ١٩٨٧.....	٢١٥	٢ - ٩
الملفات النمطية لاستقصاء ميزانيات الأسر المعيشية.....	٢١٥	٣ - ٩

٤ - ٩	قالب الملف المستوي كما استُخدم في ملف الأسرة المعيشية للاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي في سنة ١٩٨٧ .....	٢١٦
٥ - ٩	ملف الملاحظات وفيه البيانات النهائية لمتغيرات استقصاء الأسر المعيشية .....	٢١٩
١ - ميم	عدد المدارس بحسب عدد الموظفين .....	٢٤٢

### الأشكال

١ - ٢	جدول زمني لأنشطة استقصاء الأسر المعيشية في البلد "س" .....	١٢
٢ - ٢	عينة لصحيفة عمل التكاليف لبرنامج استقصاء أسر معيشية .....	١٥
١ - ٣	ترتيبات المناطق الإدارية للتقسيم الضمني إلى طبقات .....	٥٣
٢ - ٣	مثال للاختيار المنهجي للمجموعات، باحتمال متناسب مع الحجم .....	٦٤
١ - ٨	العلاقة بين أخطاء المعاينة وأخطاء غير المعاينة باعتبارها عناصر في خطأ الاستقصاء بأسره .....	١٧٨
٢ - ٨	مجموع خطأ الاستقصاء والعناصر التي يتكوّن منها .....	١٧٩
٣ - ٨	مجموع خطأ الاستقصاء المخفّض .....	١٧٩
١ - ٩	دورة استقصاء الأسر المعيشية .....	١٩٤
١ - ميم	المعاينة المنهجية الخطية (اختيار العينات) .....	٢٣٤
٢ - ميم	اختيار العينات بالمعاينة المنهجية الدائرية .....	٢٣٥
٣ - ميم	الاتجاه الخطي الرتيب .....	٢٣٦
٤ - ميم	التقلّبات الدورية .....	٢٣٧



# الفصل الأول

## مصادر البيانات للإحصاءات الاجتماعية والديمغرافية

### ١ - ١ مقدمة

١ - تشكّل استقصاءات الأسر المعيشية أحد المصادر الرئيسية الثلاثة للإحصاءات الاجتماعية والديمغرافية في كثير من البلدان. ومن المعترف به أيضاً أن تعداد السكان وتعداد المساكن هما أيضاً مصدر رئيسي للإحصاءات الاجتماعية، لكنهما يجريان عادة بعد فترات فاصلة طويلة تبلغ نحو ١٠ سنوات. ونُظِم التسجيل الإداري هي المصدر الثالث. غير أن هذا المصدر، في كثير من البلدان يبدو أفضل تطويراً في الإحصاءات الصحية والحيوية منه في الإحصاءات الاجتماعية. وتقدم استقصاءات الأسر المعيشية بديلاً أرخص لعمليات التعداد التي تهدف إلى جمع البيانات الحسنة التوقيت وبديلاً أكثر صلة وملاءمة لنظم التسجيل الإداري. وهي تُستخدَم لجمع بيانات اجتماعية - ديمغرافية مفصلة ومنوعة تتصل بالأحوال التي يعيش فيها الناس، وبحسن حالهم، والأنشطة التي يقومون بها، والخصائص الديمغرافية والعوامل الثقافية التي تؤثر في سلوكهم، وبالتغير الاجتماعي والاقتصادي. غير أن هذا لا يستثني الاستخدام التكميلي للبيانات التي تولدها استقصاءات الأسر المعيشية باستخدام بيانات من مصادر أخرى أيضاً، مثل عمليات التعداد والسجلات الإدارية.

### ١ - ٢ مصادر البيانات

٢ - المصادر الرئيسية الثلاثة المذكورة أعلاه للبيانات الاجتماعية والديمغرافية يمكن، إذا خُطت جيداً ونُفذت جيداً، وفي حالة نُظِم السجلات — إذا كانت منظمة جيداً، أن يكمل بعضها بعضاً في برنامج متكامل لجمع البيانات وتصنيفها. ومما يذكر أن الإحصاءات الاجتماعية والديمغرافية أساسية لبرامج التنمية الاجتماعية - الاقتصادية. وإن إحصاءات تكوين السكان بحسب العمر والجنس، بما في ذلك التوزيع الجغرافي للسكان، من أبسط البيانات اللازمة لوصف السكان و/أو المجموعات الفرعية من السكان. وتوفر هذه الخصائص السياق الذي يمكن أن تُدرَس فيه معلومات هامة أخرى عن الظواهر الاجتماعية، مثل التعليم، والإعاقة، ومشاركة القوة العاملة، والأحوال الصحية، وحالة التغذية، ووقوع الناس ضحايا للجرائم، والخصوبة، ونسبة الوفيات، والهجرة.

## ١ - ٢ - ١ استقصاءات الأسر المعيشية

٣ - أصبحت استقصاءات الأسر المعيشية بالعيّنة مصدراً رئيسياً للبيانات عن الظواهر الاجتماعية في الفترة الأخيرة التي تتراوح بين ٦٠ و ٧٠ سنة. وهي من بين أكثر أساليب جمع البيانات مرونةً. ويمكن، من الناحية النظرية، أن يُحقَّق في أيّ موضوع يرتكز على السكان بواسطة استقصاءات الأسر المعيشية. ومن الشائع استخدام الأسر المعيشية كوحدات للمعاينة في المرحلة الثانية في معظم استراتيجيات المعاينة على أساس المناطق الجغرافية (انظر الفصلين الثالث والرابع من هذا الدليل). ففي الاستقصاءات بالعيّنة يُختار جزء من السكان وتقدّم ملاحظات أو تُجمع بيانات عن ذلك الفرع؛ ثم تُطبق النتائج التي يُحصَلُ عليها على جميع السكان بالتناسب. ونظراً إلى أن عبء العمل في الاستطلاعات بالعيّنة يكون أخف على المقابلين (الأشخاص الذين يجرون المقابلات)، وكون الفترة المخصّصة لجمع البيانات أطول، يمكن تغطية معظم المواضيع المراد تغطيتها بتفصيل أكثر ممّا يحدث في عمليات تعداد السكان. يضاف إلى ذلك أنه، حيث إن عدد الموظفين الميدانيين اللازمين للعمل أقل، يمكن تعيين أشخاص أكثر تأهيلاً لتدريبهم تدريباً أكثر كثافة مما يحدث في عمليات التعداد. والحقيقة هي أنه لا يمكن تلبية كل احتياجات البلد من البيانات بواسطة إجراء عمليات التعداد؛ لذلك، توفر استقصاءات الأسر المعيشية آليةً لتلبية الاحتياجات الإضافية والناشئة على أساس مستمر. وإن مرونة استقصاءات الأسر المعيشية تجعلها خيارات ممتازة لتلبية احتياجات مستخدمي البيانات من المعلومات الإحصائية التي لولا هذه الاستقصاءات ما كانت ستتاح أو لكانت غير كافية.

### ١ - ٢ - ١ - ١ أنواع استقصاءات الأسر المعيشية

٤ - يوجد لدى كثير من البلدان برامج استقصاءات للأسر المعيشية تشمل استقصاءات دورية وأخرى مخصّصة. ويُنصح بأن يكون برنامج استقصاءات الأسر المعيشية جزءاً من نظام متكامل لجمع البيانات الإحصائية في البلد. وفي مجال الإحصاءات الاجتماعية والديمقراطية يمكن أن تشكّل استقصاءات الأسر المعيشية فيما بين عمليات التعداد جزءاً من هذا النظام.

٥ - كان برنامج القدرات الوطنية على إجراء استقصاءات الأسر المعيشية (NHSCP) مجهوداً كبيراً لمساعدة البلدان النامية على إنشاء القدرات الإحصائية والاستقصائية للحصول على المعلومات الاجتماعية - الاقتصادية والديمقراطية من قطاع الأسر المعيشية. وقد نُفِّدَ البرنامج على مدى نحو ١٤ سنة، من ١٩٧٩ حتى ١٩٩٢. وعندما انتهى تنفيذ البرنامج كان ٥٠ بلداً قد شارك في البرنامج. وكان إنجاز الرئيس تشجيع البلدان واعتمادها استقصاءات للأسر المعيشية متكاملة ومتعددة المواضيع. وبالإضافة إلى ذلك، شجع البرنامج على بناء قدرات إجراء استقصاءات بالعيّنة، لا سيما في البلدان الأفريقية.

٦ - توجد أنواع مختلفة من استقصاءات الأسر المعيشية يمكن إجراؤها لجمع البيانات عن الإحصاءات الاجتماعية والديمقراطية. وتشمل هذه الاستقصاءات استقصاءات مخصّصة، واستقصاءات متعددة المراحل، واستقصاءات متعددة المواضيع، واستقصاءات طويلة. ويتوقف اختيار النوع الملائم من الاستقصاءات على عدد من العوامل، من بينها المواضيع المطلوبة، والموارد، والاعتبارات السوقيّة.

٧ - تغطي الاستقصاءات المخصّصة مواضيع أو مسائل منفردة مثل استخدام الوقت، أو حالة التغذية. ويمكن أن تكون الاستقصاءات دورية أو مخصّصة.

٨ - تقتضي الاستقصاءات المتعددة المراحل جمع معلومات إحصائية على مراحل متتالية تكون كل مرحلة منها مقدمة للمرحلة التالية لها. وتشمل المرحلة الأولى عادة عينة أكبر مما تشملها المراحل التالية. وتتمثل وظيفتها في استعراض وحدات العينات فيما يتعلق بخصائص محددة لتقرير أهلية هذه الوحدات للاستخدام في المراحل اللاحقة. هذه الاستقصاءات طريقة فعّالة للتكلفة لتحديد السكان المستهدفين المراد جمع معلومات مفصلة منهم عن موضوع يكون موضع اهتمام في المراحل اللاحقة. وثمة مواضيع مثل الإعاقة واليتم من بين المواضيع المناسبة للدراسة باستخدام هذا النهج.

٩ - في الاستقصاءات المتعددة المواضيع تُغطّى مواضيع مختلفة في استقصاء واحد. وهذا النهج أكثر كفاءة كلفة بوجه عام من إجراء استقصاءات تشمل موضوعاً واحداً فقط.

١٠ - في الاستقصاءات الطولية تُجمع البيانات من وحدات العينات نفسها على مدى فترة من الزمن. ويمكن أن تكون الفترات الفاصلة بين عمليات جمع البيانات شهرية أو ربع سنوية أو سنوية. والقصد من إجراء استقصاءات من هذا النوع هو قياس التغيرات في بعض الخصائص لنفس السكان على مدى فترة من الزمن. والمشكلة الكبرى في هذا النوع من الاستقصاءات ارتفاع نسبة تناقص عدد المستجيبين. وتوجد ثمة مشكلة أخرى هي مشكلة أثر التكيّف مع الاستقصاء.

#### ١-٢-١ - ٢ - ١ - ٢ - ١ ميزات وأوجه قصور استقصاءات الأسر المعيشية مقارنة بعمليات التعداد

١١ - ليست استقصاءات الأسر المعيشية مُكلفة بقدر ما هو التعداد مُكلف، لكنها يمكن مع ذلك أن تكون باهظة التكلفة إذا ما لزم إنتاج النتائج بصورة مستقلة عن المناطق الإدارية الأقل تكلفة، مثل الأقاليم والمقاطعات. خلافاً للتعداد الذي يقتضي جمع البيانات لملايين الأسر المعيشية، يكون الاستقصاء بالعينة محدوداً بعينة من بضعة آلاف أسرة معيشية نتيجة لقيود التكاليف، مما يحدّ كثيراً من قدرته على إنتاج بيانات موثوقة لمناطق صغيرة. وسوف تُستطلع العلاقة بين حجم العينة وموثوقية البيانات لمناطق ومجالات صغيرة في الفصول التالية.

١٢ - فيما يتعلق بميزات استقصاء الأسر المعيشية مقارنة بالتعداد:

(أ) التكلفة الإجمالية للاستقصاء، كما ذكر أعلاه، أدنى بوجه عام من التكلفة الإجمالية للتعداد، لأن التعداد يحتاج إلى مقادير كبيرة من موارد القوى العاملة والموارد المالية والسوقية، والمادية. وإن العينة الاحتمالية المختارة والمستقصاة على الوجه الصحيح تعود بنتائج دقيقة وموثوقة يمكن أن تُستخدم أساساً للتوصل إلى استنتاجات بشأن مجموع السكان. نتيجة لذلك، لا توجد حاجة ماسة إلى إجراء تعداد للحصول على بعض التقديرات، مثل المعدل الإجمالي للخصوبة؛

(ب) تنتج استقصاءات بالعينة بوجه عام معلومات إحصائية أفضل نوعية لأنها، كما ذكر آنفاً، تسمح بإمكانية أكثر لتعيين أشخاص أفضل تأهيلاً وأفضل تدريباً على إجراء المقابلات. ومن الأسهل أيضاً توفير إشراف أفضل لأن المشرفين يكونون في العادة مدربين جيداً، وربما تكون نسبة المشرفين/المقابليين عالية تصل إلى نسبة: ١ إلى ٤. يُضاف إلى ذلك أن في الإمكان استخدام معدات تقنية أفضل لإجراء قياسات مادية في الاستقصاءات عندما يلزم إجراء هذه القياسات. في عمليات التعداد، تكون نوعية البيانات أقل جودة في بعض الحالات بسبب كبر حجم العملية. وهذا يؤدي إلى وقوع تقصير أو إهمال في ضمان الجودة في مراحل مختلفة، مما يسفر عن وقوع نسبة أخطاء عالية في غير أخطاء العينات؛

(ج) يوجد نطاق أوسع ومرونة أكثر في الاستقصاء بالعيّنة ممّا يوجد في التعداد من حيث عمق التحقيق وعدد البنود الواردة في الاستبيان. وربما يستحيل جمع معلومات ذات نوع أكثر تخصصاً في التعداد بسبب عدد التخصصات الذي يفوق القدرة أو بسبب مقدار المعدات اللازمة لإجراء الدراسة، الذي يفوق القدرة أيضاً. ولا يكون في الإمكان استخدام مُعاملات ترجيح لمقاييس الأغذية ومقاييس أخرى، مثلاً، في دراسة للتغذية. وكذلك لا يمكن إخضاع كل شخص من السكان لفحص طبي لتقرير عدد حالات الإصابة بفيروس نقص المناعة البشرية/الإيدز. علاوةً على ذلك، يمكن إضافة بنود إلى استقصاء الأسر المعيشية بالعيّنة تكون معقدة نسبياً إذا استُخدمت في التعداد.

١٣ - الاستقصاءات بالعيّنة أكثر مناسبة لجمع بيانات وطنية على مستوى موقع جغرافي كبير نسبياً حول مواضيع تحتاج إلى الاستطلاع بعمق، مثل الإعاقة المتعددة الجوانب، ونفقات الأسرة المعيشية، وأنشطة القوى العاملة، وضحايا الجرائم. وهذا يختلف عن عمليات التعداد التي تجمع معلومات عامة نسبياً تغطي مجالات صغيرة وتكون مصدراً لهذه المعلومات.

١٤ - بوجه عام، تشمل قوة العمليات الإحصائية لاستقصاءات الأسر المعيشية مرونة في أدوات جمع البيانات تكفي لاستيعاب عدد أكبر من الأسئلة حول أنواع شتى من المواضيع، وكذلك إمكانية تقدير مَعْلَمَات تَقَارَن بالمَعْلَمَات التي تُقاس في عمليات تعداد السكان والمساكن.

## ١ - ٢ - ٢ - ١ عمليات تعداد السكان والمساكن

١٥ - يشمل تعداد السكان، المشار إليه فيما يلي أدناه بكلمة التعداد، كامل عملية جمع وتصنيف وتقييم ونشر البيانات الديمغرافية والاجتماعية وغيرها، التي تغطي، في وقت معيّن، جميع الأشخاص في البلد أو في جزء محدد جيداً أو أجزاء محددة جيداً من البلد. وهو مصدر رئيسي للإحصاءات الاجتماعية، وفيه ميزة واضحة هي توفير بيانات موثوق بها — أي بما معناه بيانات غير متأثرة بأخطاء العيّنات — لوحدات جغرافية صغيرة. أما التعداد فهو وسيلة مثالية لتوفير معلومات عن حجم السكان وتركيباتهم وتوزعهم المكاني، إضافة إلى خصائصهم الاجتماعية - الاقتصادية والديمغرافية. ويجمع التعداد بوجه عام معلومات عن كل فرد في الأسرة المعيشية وعن كل مجموعة أماكن سكنية، في البلد كله عادةً، أو في أجزاء محدّدة جيداً من البلد.

## ١ - ٢ - ٢ - ١ الملامح الأساسية للتعداد التقليدي للسكان والمساكن

١٦ - تتميز عمليات التعداد التقليدية للسكان والمساكن بالخصائص التالية:

(أ) يجري تعداد الأفراد من السكان وكل مجموعة أماكن سكنية على حدة وتُسجَل خصائصهم كلاً على حدة؛

(ب) الهدف هو تغطية جميع السكان في إقليم محدد بوضوح. ويهدف إلى شمول كل شخص حاضر و/أو المقيمين المعتادين، وذلك يتوقف على مسألة ما إذا كان نوع تعداد السكان حسب الأمر الواقع أم حسب الوضع القانوني. وفي حالة عدم وجود سجلات شاملة للسكان أو سجلات إدارية تكون عمليات التعداد المصدر الوحيد لإحصاءات المناطق الصغيرة؛

(ج) يجري التعداد في جميع أنحاء البلد بوجه عام في الوقت نفسه ما أمكن ذلك. ويُعَدُّ جميع السكان والمساكن فيما يتعلق بالفترة المشار إليها نفسها؛

(د) تُجرى عمليات التعداد عادةً على فترات معيّنة. معظم البلدان تجري تعداداً كل ١٠ سنوات؛ وغيرها تجريه كل ٥ سنوات. وهذا يسهّل وجود معلومات قابلة للمقارنة في فترات محددة.

### ١-٢-٢ استخدام نتائج التعداد

١٧ - فيما يتعلق باستخدام نتائج التعداد:

(أ) توفر عمليات التعداد معلومات عن حجم السكان وتركيباتهم وتوزّعهم المكاني إلى جانب الخصائص الديمغرافية والاجتماعية؛

(ب) عمليات التعداد مصدر لإحصاءات المناطق الصغيرة؛

(ج) مناطق إجراء عمليات التعداد هي المصدر الرئيسي لأطر العيّنات لاستقصاءات الأسر المعيشية. وغالباً ما تُستخدم البيانات التي تُجمع في عمليات التعداد كمعلومات مساعدة في تقسيم العيّنات إلى طبقات وفي تحسين التقديرات في استقصاءات الأسر المعيشية.

### ١-٢-٣ أوجه القصور الرئيسية لعمليات التعداد

١٨ - يعتبر التعداد عادةً، بسبب تغطيته الجغرافية التي لا نظير لها، مصدراً لبيانات خط الأساس المتعلقة بخصائص السكان. لذلك، لا يمكن فيه تغطية مواضيع كثيرة بتفصيل كبير. وربما لا يكون التعداد أمثل مصدر للمعلومات المفصلة عن النشاط الاقتصادي، مثلاً. فتلك المعلومات تحتاج إلى أسئلة وتحقيقات مفصلة.

١٩ - لأنّ المقابلات التي تجرى في التعداد تعتمد اعتماداً كبيراً على ردود أشخاص غير الذين تتصل بهم المعلومات، فهي لا تعطي دائماً معلومات دقيقة عن خصائص لا يعرفها إلا الفرد المعني نفسه، مثل الوظيفة، وساعات العمل، والدخل، إلخ.

٢٠ - أُجريت عمليات تعداد سكان في كثير من البلدان خلال العقود القليلة الماضية. فقد أجرى نحو ١٨٤ بلداً ومنطقةً، مثلاً، تعداد سكان في دورة الـ ٢٠٠٠ (١٩٩٥ - ٢٠٠٤).

### ١-٢-٣ السجلات الإدارية

٢١ - يُجمَع كثير من الإحصاءات الاجتماعية من سجلات إدارية مختلفة كمنتجات جانبية للعمليات الإدارية. ومن الأمثلة على ذلك الإحصاءات الصحية التي تُجمَع من سجلات المستشفيات، وإحصاءات العمل من دوائر توفير الوظائف، والإحصاءات الحيوية من نظام السجل المدني، وإحصاءات التعليم من تقارير الالتحاق بالمدارس التي تقدمها وزارات التربية والتعليم. وتتوقف موثوقية الإحصاءات الواردة من السجلات الإدارية على مدى اكتمال السجلات الإدارية واتساق التعاريف والمفاهيم المستخدمة فيها.

٢٢ - يمكن أن تكون السجلات الإدارية مصادر ذات كفاءة عالية من حيث التكلفة لكن هذه النظم، مع ذلك، غير راسخة في معظم البلدان النامية. وهذا يعني ضمناً أن هذه المعلومات غير دقيقة في معظم الحالات. حتى إذا كانت عمليات السجلات الإدارية مستمرة لأغراض إدارية، يكون تجميع الإحصاءات، في معظم الحالات، موضع اهتمام ثانوي لمعظم المنظمات، ونتيجةً لذلك، يؤثر هذا في جودة البيانات. والمتطلبات الإحصائية، التي

ينبغي الوفاء بها، مثل توحيد المفاهيم والتعريف، وحسن التوقيت، وكمال التغطية، لا تؤخذ عادةً في الاعتبار ولا يلتزمُ بها.

٢٣ - في معظم البلدان، تكون المعلومات الواردة من السجلات الإدارية في أحيان كثيرة محدودةً من حيث المضمون، لأنها أكثر ما تكون فائدة للأغراض القانونية أو الإدارية. ونظم التسجيل المدني هي أمثلة على النظم الإدارية التي وضعت في بلدان كثيرة. لكن لم تكن كل البلدان ناجحة في هذا المجهود. البلدان التي توجد لديها نظم تسجيل حيوي كاملة تستطيع إنتاج تقارير دورية عن الأحداث الحيوية، مثل عدد الأطفال المولودين أحياناً بحسب نوع الجنس، وتاريخ الولادة ومكان الولادة، وعدد الوفيات بحسب العمر؛ ونوع الجنس، ومكان الوفاة، وسبب الوفاة، وحالات الزواج والطلاق، إلخ.

٢٤ - يتضمّن سجلّ السكان قواعد بيانات عن حياة كل شخص وكل أسرة معيشية في البلد. ويجري تحديث السجل على أساس مستمر عندما تحدث تغييرات في خصائص الفرد و/أو الأسرة المعيشية. وإذا ما اقترنت هذه السجلات مع سجلات اجتماعية أخرى يمكن أن تكون مصدراً غنياً للمعلومات. من بين البلدان التي وضعت نظماً من هذا القبيل الدانمرك، والنرويج، وهولندا، وألمانيا، والسويد. وفي معظم هذه البلدان تستند عمليات التعداد إلى نظام التسجيل.

٢٥ - في كثير من البلدان النامية، يمكن أن تكون السجلات الإدارية لمختلف البرامج الاجتماعية مصدراً للبيانات ذا كفاءة من حيث التكلفة، ومقترحاً جذاباً، لكنها ليست مطورة تطويراً جيداً. وغالباً ما تكون السجلات الإدارية محدودة في مضمونها وغير قابلة في العادة للتكييف كما تُكَيّف استقصاءات الأسر المعيشية، من حيث المفاهيم أو تفصيل المواضيع. وفي هذه الحالة، يشكّل استخدامها كمكتملة لمصادر أخرى تحدياً كبيراً، لقلّة المفاهيم الموحدة فيها، ونظم التصنيف، والتغطية الاختيارية، وقلّة التغطية.

## ١ - ٢ - ٤ أوجه التكامل بين المصادر الثلاثة للبيانات

٢٦ - لاحظ هذا الفصل طرقاً مختلفة يمكن بها استخدام عمليات التعداد والاستقصاءات ونظم التسجيل الإداري مجتمعة. ويدرس هذا الفرع موضوع الجمع بين معلومات آتية من مصادر بيانات مختلفة بطريقة تكاملية بمزيد من التفصيل. وما يثير الاهتمام في هذا المجال هو ضرورة الحد من تكاليف التعداد والاستقصاء وتخفيض عبء الإجابة على الأسئلة، وتوفير البيانات في مجالات على مستويات أدنى، ربما لا تغطيها بيانات الاستقصاء، مثلاً، وزيادة استعمال البيانات المتاحة في البلد إلى الحد الأقصى.

٢٧ - نظراً إلى أنه لا يمكن تكرار عمليات التعداد مراراً وتكراراً، توفر استقصاءات الأسر المعيشية أساساً لتحديث بعض معلومات التعداد، لا سيما على الصعيد الوطني وعلى مستوى مواقع أخرى ذات مساحة واسعة. وفي معظم الحالات، لا تُدرس في التعداد إلا المواضيع البسيطة نسبياً ويكون عدد الأسئلة في العادة محدوداً. لذلك، يمكن تكملة معلومات التعداد بمعلومات مفصلة حول مواضيع معقدة من استقصاءات الأسر المعيشية، حيث يُستفاد من صغر حجمها وقابليتها للمرونة.

٢٨ - كانت عمليات التعداد واستقصاءات الأسر المعيشية، في كثير من الحالات، يكمل بعضها بعضاً. وإن جمع المعلومات عن مواضيع إضافية أثناء التعداد من عيّنة للأسر المعيشية طريقة فعّالة لتكثيف نطاق التعداد بغية تلبية الطلبات المتزايدة على الإحصاءات الاجتماعية. ومن شأن استخدام أساليب و تقنيات المعاينة (أخذ

العينات) أن يمكن من إنتاج بيانات توجد حاجة ملحة إليها، بقدر مقبول من الدقة، عندما تجعل ضائقات الوقت والتكلفة من غير العملي الحصول على هذه البيانات بواسطة إجراء تعداد كامل.

٢٩ - يقدم التعداد أيضاً إطار معيّن، وهيكل إحصائي، وقدرة إحصائية، وإحصاءات قياسية تلزم في إجراء استقصاءات للأسر المعيشية. ومن الشائع وضع عيّنة للأسر المعيشية في سياق تعداد بغية جمع معلومات عن مواضيع أكثر تعقيداً، مثل الإعاقة، والوفيات النفاسية، والنشاط الاقتصادي، والخصوبة.

٣٠ - تدعم عمليات التعداد استقصاءات الأسر المعيشية بتقديمها أطراً معيّنة: يقدم التعداد قائمة صريحة بجميع وحدات المجالات، كمجالات التعداد، مثلاً، المستخدمة عموماً كوحدات المرحلة الأولى في عملية اختيار استقصاءات الأسر المعيشية بالعيّنة. علاوة على ذلك، يمكن أن تُستخدم بعض المعلومات المساعدة المتاحة من التعداد لتصميم الاستقصاءات بكفاءة. وكذلك يمكن أن تستخدم المعلومات المساعدة المستمدة من عمليات التعداد لتحسين تقديرات العينات بواسطة تقديرات التراجع والنسب، وبذلك تحسّن دقة تقديرات الاستقصاء.

٣١ - لتحقيق التكامل بين مصادر البيانات يلزم تعيين وحدات التعداد بوضوح واعتماد وحدات جغرافية متسقة في جمع الإحصاءات والإبلاغ عنها بواسطة مصادر مختلفة. يضاف إلى ذلك أن من الأساسي اعتماد تعريفات ومفاهيم وتصنيفات مشتركة من مختلف مصادر البيانات، بما في ذلك السجلات الإدارية.

٣٢ - يمكن أيضاً استخدام البيانات الواردة من استقصاءات الأسر المعيشية لمراجعة تغطية التعداد ومضمونه. والهدف من ذلك هو تقرير حجم الأخطاء واتجاهها. وقد استُخدمت الاستقصاءات اللاحقة للتعداد، مثلاً، لهذا الغرض أثناء دورة الـ ٢٠٠٠ من عمليات التعداد في زامبيا وكمبوديا لتقييم أخطاء التغطية. ويمكن كذلك استخدام بيانات التعداد لتقييم بعض نتائج الاستقصاء.

٣٣ - تقدير المناطق الصغيرة، الذي لقي اهتماماً كبيراً نتيجة لتزايد الطلب على تقديرات موثوقة للمناطق الصغيرة، مجال تُستخدم فيه البيانات الآتية من الاستقصاءات والسجلات الإدارية لإنتاج تقديرات متزامنة. ومما يُذكر أن المقدرين التقليديين المباشرين لمجالات محددة لا يعطون دقة كافية لأن أحجام العينات في المناطق الصغيرة نادراً ما تكون كبيرة إلى الحد الكافي. ويستند تقدير المناطق الصغيرة إلى سلسلة من التقنيات الإحصائية تُستخدم لإنتاج تقديرات للمناطق عندما تكون الاستقصاءات التقليدية لهذه المناطق غير موثوقة، أو لا يمكن حسابها. تشمل التقنيات نماذج توفر رابطاً مع مناطق صغيرة ذات صلة بواسطة بيانات تكميلية أو مساعدة منها، على سبيل المثال، بيانات تعداد حديث للسكان. لذلك، كان المجال الأساسي للإجراءات المتعلقة بالمناطق الصغيرة الاستعارة من مصادر بيانات مختلفة والجمع بين نقاط القوة في هذه المصادر سعياً إلى إنتاج تقديرات أكثر دقة وموثوقية.

٣٤ - في البلدان التي توجد فيها نُظم تسجيل مدني متطورة جداً يمكن استخدام البيانات المستمدة من عمليات التعداد والاستقصاءات بنجاح، مقرونة ببيانات مستمدة من السجلات الإدارية. ففي تعداد السكان الذي أُجري في سنغافورة في سنة ١٩٩٠، مثلاً، كان لدى المقابِلين استمارات مملوءة مسبقاً بمعلومات أساسية منقولة من السجلات الإدارية لكل فرد في الأسرة. وأدى هذا النهج إلى تخفيض الوقت اللازم لإجراء المقابلة ولتكاليف التعداد. ولما كان التعداد المستند إلى السجلات لا يعطي إلا مجموع عدد السكان وخصائصهم الأساسية فقط، تُجمَع الخصائص الاجتماعية - الاقتصادية المفصلة على أساس العينات.

٣٥ - يمكن استخدام البيانات المستمدة من السجلات الإدارية للتحقق من نتائج الاستقصاءات وعمليات التعداد وتقييمها. على سبيل المثال، في البلدان التي توجد فيها نُظم تسجيل حيوي كاملة، يمكن التحقق من بيانات الخصوبة ونسب الوفيات الواردة من عمليات التعداد بمقارنتها بالبيانات الآتية من نظام التسجيل.

## ١ - ٢ - ٥ ملاحظات ختامية

٣٦ - يمكن القول في الختام إنه يجب النظر إلى استقصاءات الأسر المعيشية وعمليات التعداد والمصادر الإدارية باعتبارها مكّلة بعضها لبعض. ويعني ذلك ضمناً أنه ينبغي، حيثما أمكن ذلك، استخدام مفاهيم وتعريفات مشتركة في تخطيط عمليات التعداد والاستقصاء. وينبغي أيضاً فحص الإجراءات الإدارية فحصاً دورياً لضمان استخدام المفاهيم والتعريفات المشتركة.

٣٧ - يجب أن يكون برنامج استقصاءات الأسر المعيشية جزءاً من نظام متكامل لجمع البيانات الإحصائية داخل البلد، يشمل عمليات التعداد والسجلات الإدارية أيضاً، لكي تتسنى تلبية الاحتياجات الإجمالية للإحصاءات الاجتماعية - الديمغرافية تلبية كافية.

## المراجع ومواد أخرى للقراءة

- Ambler, R., and others (2001). Combining unemployment benefits data and LFS to estimate ILO unemployment for small areas: an application of the modified Fay-Herriot method. Invited paper. International Statistical Institute session, Seoul.
- Banda, J. (2003). Current status of social statistics: an overview of issues and concerns. Presented at the Expert Group Meeting on Setting the Scope of Social Statistics organized by the United Nations Statistics Division in collaboration with the Siena Group on Social Statistics, New York, 6-9 May 2003.
- Bee-Geok, L., and K. Eng-Chuan (2001). ESA/STAT/AC.88/05 7 April. Combining survey and administrative data for Singapore's census of population 2000. Invited paper, International Statistical Institute session, Seoul.
- Kiregyera, B. (1999). *Sample Surveys: With Special Reference to Africa*. Kampala: PHIDAM Enterprises.
- Rao, J. N. K. (1999). Some recent advances in model-based small area estimation. *Survey Methodology* (Statistics Canada, Ottawa), vol.25, No.2, pp. 175-186.
- Singh, R. and N. Mangat (1996). *Elements of Survey Sampling*. Boston, Massachusetts: Kluwer Academic Publishers.
- Statistics Canada (2003). *Survey Methods and Practices*. Ottawa.
- United Nations (1982). *National Household Survey Capability Programme: Non-sampling errors in household surveys: Sources, Assessment and Control*. Preliminary version. DP/UN/INT-81-041/2., New York: United Nations Department of Technical Co-operation for Development and Statistical Office.
- \_\_\_\_\_ (1984). *Handbook of Household Surveys*, (Revised Edition). Studies in Methods No. 31 Sales No. E.83.XVII.13.
- \_\_\_\_\_ (1998). *Principles and Recommendations for Population and Housing Censuses*, Revision 1. Statistical Papers, No. 67/Rev.1 Sales No. E.98.XVII.8.
- \_\_\_\_\_ (2001). *Principles and Recommendations for a Vital Statistics System*, Revision 2. Sales No. E.01.XVII.10. ST/ESA/STAT/SER.M/19/Rev.2.
- \_\_\_\_\_ (2002). Technical report on collection of economic characteristics in population censuses. New York and Geneva: Statistics Division, Department of Social and Economic Affairs, and Bureau of Statistics, International Labour Office. ST/ESA/STAT/119. English only.

Whitfold, D., and J. Banda, (2001), Post enumeration surveys (PES's): are they worth it? Presented at the Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid-decade Assessment and Future Prospects, New York, 7-10 August organized by the Statistics Division, Department of Economic and Social Affairs, United Nations Secretariat. Symposium 2001/10. English only.



## الفصل الثاني

# تخطيط الاستقصاءات وتنفيذها

١ - بينما يَنْصَبُ تركيز هذا الدليل على جوانب المعاينة في استقصاءات الأسر المعيشية، من الضروري إلقاء نظرة عامة على تخطيط استقصاءات الأسر المعيشية وعملياتها وتنفيذها، بغية وضع الفصول والفروع المتعلقة بالمعاينة في سياقها الصحيح، يوجد الكثير من الكتب المدرسية والدليلية واليدوية التي تتناول موضوع تخطيط استقصاءات الأسر المعيشية وتنفيذها بتفصيل كثير، ويَحْتُ القارئ على الرجوع إليها للحصول على مزيد من المعلومات. غير أن كثيراً من النقاط الرئيسية أُبرِزت ووصفت وصفاً موجزاً في هذا الفصل، بما في ذلك ملامح رئيسية للتخطيط والتنفيذ، ما عدا تصاميم العينات واختيارها، التي تناقش في الفصلين الثالث والرابع والمرفق الأول.

### ٢ - ١ تخطيط الاستقصاءات

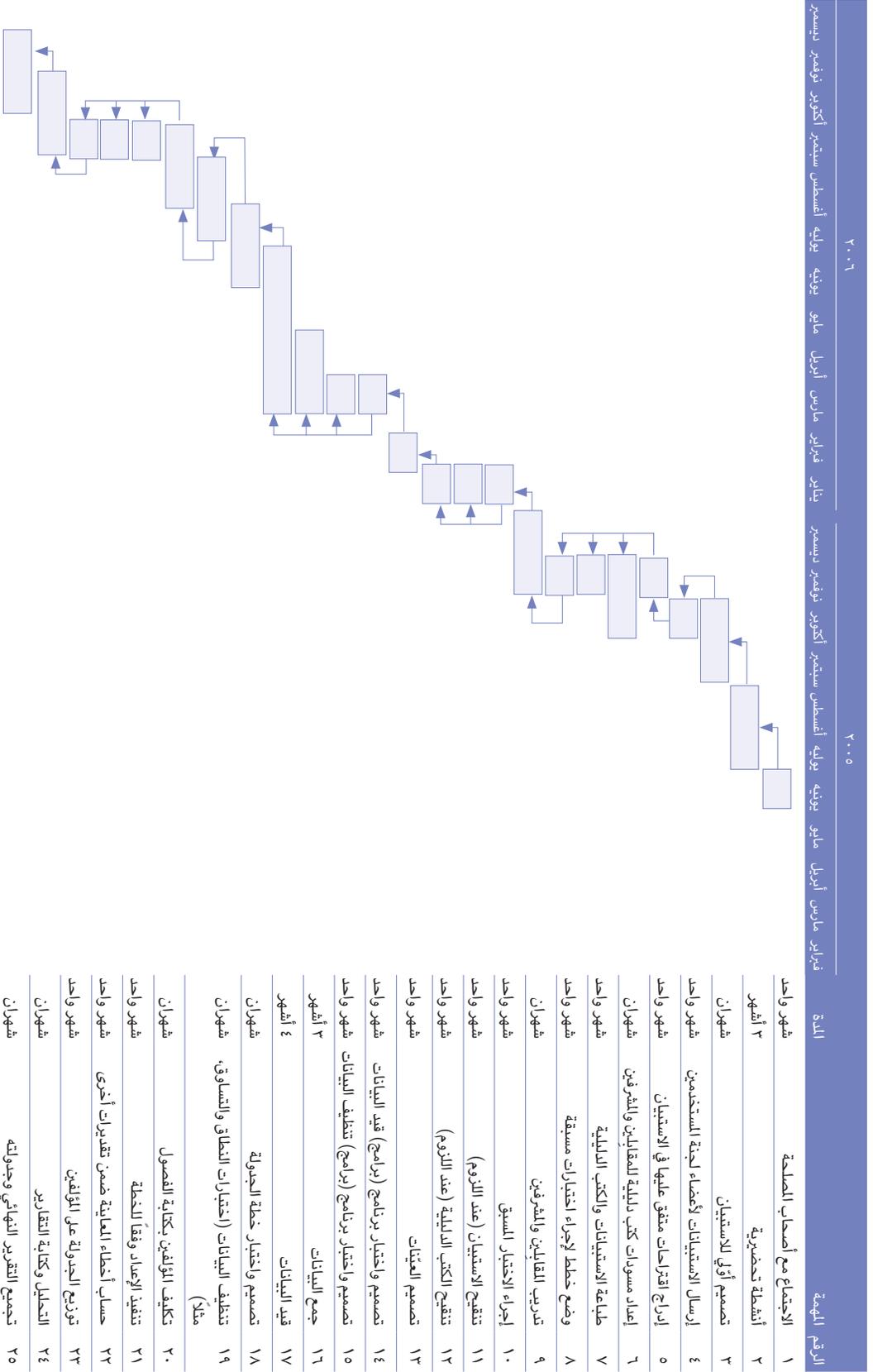
٢ - لكي يعود الاستقصاء بالنتائج المبتغاة يلزم إيلاء انتباه خاص للأعمال التحضيرية التي تسبق العمل الميداني. وفي هذا الصدد، تتطلب جميع الاستقصاءات أعمالاً تحضيرية تُجرى بعناية وحكمة إذا أُريد لها النجاح. غير أن مقدار التخطيط يتفاوت بحسب نوع الاستقصاء والمواد والمعلومات اللازمة له. ولما كان وضع خطة لاثقة للاستقصاء يتطلب وقتاً كافياً وموارد كافية (انظر الشكل ٢ - ١)، ليست دورة التخطيط البالغة سنتين غير شائعة في إجراء استقصاء مركب (للاطلاع على مناقشة تفصيلية لتخطيط الاستقصاء، انظر: الأمم المتحدة، ١٩٨٤).

يعرض الشكل ٢ - ١ التالي توضيحاً للجدول الزمني.

### ٢ - ١ - ١ أهداف الاستقصاء

٣ - لا بد من ذكر أهداف الاستقصاء بوضوح منذ بداية المشروع. ويجب أن يصدر بيان إحصائي مَصوغ صياغةً جيدة بشأن المعلومات المبتغاة، يعطي وصفاً واضحاً للتغطية السكانية والجغرافية. ومن الضروري أيضاً، في هذه المرحلة، اشتراط الكيفية التي ستُستخدم بها النتائج. ويجب أن يسترشد الخبير الإحصائي الذي سيجري الاستقصاء بالميزانية المحددة له لدى وضعه للأهداف. فمعرفة القيود على الميزانية تسهّل النجاح في تخطيط الاستقصاء وتنفيذه.

٤ - لا تُذكر أهداف الاستقصاء بوضوح في بعض الحالات. فقد يُطلب من منظمة استقصاءات أن تجري استقصاءً للأنشطة في القطاع غير النظامي. وإذا لم يُبيّن الغرض من الاستقصاء بوضوح يكون من شأن الخبير الإحصائي أو مدير الاستقصاء أن يعرف القطاع غير النظامي من حيث عملياته لأغراض إجراء الاستقصاء، ويبيّن



الشكل ١-٢ جدول زمني لأنشطة استقصاء الأسر المعيشية في البلد "س"

بالتفصيل الأنشطة الاقتصادية الخاصة التي هي أقرب ما تكون إلى تصوير متطلبات الجهة الممولة للاستقصاء. وينبغي أن يُذكر أنه يوجد احتمال كبير لأن تقع في الاستقصاء الذي تكون أهدافه غامضة وغير واضحة أخطاءً غير أخطاء العينات.

٥ - من الأهمية بمكان أن يُشرك أصحاب المصلحة، أي مختلف المستخدمين للإحصاءات ومنتجها في تعيين أهداف الاستقصاء وكذلك مداه وتغطيته. وتساعد المشاورات على الوصول إلى توافق في الآراء أو إلى حلول وسطى بشأن ماهية البيانات اللازمة، والشكل الذي يلزم أن تتخذه تلك البيانات، ومستويات تجزئة البيانات واستراتيجيات نشرها وتواتر جمع البيانات.

٦ - بعض الاستقصاءات التي تجريها منظمات متخصصة في إجراء الاستقصاءات تكون لها أهداف واضحة وجليّة. فالاستقصاء التجريبي الذي أجرته زامبيا للقوى العاملة في سنة ١٩٨٣، مثلاً، كانت له الأهداف التالية:

(أ) جمع معلومات عن حجم وتكوين القطاع السكاني العامل في القطاع النظامي؛

(ب) تقدير العرض والطلب على اليد العاملة؛

(ج) أن يُنحَدَّ أساساً للقيام بإسقاطات اليد العاملة في مهن معيّنة؛

(د) المساعدة في توسيع نطاق التعليم في الميادين الهامة جداً للتنمية الاقتصادية.

٧ - ينبغي ملاحظة أن وجود أهداف معلنة بوضوح هو الخطوة الأولى في تحديد أسئلة الاستقصاء التي يلزم تقديم أجوبة إحصائية عليها.

## ٢ - ١ - ٢ عالم الاستقصاء

٨ - من الضروري، لدى تخطيط استقصاء، تعيين المناطق الجغرافية التي سيغطيها والسكان الذين سوف يستهدفهم. ففي استقصاء إيرادات الأسر المعيشية ونفقاتها، مثلاً، يمكن أن يغطي الاستقصاء المناطق الحضرية وربما يستثنى المناطق الريفية.

٩ - في تعريف العالم، ينبغي تعيين السكان الذين ستؤخذ منهم العينات بالضبط. وفي استقصاء إيرادات الأسر المعيشية ونفقاتها سيكون عالم وحدات المرحلة الأولى مناطق التعداد (وهذه هي مدى انتشار وحدات المناطق الجغرافية، مثلاً، في كل أنحاء البلد) وستكون وحدات المرحلة الثانية أسراً معيشية في مناطق تعداد مختارة (يناقش الفصلان الثالث والرابع المجموعات بمزيد من التفصيل؛ انظر أيضاً المرفق الأول للاطلاع على تعريف المجموعة).

١٠ - غير أنه ينبغي الإشارة إلى أن عدد السكان المستهدفين، في الواقع، أصغر إلى حد ما من السكان الذين يشكّلون العالم. فقد جرت العادة بتقييد عدد السكان المستهدفين لعدد من الأسباب. في بعض الاستقصاءات، ربما تُستثنى من الاستقصاء أسر عسكرية معيّنة ساكنة في المعسكرات. وفي استطلاعات القوى العاملة ربما يُظهِر الأطفال الذين هم دون سن معيّنة بوصفهم أفراداً في الأسرة المشمولة بالاستقصاء، لكنهم ليسوا جزءاً من القوى العاملة.

١١ - من المهم أن نلاحظ أنه عندما يكون السكان الفعليون مختلفين عن السكان المستهدفين تنطبق النتائج على السكان الذين اختيرت منهم العينات على وجه التحديد. وكما نُوقش في الفصل الرابع، ينبغي وضع أُطر شاملة يُجَبُّ بعضها بعضاً لكل مرحلة من مراحل الاختيار.

## ٢ - ١ - ٣ المعلومات المراد جمعها

١٢ - يمكن أن توضع، من قائمة الأسئلة التي تحتاج إلى ردود إحصائية، قائمة بالبنود التي يمكن أن توفر معلومات وقائية تؤثر في المسائل التي هي قيد التحقيق. ومن المهم دائماً أن يؤخذ في الاعتبار أن بعض البيانات اللازمة ربما تكون متاحة من مصادر موجودة. وينبغي لدى وضع قائمة البنود اتخاذ ترتيبات لإدخال بنود تكميلية مرتبطة بالبنود الرئيسية. وفي استقصاء الوظائف والإيرادات، مثلاً، يمكن جمع معلومات إضافية عن السن ونوع الجنس والتعليم. ومن شأن هذه المعلومات أن تهيئ معرفة متعمقة إضافية بالمسائل ذات الصلة، ومن ثم تُثري التحليل.

١٣ - يمكن أن نضيف أنه يجب وضع خطة جدول في وقت تخطيط الاستقصاء. وينبغي نشر الجداول الفارغة للتعليق عليها وتحسينها.

## ٢ - ١ - ٤ ميزانية الاستقصاء

١٤ - تشير ميزانية الاستقصاء إلى المتطلبات المالية للاستقصاء المراد إجراؤه. الميزانية لازمة لدعم وإرشاد تنفيذ الاستقصاء ووضع الجدول الزمني لإنتاج نتائج الاستقصاء. ويجب أن تكون تقديرات التكاليف مفصلة قدر الإمكان. لذلك، من الضروري أن تُفهم جميع الخطوات المفصلة التي تنطوي عليها عملية الاستقصاء، وتبين الميزانية تكلفة الموظفين والمعدات وجميع بنود التكاليف الأخرى. وإذا كان ثمة سقف محدد مسبقاً للأموال المتاحة (وهذا ما يحدث في العادة)، يجب أن يكون إجمالي الميزانية في حدود الإطار المحدد. ويُنصح باتباع المبادئ التوجيهية العامة للوكالة الممولة في إعداد الميزانية. فربما يسهّل هذا الموافقة على تقديرات الميزانية. وإذا دعت الحاجة إلى الخروج عن الميزانية المقررة، ينبغي الحصول على إذن مسبق من المنظمة (أو المنظمات) ذات الصلة. وينبغي تحضير الطلبات المالية للاستقصاء في مرحلة مبكرة. وتتوقف الميزانية، بوجه عام، على تصميم الاستقصاء، والدقة المطلوبة في التغطية الجغرافية. يبيّن الشكل ٢ - ٢ أدناه صحيفة عمل ممكنة للتكاليف.

١٥ - من الأساسي وضع نظام فعّال لمراقبة التكاليف في المنظمة التي تُجري الاستقصاء. وفي معظم عمليات الاستقصاء الكبيرة الحجم توجد خطورة كبيرة من فقد السيطرة على رصد صرف الأموال بمجرد بدء العمل الميداني. وفي هذه الظروف، يميل مقدار كبير من الأموال إلى أن يُوجّه إلى مجالات لا صلة لها بعمليات الاستقصاء الرئيسية. وتساعد الحكمة في مراقبة التكاليف على رصد النفقات الفعلية بالنسبة إلى التكاليف المقدرة والعمل الفعلي المنجز. ومما لا بد منه أن تضمن الإدارة المسؤولة عن الاستقصاء المساءلة عن الأموال. فهذا يعزز إلى حد كبير مصداقية المنظمة المُجرية للاستقصاء.

## ٢ - ٢ تنفيذ الاستقصاءات

### ٢ - ٢ - ١ طرق جمع البيانات

١٦ - يوجد عدد من الأساليب المستخدمة في جمع البيانات، من بينها الملاحظة المباشرة والقياس؛ والاستبيان البريدي؛ والمقابلات الهاتفية والشخصية.

## الشكل ٢ - ٢

## عينة لصحيفة عمل التكاليف لبرنامج استقصاء أسر معيشية

مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	تكلفة الوحدة (وحدة العملة ذات الصلة لكل شخص ~ شهر	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	العملة ذات الصلة لكل شخص ~ شهر	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
	إلا حيث يُذكر غير ذلك (ذات الصلة)	إلا حيث يُذكر غير ذلك (ذات الصلة)
<b>أولاً - التخطيط والأنشطة التحضيرية</b>		
<b>ألف - التخطيط الأولي والرصد اللاحق (كبار الموظفين)</b>		
<b>باء - اختيار الموضوع وتحديد</b>		
١ - تخطيط الموضوع		
٢ - إعداد خطط الجدولة		
٣ - خدمات سكرتارية وخدمات أخرى		
<b>جيم - وضع الاستقصاء وتصميمه</b>		
١ - تخطيط التصميم الأولي: هيكل الاستقصاء، تغطية السكان، إجراءات أخذ العينات، طرق جمع البيانات، إلخ. (موظفون فنيون)		
٢ - تطوير مواد العينات:		
(أ) مواد الخرائط (على افتراض أن مواد التعداد موجودة):		
تكاليف الموظفين		
الخرائط واللوازم		
(ب) القوائم الميدانية للأسر المعيشية (٢٠٠٠ منطقة تعداد):		
تكاليف الموظفين (معظمها للمقابلين)		
تكاليف سفر		
(ج) اختيار العينات وإعداد القوائم الميدانية		
<b>دال - تصميم وطبع الاستبيانات وغيرها</b>		
١ - موظفون فنيون		
٢ - خدمات سكرتارية وخدمات أخرى		
٣ - تكاليف طباعة (بعد الاختبارات المسبقة)		
<b>هاء - الاختبار المسبق</b>		
١ - تخطيط من قبل موظفين فنيين:		
(أ) تحضيرات أولية		
(ب) تحليل النتائج ومراجعة المواد		
٢ - مشرف ميداني:		
(أ) تكاليف الموظفين		
(ب) تكاليف سفر		
٣ - مقابلون:		
(أ) تكاليف الموظفين		
(ب) تكاليف السفر		

## الشكل ٢ - ٢

## عيّنة لصحيفة عمل التكاليف لبرنامج استقصاء أسر معيشية (تابع)

مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	تكلفة الوحدة (وحدة العملة ذات الصلة) لكل شخص ~ شهر (إلا حيث يُذكر غير ذلك)	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
<b>واو - تحضير مواد التعليم والتدريب لاستخدامها في الميدان</b>		
		١ - الموظفون الفنيون
		٢ - الخدمات السكرتارية وخدمات أخرى
		٣ - تكاليف الاستنساخ
<b>زاي - أنشطة تخطيط متنوعة (العلاقات العامة والدعاية، مثلاً)</b>		
		١ - كبار الموظفين
		٢ - الموظفون الفنيون
		٣ - الموظفون التقنيون
		٤ - موظفو الخدمات
		٥ - تكاليف السفر
		٦ - الطباعة
		٧ - الخرائط ولوازم أخرى
<b>المجموع الفرعي</b>		
<b>ثانياً - العمليات الميدانية</b>		
<b>ألف - تدريب المشرفين الميدانيين</b>		
		١ - تكاليف الموظفين
		٢ - المأوى والمأكل
		٣ - تكاليف السفر
<b>باء - تدريب المقابليين</b>		
		١ - تكاليف المشرفين
		٢ - تكاليف المقابليين
		(أ) تكاليف الموظفين
		(ب) تكاليف السفر
<b>جيم - جمع البيانات (بما في ذلك مراقبة الجودة)</b>		
		١ - تكاليف المشرفين
		٢ - تكاليف المقابليين
		(أ) تكاليف الموظفين
		(ب) تكاليف السفر
<b>دال - الإدارة الميدانية</b>		
		١ - التوجيه الميداني
		٢ - تكاليف السفر
		٣ - تكاليف أخرى (مراقبة المواد وشحنها، مثلاً)
<b>هاء - عناصر المجموع الفرعي</b>		
		١ - الموظفون الفنيون

## الشكل ٢ - ٢

## عيّنة لصحيفة عمل التكاليف لبرنامج استقصاء أسر معيشية (يتبع)

مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	تكلفة الوحدة (وحدة العملة ذات الصلة) لكل شخص ~ شهر (إلا حيث يُذكر غير ذلك)	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
		٢ - الموظفون التقنيون
		٣ - موظفو الخدمات
		٤ - تكاليف السفر
		٥ - علاوة الإقامة أثناء السفر
		٦ - المقابلات
		٧ - تكاليف متفرقة
		المجموع الفرعي
		ثالثاً - إعداد البيانات
		ألف - تخطيط النظم
		باء - البرمجة الحاسوبية
		جيم - التشفير الكتابي
		١ - التشفير الأوّلي
		٢ - مراقبة الجودة
		٣ - الإشراف
		دال - العمليات من من لوحة المفاتيح إلى قرص الحاسوب
		١ - الطباعة الأوّلية
		٢ - مراقبة الجودة
		٣ - الإشراف
		هاء - وقت الحواسيب (بما في ذلك تكاليف المشغل والصيانة)
		واو - تكاليف التجهيز المتفرقة (اللوازم، إلخ.)
		زاي - عناصر المجموع الفرعي
		١ - الموظفون الفنيون
		٢ - الموظفون التقنيون
		٣ - موظفو مراقبة الجودة
		٤ - موظفو الخدمات
		٥ - عمل الحاسوب
		٦ - متفرقات
		المجموع الفرعي
		رابعاً - مراجعة البيانات ونشرها
		ألف - وقت الموظفين الفنيين
		باء - تكاليف النشر
		خامساً - توجيه وتنسيق الاستقصاء (مواصلة الإشراف على جميع الأنشطة)

الشكل ٢ - ٢

## عيّنة لصحيفة عمل التكاليف لبرنامج استقصاء أسر معيشية (انتهى)

مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	تكلفة الوحدة (وحدة العملة ذات الصلة)	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
مجموع التكلفة المقدرة (وحدة العملة ذات الصلة)	العملة ذات الصلة لكل شخص ~ شهر إلا حيث يُذكر غير ذلك	وحدات العمل المقدرة (شخص ~ أشهر إلا حيث يُذكر غير ذلك)
سادساً - المجموع الفرعي		
سابعاً - تقييم الدراسات والبحوث المنهجية (ربما تقدر بـ ١٠ في المائة من المجموع التراكمي)		
ثامناً - مصاريف إدارية عامة (يمكن تقديرها بـ ١٥ في المائة من المجموع التراكمي للتكاليف الإدارية، وإيجار أماكن العمل، واللوازم العامة، وما أشبه ذلك)		
تاسعاً - المجموع		

المصدر: الأمم المتحدة (١٩٨٤).

١٧ - الملاحظة المباشرة والقياس: تشكّل الملاحظة المباشرة والقياس الطريقة المثالية لأنهما في العادة أكثر موضوعية. فلا يُخشى من ضعف الذاكرة ولا الذاتية، سواء من قبل المستجيبين أم المقابِلين. ومن الأمثلة على المجالات التي استُخدمت فيها الملاحظة المباشرة ما يلي:

(أ) بعض جوانب استقصاءات استهلاك الأغذية؛

(ب) عمليات جمع الأسعار، حيث يمكن لعدّادين أن يشتروا المنتجات ويُسجّلوا أسعارها.

١٨ - في هذه الطريقة، وإن كانت مفيدة، عيب هو أنها باهظة التكلفة من حيث الموارد والوقت. وفي معظم الحالات، يحتاج المقابِلون إلى استخدام بعض المعدات. وقد دلت التجربة على أن طريقة الملاحظة المباشرة والقياس تميل إلى أن تكون مفيدة وعمليّة عندما تكون أحجام العيّنات أو السكان صغيرة نسبياً.

١٩ - الاستبيانات البريدية: طريقة استخدام الاستبيانات البريدية رخيصة وسريعة نسبياً. فعنصر التكلفة الرئيسي في مرحلة جمع البيانات هو طابع البريد. وعندما يُصمّم الاستبيان ويُطبع يرسل إلى المستجيبين (هؤلاء هم الأشخاص الذين يُتوقّع منهم أن يملأوا الاستبيان). ويُفترض في المستجيبين أن يكونوا مُلمّين بالقراءة والكتابة لأنه يُتوقّع منهم أن يملأوا الاستبيان هم أنفسهم دون مساعدة. غير أن هذا افتراض خاطئ، لا سيما في البلدان النامية، حيث ما زالت مستويات الإلمام بالقراءة والكتابة منخفضة. الضعف الرئيسي في هذه الطريقة هو ارتفاع نسب عدم الرد (أي أن نسبة الذين لا يملأون الاستبيان و/أو لا يردّون على بعض الأسئلة الواردة في الاستبيان كبيرة) المرتبطة به، وربما يعود ذلك إلى تعقّد الاستبيان المستخدم. غير أنه لا يمكن استبعاد اللامبالاة كلياً. وفي بعض الحالات، تكون نسبة الردّ على الاستبيان مرتفعة لكن نسبة عدم الردّ على بعض الأسئلة الموجودة فيه تكون مرتفعة.

٢٠ - ربما يلزم، لتحسين نسبة الردّ، إرسال رسائل تذكّرة لغير المستجيبين (الأشخاص الذين لا يردّون على أسئلة الاستقصاء). غير أنه يُنصح باختيار عدد من أفراد العيّنة غير المستجيبين وإجراء مقابلات شخصية معهم. وربما يكون هذا ضرورياً لأن خصائص الوحدات غير المستجيبة قد تكون مختلفة اختلافاً تاماً عن خصائص

الوحدات المستجيبة (انظر مناقشة التقسيم إلى طبقات وميزاته في الفصل الثالث والمرفق الأول). في هذه الحالة تُعالج الوحدات المستجيبة والوحدات غير المستجيبة باعتبارهما طبقتين مختلفتين/مجالين مختلفين لاحقين ينبغي إعطاؤهما ترجيحات مختلفة لدى إعداد التقديرات (تناقش مُعاملات الترجيح بمزيد من التفصيل في الفصول اللاحقة، لا سيما الفصل السادس). وينبغي، لزيادة نسبة الاستجابة أن يكون الاستبيان جذاباً، وقصيراً، وبسيطاً قدر الإمكان. وربما يساعد إرسال ظرف عليه طابع بريد وعنوان الجهة التي يُعاد إليها الاستبيان بعد ملئه لتحسين نسبة الاستجابة.

٢١ - لتحسين هذه الطريقة تحسيناً مرضياً يجب أيضاً أن يكون ثمة إطار للمعاينة حديث قدر الإمكان. لذلك ينبغي أن تكون عناوين المستجيبين المكتوبة على مغلف الاستبيان حديثة. ويجب أيضاً أن تكون المنظمة التي تجري الاستقصاء مقتنعة بقدرة المستجيبين على ملء الاستبيان دون مساعدة.

٢٢ - فيما يلي موجز لبعض الميزات وأوجه القصور للاستقصاءات المُجرّاة بالاستبيان البريدي:

الميزات:

- (أ) إنها أرخص؛
- (ب) يمكن نشر العيّنة على نطاق واسع؛
- (ج) إزالة تحيُّز الشخص المقابل؛
- (د) إنها سريعة.

أوجه القصور:

- (أ) نسبة عدم الاستجابة تكون عالية في العادة؛
- (ب) تؤخذ الردود على الأسئلة بقيمتها الاسمية ولأنه لا تتاح فرصة للتحقق منها؛
- (ج) في استقصاء المواقف يصعب التحقق ممّا إذا كان المستجيب قد رد على الأسئلة بمفرده دون مساعدة؛
- (د) تكون الطريقة مفيدة فقط في الحالات التي تكون فيها الاستبيانات بسيطة إلى حد ما؛ وهي لذلك ليست طريقة مناسبة لإجراء استقصاءات معقدة.

٢٣ - طريقة المقابلة الشخصية: هذه هي أكثر الطرق شيوعاً في البلدان النامية لجمع البيانات بواسطة استقصاءات بالعيّنة على نطاق واسع. بالإضافة إلى نسبة الاستجابة العالية التي تسفر عنها المقابلات الشخصية، هذه الطريقة مناسبة بسبب ارتفاع نسب الأمية في هذه البلدان. تستدعي هذه الطريقة إجراء مقابلات مع أشخاص مختارين لجمع المعلومات بطرح أسئلة عليهم. والميزة الرئيسية لهذه الطريقة أن المقابل يستطيع إقناع الشخص الذي يجري معه المقابلة بالرد على الأسئلة (بشيء من الحوافز)، ويمكنه أن يشرح له أهداف الاستقصاء. علاوةً على ذلك، عندما تُستخدم طريقة المقابلة الشخصية يوجد احتمال أكبر بجمع المعلومات الإحصائية عن بنود صعبة من حيث المفاهيم، يرجّح أن تكون الإجابة عليها غامضة لو أرسلت في استبيان بريدي.

٢٤ - غير أنه توجد بعض الحدود التي هي أصيلة في استخدام طريقة المقابلات الشخصية، مثل:

- (أ) يمكن للمقابلين المختلفين أن يفسروا الأسئلة تفسيرات مختلفة، وبذلك يُدخلون تحيزاً في نتائج الاستقصاء لأنّ المقابلين الذين يرجعون إلى دليل التعليمات باستمرار قليلون جداً؛
- (ب) في عملية التحقّق، قد يوحي بعض المقابلين بإجابات إلى المستجيبين؛
- (ج) ربما تؤثر خصائص المقابلين، مثل السن ونوع الجنس وأحياناً حتى العنصر، في مواقف المستجيبين؛
- (د) ربما يخطئ المقابلون في قراءة الأسئلة لأنّ انتباههم يكون موزّعاً بين المقابلة والتسجيل.
- ٢٥ - أوجه القصور الواردة أعلاه تشكّل، مجتمعة، ما يُسمّى بتحيز المقابل ويمكن أن تسبّب، كما أظهرت الدراسات، أخطاءً خطيرة في الاستقصاءات غير أخطاء المعاينة (ارجع إلى مناقشة الأخطاء غير أخطاء المعاينة في الفصل الثامن).

٢٦ - ينبغي أن تؤخذ النقاط التالية في الاعتبار لدى توجيه الأسئلة إلى المستجيبين:

- (أ) ينبغي أن يفهم المقابل بوضوح الغرض من كل سؤال كما يشرحه دليل المقابلين. ومن الأهمية بمكان أن يرجع المقابلون إلى الدليل باستمرار؛
- (ب) دلت التجربة على أن من الأفضل للمقابل أن تطرح الأسئلة بحسب تتابعها في الاستبيان. ففي معظم الاستبيانات، أوليت عناية كبيرة لترتيب الأسئلة، وأُخذ في الحسبان حفز المستجيب، والربط بين المواضيع، وتيسير تنشيط ذاكرة المستجيب ليتذكر الأحداث التي وقعت في الماضي، وأكثر الأسئلة حساسية؛
- (ج) ينبغي أن يمتنع المقابلون، ما أمكنهم ذلك، عن الإيحاء بأجوبة للمستجيبين؛
- (د) يجب توجيه جميع الأسئلة. فبهذه الطريقة تقل البنود التي لا يُجاب عليها إلى الحد الأدنى. علاوةً على ذلك، لا ينبغي ترك فراغ أمام أي بند في الاستبيان ما لم يكن في ذلك استجابة لنمط البنود التي تُترك خالية. وإذا كان ثمة سؤال لا صلة له بمستجيب معين ينبغي عندئذٍ وضع تعليق في الفراغ المذكور. وهذا النهج يطمئن مدير الاستقصاء إلى أن جميع الأسئلة الواردة في الاستبيان قد عُولجت على الوجه الصحيح.

## ٢ - ٢ - ٢ تصميم الاستبيان

٢٧ - حالما يتم تحديد أهداف الاستقصاء وخطة جداوله يمكن وضع الاستبيان ذي الصلة. فالاستبيان يؤدي دوراً مركزياً في عملية الاستقصاء تُنقل به المعلومات ممن يملكونها (المستجيبين) إلى من يحتاجون إليها (المستخدمين). فهو الأداة التي يُعبّرُ بها عن حاجة المستخدمين إلى المعلومات بطرق عملية وهو كذلك الأساس الرئيسي للمدخلات في نظام إعداد البيانات للاستقصاء المعني.

٢٨ - يحتاج حجم الاستبيان وشكله إلى النظر فيهما بمنتهى الجديّة. ويُنصح بتصميم الاستبيانات في وقت تخطيط الاستقصاء. وإذا أُريدَ إرسال الاستبيانات إلى المستجيبين بالبريد فيجب أن تكون جذابة وسهلة. فهذا يؤدي إلى زيادة نسبة الاستجابة. أما إذا أُريدَ استخدام الاستبيان لتسجيل ردود المستجيبين من قبل المقابلين في الميدان فيجب أن يكون صلباً بما فيه الكفاية ليتحمل المناولة.

٢٩ - الشيء المثالي هو أن يُصمَّم الاستبيان على نحو يسهل معه جمع بيانات ذات صلة ودقيقة. ولتحسين دقة بيانات الاستقصاء ينبغي إيلاء اعتبار خاص لترتيب بنود الاستبيان ولصياغته. وينبغي إعطاء حافز للمستجيب. وينبغي نشر الاستبيان جيداً لتسهيل قراءة الأسئلة إما من قبل المستجيب وإما من قبل المقابل. ولن نكون مُفْرِطِينَ في التأكيد، مهما أكدنا، أن كل استبيان يجب أن يحتوي على تعليمات واضحة.

٣٠ - لذلك يجب أن يبذل فريق الاستقصاء عناية خاصة لإعطاء تعاريف دقيقة للبيانات المراد جمعها ومواصفات دقيقة فيما يتعلق بترجمة البيانات المطلوبة والمفاهيم ذات الصلة إلى أسئلة عملية. وفي هذا الصدد، يصبح اختبار الاستبيان مسبقاً إجراءً عادياً، وبوجه عام، ضرورياً، ما لم تثبت صلاحية الاستبيان تماماً في استقصاءات سابقة.

٣١ - باختصار، يجب أن يتوفر في الاستبيان الجيد ما يلي:

(أ) أن يمكن من جمع معلومات دقيقة لتلبية احتياجات مستخدمي البيانات المحتملين بطريقة حسنة التوقيت؛

(ب) أن يسهل عمل جمع البيانات، وتجهيز البيانات وجدولتها؛

(ج) أن يضمن الاقتصاد في جمع البيانات، أي بمعنى تجنب جمع أي معلومات غير أساسية؛

(د) أن يسمح بتحليل شامل وذو معنى وباستخدام البيانات المجموعة لتحقيق مقاصدها.

٣٢ - ينطوي هذا ضمناً على أنه يجب تطوير استبيانات الاستقصاء على نحو تنتج معه معلومات من أفضل نوعية ممكنة مع التوكيد بوجه خاص على كونها ذات صلة وحسنة التوقيت ودقيقة. ولكي تُنَجَزَ هذه العملية بكفاءة يجب تخفيض تكلفة جمع المعلومات الضرورية والعبء الذي ينطوي عليه جمعها إلى الحد الأدنى.

## ٢-٢-١ وضع الأسئلة

٣٣ - تُستخدم أسئلة مفتوحة ومغلقة في استبيانات الاستقصاء بالعيَّة. ففي السؤال المفتوح يُطلب من المستجيب، رجلاً كان أو امرأة، أن يعطي جوابه هو نفسه على السؤال. أما في استقصاء المواقف فيمكن أن يُطلب منه تعريف ما يعتبره نوعية جيدة للحياة. ومن البديهي أن المستجيبين المختلفين سيعرّفون بطرقهم الخاصة ما يشكّل نوعية جيدة للحياة. ومن جهة أخرى، يقيّد السؤال المغلّق المستجيبَ باختيار جواب من قائمة أجوبة وضعها فريق الاستقصاء. وفيما يلي عينات من الأسئلة المغلقة:

هل توجد لديك أية إعاقة عقلية دائمة تحدُّ من أنشطتك اليومية؟

نعم  لا

كيف تقيّم قدرتك على الرؤية (حتى باستخدام النظارات أو العدسات اللاصقة، إذا استُخدمت)؟

١ -  لا أستطيع

٢ -  بصعوبة شديدة دائمة

٣ -  بشيء من الصعوبة الدائمة

٤ -  بدون صعوبة

٣٤ - ميزات الأسئلة المغلقة هي أنها: (أ) تعطي إجابات أكثر توحيداً؛ و (ب) سهلة التجهيز. القصور الرئيسي في هذه الأسئلة هو أن مُصمّم الاستقصاء هو الذي يجب أن يضع الأجوبة. ويمكن في هذه الحالة أن يغفل عن أجوبة هامة محتملة. وفي معظم الاستقصاءات، تكون أفضل طريقة لتناول المسائل والأسئلة المركّبة التي تتعلق بالمواقف والمفاهيم هي الأسئلة المفتوحة.

#### ٢-٢-٢-٢ صياغة الأسئلة

٣٥ - يجب أن تكون الأسئلة واضحة ودقيقة وغير غامضة. وينبغي ألا يُترك المستجيب (أو المستجيبة) ليخمن ما يريد المقابل أن يعلم منه/منها. ربما تبدو التعاريف والمفاهيم المستخدمة واضحة لمدير الاستقصاء، لكنها غير واضحة للمستجيب. ونتيجة لذلك ربما يستخدم المستجيب، رجلاً كان أم امرأة، رأيه الشخصي عند الإجابة على الأسئلة. ويمكن أن تكون النتيجة النهائية تفشي الأخطاء من غير أخطاء العيّنات. انظر مثلاً بسيطاً. السؤال، ما عنوان بيتك؟ يخلق تشويشاً في كثير من البلدان الأفريقية، لا سيما لسكان المناطق الحضرية، ما لم تكن كلمة "بيت" معرفة بوضوح. فكثير من المستجيبين يفسرون كلمة "بيت" بأنها القرية الأصلية التي جاؤوا منها.

#### ٢-٢-٢-٣ الأسئلة "الضمنية"

٣٦ - السؤال المسمّى بالضمني يقنع المستجيب بأن يرد عليه بطريقة معيّنة. وهذا يعني أن السؤال يكون متحيزاً إلى جواب معين. إليك سؤالاً ضمنياً في استقصاء صحي: كم يوماً في الأسبوع تشرب فيها أكثر من زجاجتي بيرة في اليوم؟ هذا السؤال يرغم المستجيب على الاعتراف بأنه يشرب البيرة، في الواقع، ما لا يقل عن زجاجتين في اليوم. هذه الأسئلة تميل إلى جعل المستجيب يعطي جواباً متحيزاً. من المهم أن نتجنب إنشاء البيانات: وإنما الهدف هو، ببساطة، جمع البيانات.

#### ٢-٢-٢-٤ صلة الأسئلة بالموضوع

٣٧ - الغرض من الاستبيان هو استخلاص معلومات تُستخدم في دراسة الوضع. لذلك، لا بد للمنظمة التي تجري الاستقصاء من أن تطرح أسئلة ذات صلة بالموضوع لكي تحصل على صورة صحيحة للوضع المعين الذي هو قيد الدرس. ويجب أن تكون الأسئلة الواردة في الاستبيان ذات صلة بمعظم المستجيبين. فلا معنى، مثلاً، لتقديم استبيان مليء بالأسئلة عن منجزات الفرد في التعليم العالي (الجامعي) في بيئة ريفية، وهي الشائعة، في معظم البلدان الأفريقية اليوم. كذلك، ليس من المناسب أن تشمل في استقصاء للخصوبة إناثاً في سن العاشرة أو دون ذلك وتساءل الواحدة منهن كم طفلاً ولدت، أو هل هي متزوجة، أم مطلقة، أم أرملة. فهذه الأسئلة تكون ذات صلة بإناث فوق سن معيّنة، لا للبنات اللاتي لم يبلغن سن الحمل والولادة بعد.

#### ٢-٢-٢-٥ تتابع الأسئلة

٣٨ - يجب أن تكون البنود في الاستبيان مرتبة على نحو تحفز معه المستجيب أو تسهل عليه أن يتذكّر وتساعد على طلب معلومات دقيقة منه. ويُقترح أن تكون الأسئلة الأولى سهلة ومثارة اهتمام وليست حسّاسة. فهذا يعزز ثقة المستجيب ويمكنه بالتالي من المضي قُدماً في المقابلة التي يقبل إجراؤها معه طائِعاً في معظم الحالات. وقد أصبح من المعتاد إلى حد ما أن تبدأ الأسئلة المعدة في استقصاءات الأسر المعيشية بأسئلة تلتمس تعريف الوحدة البسيطة، كالعنوان مثلاً، تليها أسئلة تلتمس وصفاً للأسرة المعيشية وأفرادها، وتشمل خصائص ديمغرافية على

سبيل المثال. وفي النهاية، تُسأل الأسئلة التفصيلية التي تشكّل الموضوع الرئيسي للاستقصاء<sup>١</sup>. وينبغي، عموماً، أن تكون الأسئلة الحساسة من بين الأسئلة التي تُسأل في الأخير. وينبغي أن نؤكد أنه يجب أن يكون ثمة صلة منطقية بين الأسئلة، لا سيما بين الأسئلة التي تكون محتملة أو متوقفة على شيء آخر.

## ٢ - ٢ - ٣ خطة الجدولة والتحليل

٣٩ - توجد تقنية مفيدة تساعد مصمّم الاستقصاء على توفير دقة أكثر في الأدوات التي يُقصد بها تلبية حاجة المستخدم إلى المعلومات، كما تتجلى في مجموعة الأسئلة أو الأهداف التي يشملها الاستقصاء. وتنطوي هذه التقنية على إنتاج خطط جدولة وجداول صورية. الجداول الصورية جداول مكتوبة تشمل كل شيء ما عدا البيانات الفعلية. وينبغي أن تبين الجدولة، على الأقل، عناوين الجداول وأرومات الأعمدة، وتعريف المتغيرات الموضوعية المراد جدولتها، ومتغيرات الخلفية المراد استخدامها للتصنيف، والمجموعات السكانية (موضوع الاستقصاء أو عناصره أو وحداته) التي تنطبق عليها الجداول المختلفة. ومن المستصوب أيضاً أن تُبيّن فئات التصنيف بأكبر تفصيل ممكن، وإن كان يمكن تعديل هذه الفئات حينما يُعرّف توزيع العيّنات على فئات المستجيبين، معرفة أفضل.

٤٠ - يمكن النظر إلى أهمية خطة الجدولة من عدد من المنظورات. فمن شأن الجداول الصورية، مثلاً، أن تبين إن كانت البيانات المراد جمعها سوف تؤدي إلى جداول قابلة للاستعمال. فهي لا تشير فقط إلى ما هو مفقود، وإنما تكشف أيضاً عما هو فائض عن الحاجة. علاوةً على ذلك، يُعوّض الوقت الزائد الذي يُصرف على وضع الجداول الصورية في العادة بأكثر منه في مراحل جدولة البيانات بتخفيض الوقت الذي يُصرف على تصميم الجداول الفعلية وإنتاجها.

٤١ - توجد أيضاً العلاقة الوثيقة التي يجب أن تؤخذ في الاعتبار بين خطة الجدولة وتصميم العيّنات المستخدمة في الاستقصاء. فالتقسيم الجغرافي للجداول، مثلاً، لا يمكن إجراؤه إلا إذا كانت العيّنة قد صُمّمت على نحو يسمح بهذه التجزئة. وكذلك حجم العيّنة أيضاً، يمكن أن يجعل من الضروري الحد من عدد الخلايا في الجداول المشتركة لتجنّب إنتاج جداول متباعدة أكثر من اللازم. وربما يلزم أحياناً تعديل الخطة أثناء عمل الجدولة. وربما تدعو الحاجة إلى تجميع الفئات بغية تقليل عدد الخانات الفارغة؛ أو ربما تتحقق نتائج طريفة في مسودة البيانات تدفع إلى إنتاج جداول جديدة. وسوف تُستخدم الطريقة التي تُجمع بها البيانات في استقصاءات الأسر المعيشية للرد على الأسئلة (تحقيق الأهداف)، وقد يُشار إليها بعبارة "خطة تحليل البيانات". وتفسّر هذه الخطة بالتفصيل ماهية البيانات اللازمة لتحقيق أهداف الاستقصاء. ويجب أن يرجع إليها مصمّم الاستقصاء باستمرار بينما هم يضعون تفاصيل استبيان الاستقصاء. وربما يكون من البديهي أن خطة التحليل يمكن أن تكون النقطة المرجعية الرئيسية التي يُستَرشدُ بها في تحليل نتائج الاستقصاء.

## ٢ - ٢ - ٤ تنفيذ العمل الميداني

٤٢ - تُقيّد شحة الموارد تنفيذ العمل الميداني غالباً في معظم البلدان النامية. غير أنه إذا أُريدَ إجراء استقصاء ينبغي تنظيم العمل الميداني وتنفيذه جيداً لكي تُستخدم الموارد المحدودة المتاحة لفريق الاستقصاء بكفاءة. ولكي تتجح عمليات الاستقصاء، يجب أن يفهم مصمّمو عمليات الاستقصاء الجوانب المفاهيمية لموضوع الاستقصاء

<sup>١</sup> انظر الأمم المتحدة، ١٩٨٤.

بوضوح. علاوةً على ذلك، يجب أن يُتقن المقابلون الإجراءات العملية التي ربما تؤدي إلى جمع البيانات الدقيقة بنجاح. ولكي يتم تحقيق عمليات الاستقصاء بنجاح يلزم دائماً أن يكون ثمة تنظيم ميداني فعّال ومنظم جيداً.

### ٢-٢-٤-١ المعدات والمواد

٤٣ - من الضروري في كثير من البلدان النامية أن تكون المعدات كالمركبات، والقوارب، والدراجات، إلخ. متاحة وتعمل جيداً قبل البدء بفترة كافية. ومن الضروري أيضاً أن تكون ثمة بعض قطع الغيار. فالمركبات والدراجات تيسر سرعة الحركة لقادة الأفرقة والمشرفين/المقابلين.

٤٤ - ينبغي أن توجد مواد كافية، مثل الملفات واللوحات وأقلام الرصاص وبرّيات أقلام الرصاص، ودفاتر الملاحظات والوقود (للمركبات)، بكميات كافية لاستخدامها أثناء عمليات الاستقصاء.

### ٢-٢-٤-٢ إدارة عمليات الاستقصاء

٤٥ - الاستقصاء بالعيّنة على نطاق واسع عملية معقدة في العادة وتتطلب براعة فائقة. لذلك، مهما كان التوكيد على الإدارة الحكيمة والفعّالة والكفاءة للأنشطة على مختلف المستويات لا يمكن وصفه بالمبالغة.

٤٦ - يجب أن يكون ثمة خط قيادة واضح ومُعرّف تعريفاً جيداً من مدير الاستقصاء حتى المقابلين. وينبغي ملاحظة أن استمارات الرقابة لرصد التقدم المحرز في الاستقصاء قد أثبتت فائدتها.

### ٢-٢-٤-٣ الدعاية

٤٧ - كان نجاح بعض الاستقصاءات محدوداً، ويعود جانب من السبب في ذلك إلى ارتفاع نسبة الأشخاص الذين لم يردّوا على الاستبيان لرفضهم المشاركة. لذلك يجب على منظمي الاستقصاء أن يقوموا بحملة دعائية للاستقصاء. وقد دلّت التجربة على أن الدعاية تؤدي دوراً هاماً في طلب تعاون المستجيبين، حتى وإن كانت بعض المنظمات/الوكالات الممولة تعتبر الإنفاق على الدعاية مضيعة للموارد.

٤٨ - يمكن اعتماد نُهج مختلفة للدعاية، وذلك يتوقف على الظروف السائدة. ففي المناطق الحضرية من بعض البلدان، مثلاً، يمكن للإذاعة والتلفزيون والصحف أن تبث رسائل تكمل الملصقات على الجدران، بينما في المناطق الريفية يمكن استخدام رسائل الإذاعة والملصقات.

٤٩ - ربما يكون من الضروري كذلك ترتيب عقد اجتماعات مع قادة الرأي المحليين في مناطق مختارة. ويجري في هذه الاجتماعات إعلام الحاضرين بأهداف الاستقصاء. وبالإضافة إلى ذلك، يُطلب من القادة إقناع الناس في منطقتهم بتقديم المعلومات التي يطلبها الأشخاص الذين يقابلونهم في عملية الاستقصاء.

٥٠ - من المهم قبل دخول الميدان أن يُنشرَ النص القانوني ذو الصلة بإجراء الاستقصاء. وينبغي أن يعطي الإعلان، ضمن معلومات أخرى، أهداف الاستقصاء ومدته والمواضيع التي سيشملها.

### ٢-٢-٤-٤ اختيار المقابلين

٥١ - الشخص المقابل هو صلة الوصل مع المستجيبين. وكون المقابل، رجلاً كان أم امرأة، ممثل منظمة الاستقصاء الذي يتصل دائماً مع المستجيب يبين بوضوح أهمية وظيفته البالغة لنجاح برنامج الاستقصاء. لذلك

ينبغي إيلاء اعتبار هام جداً وعناية جيدة لاختيار المقابل. وينبغي أن يكون المقابل قادراً على الاتصال بفعالية مع المستجيب. وينبغي أن تتوفر فيه الصفات اللازمة لأخذ جميع المعلومات بدقة في خلال وقت معقول.

٥٢ - ينبغي أن يكون لدى المقابل مستوى كاف من التعليم ويتوقف مستوى كفايته على نوع الاستقصاء. يضاف إلى ذلك أنه يجب أن يكون المقابل قادراً على تسجيل المعلومات بأمانة دون أن "يصطنع الأرقام". ويجب أن يتبع المقابلون المختارون التعليمات وأن يستخدموا التعاريف والمفاهيم الموجودة في الدليل الميداني للمقابلين.

٥٣ - ربما تساعد الإجراءات التالية على اختيار المقابلين المناسبين:

(أ) ينبغي أن يملأ المقابل المحتمل طلباً يبين فيه السن والوضع العائلي والعنوان الحالي والتحصيل التعليمي وتاريخ الوظائف التي شغلها؛

(ب) ربما يخضع الأشخاص الذين يقع عليهم الاختيار لفحص ذكاء وفحص إضافي في العمليات الحسابية البسيطة؛

(ج) نظراً إلى أنه يوجد في العادة حاجة إلى مقابلة المرشحين للتعين، بالإضافة إلى الفحص الكتابي، يجب أن يُجرى المقابلات مع المرشحين فريقاً يضع كل عضو فيه تقييمه المستقل لكل مرشح. ومن الصفات التي تؤخذ في الاعتبار لدى تقييم المرشحين دماثة الخلق، والاهتمام بالعمل، وحسن التعبير عن النفس، والفتنة.

٥٤ - العمل الميداني يمكن أن يكون متعباً، ففيه صعوبات في السفر في طرق وعرة؛ لذلك ينبغي أن يكون المقابل المختار ملتزماً ومستعداً للعمل في أحوال صعبة.

## ٢-٤-٥ تدريب المقابلين

٥٥ - ينبغي تدريب المقابلين المختارين تدريباً جيداً قبل إرسالهم إلى الميدان. والغرض الرئيسي من برنامج التدريب هو توحيد إجراءات المقابلة في الاستقصاء. وهذا أمر ضروري طبعاً لمنع المقابلين من إعطاء تفسيرات مختلفة للتعاريف والمفاهيم وأهداف الاستقصاء، ومن ثم تقليل تحيز المقابلين إلى الحد الأدنى.

٥٦ - ينبغي أن يتولى عملية التدريب مدربون مؤهلون. ويجب أن يكون هؤلاء المدربون على معرفة جيدة بمقاصد الاستقصاء وأهدافه. ويفضل أن يكونوا جزءاً من فريق الاستقصاء الذي يقوم بإجراء عملية الاستقصاء.

٥٧ - ينبغي تعليم المقابلين بعناية وتعريفهم بمقاصد الاستقصاء وكيف ستستخدم نتائجه. ولكي يعرف المقابلون أهداف الاستقصاء معرفة جيدة يجب أن يُدربوا تدريباً جيداً على معرفة المفاهيم والتعاريف الموجودة في الاستبيان.

٥٨ - يجب أن يتناوب المقابلون، كجزء من عملية التدريب، على شرح بنود الاستبيان لزملائهم أمام المدرب. فيمكن، مثلاً، أن يتناوب المقابلون على توجيه أسئلة بعضهم لبعض في جو الفصل الدراسي، ثم يؤخذون إلى الميدان في حي قريب ويقابل المقابلون المدربون بضع أسر. وينبغي أن يكون المدرب دائماً موجوداً كمرشد لتصويب المقابلين إذا أخطأوا. وبعد المقابلات الميدانية يناقش المدربون النتائج تحت إشراف المدرب. وينبغي أن يسفر برنامج التدريب عن قرار من مدير الاستقصاء بشأن أيّ المدربين يحتاجون إلى تدريب إضافي، وما إذا كان أيّ منهم لا يصلح بتاتاً لأداء المهمة.

## ٢ - ٤ - ٦ - ٢ - ٦ الإشراف الميداني

٥٩ - من المتفق عليه بوجه عام أن التدريب شرط أساسي للعمل الميداني الفعّال والناجح. غير أن التدريب من دون إشراف صحيح ربما لا يعود بالنتائج المرجّوة. ويتطلب العمل الميداني إشرافاً مكثراً وفعّالاً مستمراً من قبل موظفين كبار أكثر خبرة وأفضل تأهيلاً من المقابلين. وينبغي أن يتلقّى المشرفون تدريباً في جميع جوانب الاستقصاء. ولن نكون مبالغين في التوكيد، مهما أكدنا أن المشرف يشكّل حلقة وصل هامة بين المنظمة التي تجمع البيانات والمقابل. ويُفترض بالمشرف أن ينظم عمل المقابلين بتحديد المهام الميدانية والمواقع؛ ويستعرض العمل المنجز ويحافظ على مستوى عالٍ من الالتزام ببرنامج الاستقصاء بين المقابلين. ونوصي بأن تكون نسبة المشرفين إلى المقابلين مرتفعة نسبياً، إذا أمكن. واقترح أن تكون النسبة مشرفاً واحداً لكل أربعة مقابلين أو خمسة، فتلك هي النسبة المثالية في استقصاءات الأسر المعيشية. لكن هذا مجرد مبدأ توجيهي.

## ٢ - ٤ - ٧ - ٢ - ٧ متابعة الأشخاص غير المستجيبين

٦٠ - لا بد من أن توجد حالات، في معظم الاستقصاءات، لا يردّ فيها الناس على الاستبيان (ارجع إلى الفصل الثامن المتعلق بالأخطاء غير أخطاء العيّنات). فقد يرفض بعض المستجيبين التعاون مع المقابلين؛ وفي بعض الحالات، لا يهتمون ببند معيّن في الاستبيان. وعندما يُبلّغ المشرف عن وحدة غير متعاونة ينبغي أن يتصل بهذه الوحدة المشمولة بالعيّنة ويحاول الحصول على المعلومات منها، نظراً لما لديه من مؤهلات أفضل وخبرة أكثر. ولما كان الهدف العملي لأيّ استقصاء هو تحقيق أعلى نسبة استجابة ممكنة، يُوصى بجمع المعلومات من عيّنة فرعية من المجموعة الأولى من غير المستجيبين. ففي هذه الحالة، يُوجّه جهد الاستقصاء آنئذٍ نحو العيّنة الفرعية، ويُفضّل أن يقوم مشرفون بإجراء المقابلات .

## ٢ - ٤ - ٨ - ٢ - ٨ تقليل عدد حالات عدم الاستجابة

٦١ - من المهم، لدى تصميم استقصاء للأسر المعيشية وتنفيذه، وضع إجراءات استقصاء جيدة تهدف إلى الحصول على أقصى استجابة ممكنة. وإننا نوّكد أهمية تجهيز هذه الإجراءات لتقليل عدد حالات الرفض، مثل ترتيب أمر العودة إلى إجراء مقابلة في وقت يناسب المستجيب. وكذلك يجب شرح أهداف الاستقصاءات واستعمالاتها بعناية للمستجيبين الراضين بغية مساعدتهم على التعاون. ومن شأن ضمان السرية أن يساعد أيضاً على تخفيف ما يساور المستجيبين من مخاوف من أن تُستخدم إجاباتهم، لأغراض غير الأغراض المذكورة في الاستقصاء.

٦٢ - ينبغي معاودة الاتصال الهاتفي في أوقات مختلفة إذا لم يكن في البيت أحد. ويُوصى بمعاودة الاتصال ما لا يقل عن أربع مرات.

٦٣ - من المهم أيضاً تجنّب مشكلة عدم التمكن من الوصول إلى وحدات مختارة للعيّنة، يمكن أن تكون مصدراً هاماً لعدم الاستجابة. وأفضل طريقة لمعالجة هذه المشكلة هي استخدام أحدث إطار عيّنات (انظر الفصل الرابع للاطلاع على مناقشة مفصّلة).

## المراجع ومواد أخرى للقراءة

- Kiregyera, B. (1999). *Sample Surveys: With Special Reference to Africa*. Kampala: PHIDAM Enterprises.
- Statistics Canada (2003). *Survey Methods and Practices*. Ottawa.
- United Nations (1982). Preliminary version. National Household Survey Capability Programme: DP/UN/INT-81-041/2 non-sampling errors in household surveys: sources, assessment and control. New York:
- \_\_\_\_\_ (1984). *Handbook of Household Surveys*, (Revised Edition) Studies in Methods, No. 31. Sales No. E: 83.XVIII.13.
- \_\_\_\_\_ (1998). *Principles and Recommendations for Population and Housing Censuses*, Revision 1. Statistical Papers, No. 67/Rev.1. Sales No. E.98.XVII.8.
- \_\_\_\_\_ (2002). Technical report on collection of economic characteristics in population censuses. New York and Geneva: ST/ESA/STAT/119. English only. Statistics Division, Department of Social and Economic Affairs, and Bureau of Statistics, International Labour Office.
- Zanutto, E., and A. Zaslavsky (2002) Using administrative records to improve small area estimation: an example from the U.S. Decennial Census. *Journal of Official Statistic* (Statistics Sweden), vol. 18, No. 4, pp. 559-576.



## الفصل الثالث

# استراتيجيات المعاينة

### ٣ - ١ مقدمة

١ - بينما يلقي الفصل الثاني المتعلق بتخطيط الاستقصاء نظرة عامة على مختلف مراحل عمليات استقصاء الأسر المعيشية، هذا الفصل هو الأول من عدة فصول تركز فقط على جوانب المعاينة (أخذ العينات) — التي هي محور التركيز الرئيسي لهذا الدليل. يبحث هذا الفصل بإيجاز المعاينة الاحتمالية مقابل المعاينة غير الاحتمالية، ويقدم حججاً داعمة لسبب استخدام المعاينة الاحتمالية دائماً في استقصاءات الأسر المعيشية. يولى انتباه كبير لحجم العيّنة — المَعْلَمَات الكثيرة التي تحدده وكيفية حسابها. ونقدّم تقنيات بشأن تحقيق كفاءة المعاينة. تشمل هذه التقنيات التقسيم إلى طبقات، وأخذ عينات من المجموعات، وأخذ عينات على مراحل، مع التوكيد بوجه خاص على تصميم عينات على مرحلتين (انظر تعاريف هذه المفاهيم ومواصفاتها في الجدول ٣ - ١ والمرفق الأول). تعرض هنا خيارات معاينة مختلفة، ويرد وصف بالتفصيل لتصميمين رئيسيين للعينات استُخدما في بلدان كثيرة. ويُبحث أيضاً الموضوعان الخاصان (أ) المعاينة على مرحلتين للوصول إلى سكان "نادرين" و(ب) أخذ عينات لتقدير التغيير أو الاتجاه. وينتهي الفصل إلى عرض موجز بالتوصيات المقدمة.

### ٣ - ١ - ١ نظرة عامة

٢ - تكاد تصاميم العينات لاستقصاءات الأسر المعيشية كلها، في البلدان النامية والمتقدمة النمو على حدٍ سواء، تكون مُركَّبة بسبب مراحلها المتعددة وتقسيمها إلى طبقات وضمّها في مجموعات أو عناقيد. يضاف إلى ذلك، حقيقة أن استقصاءات الأسر المعيشية بالعيّنة على الصعيد القومي غالباً ما تكون عامة الغرض في نطاقها، وتشمل مواضيع متعددة تهم الحكومة، تضيف إلى تعقدها. لذلك، يركز الدليل على استراتيجيات معيّنة متعددة المراحل.

٣ - لتحقيق النتائج المرجوة، وهي تصميم جيد لاستقصاء الأسرة المعيشية، يجب أن يوضع ترتيب متناغم يجمع بصورة متوائمة عناصر متعددة. ويجب اختيار العيّنة على مراحل لكي تُحدّد المواقع التي ستُجرى فيها المقابلات وتُختار الأسر المعيشية بكفاءة. ويجب أن يُقسم التصميم إلى طبقات بطريقة يُضمن بها أن تُنشر العيّنة المختارة فعلاً نشراً صحيحاً على مناطق جغرافية فرعية ومجموعات فرعية من السكان. ويجب أن تستخدم خطة المعاينة مجموعات تكون وحداتٍ مُعرّفة جغرافياً تُختار منها الأسر، لكي تبقى التكاليف في حدود مستويات يمكن إدارتها. ويجب في الوقت نفسه، تجنّب الإفراط في تجميعها لأن الإفراط في خطة التجميع يترك آثاراً ضارة على الموثوقية (انظر مناقشة أثر التقسيم إلى مجموعات في الفرع ٣ - ٣ - ٥). ويجب أن يأخذ حجم العيّنة في الحساب الاحتياجات التنافسة بغية إيجاد توازن مثالي بين التكلفة والدقة. ويجب أيضاً أن يتناول حجم العيّنة الاحتياجات

الملحة للمستخدمين الذين يريدون بيانات لمجالات منها المجموعات الفرعية من السكان أو المناطق الفرعية. ويجب أن يلتزم تصميم العيّنة منتهى الدقة بطريقتين اثنتين مختلفتين هما: أولاً يجب أن يكون إطار العيّنة المستخدم (أو الموضوع) كاملاً، وصحيحاً، وحديثاً قدر الإمكان؛ ثانياً، يجب أن تُستخدم تقنيات لاختيار العيّنة تقلل إلى الحد الأدنى التحيز غير المقصود الذي يحدثه المنفذون أحياناً. ويجب أيضاً أن يكون التصميم قادراً على تقييم نفسه ذاتياً؛ بعبارة أخرى، يجب أن يكون التصميم على نحو يمكن معه تقدير أخطاء العيّنة ويجري تقديرها فعلاً، لتوجيه المستخدمين إلى قياس موثوقية النتائج الرئيسية. تنشأ أخطاء العيّنة من تقدير الخصائص السكانية على أساس بيانات واردة من جزء من السكان فقط لا من جميع السكان.

٤ - الغرض الرئيسي من الاستقصاء هو التمكن من الحصول على استنتاجات عن السكان المستهدفين، بناءً على عيّنة عشوائية. يهدف المستقصي/الباحث، وهو يفعل ذلك، إلى تقدير سمات غير معروفة للسكان. ومن بين الخصائص/المعلّيمات المشتركة بين السكان مجاميع، ومتوسطات، ونسب، وتباينات. فإذا كانت مثلاً،  $Y_1, Y_2, Y_3, \dots, Y_N$  قيماً للمتغير  $y$  في السكان،

$$(١ - ٣) \quad \bar{Y} = \frac{1}{N} \sum Y_i \quad \text{فإن متوسط السكان هو}$$

$$(٢ - ٣) \quad \sigma^2 = \frac{1}{N} \left( \sum Y^2 - N\bar{Y}^2 \right) \quad \text{ويكون متغير السكان هو}$$

تستخدم تقديرات العيّنة في معظم الحالات لتقدير معلّيمات السكان. مثال ذلك أنه إذا كان المتوسط والمتغير لحجم عيّنة مقداره  $n$  لعينة عشوائية بسيطة مع البديل قد أعطيا بهذا الشكل

$$(٣ - ٣) \quad \bar{y} = \frac{1}{n} \sum y_i$$

$$(٤ - ٣) \quad s^2 = \frac{1}{n-1} \left( \sum y_i^2 - n\bar{y}^2 \right)$$

وكانت  $y_1, y_2, y_3, \dots, y_n$  قيماً للمتغير لعدد من وحدات العيّنة مقداره  $n$ . في الاستقصاءات التي تُجرى بالعيّنة بحسب الباحث تباينات المتغيرات المختارة عشوائياً لتحديد خطأ العيّنة في التقدير (انظر تعريف خطأ العيّنة في الجدول ٣ - ١ ولزيادة البحث في خطأ العيّنة انظر الفصل السابع والمرفق الأول). من بين العوامل التي تؤثر في حجم تباين العيّنة اختلاف خواص المتغير الذي هو قيد الدراسة، وحجم العيّنة، وتصميم العيّنة (تُبحث هذه الجوانب في أجزاء مختلفة من هذا الفصل، والفصل السابع، وتُعرض المبادئ الأساسية للمعاينة للاستقصاء في المرفق الأول).

٥ - يبحث الفصلان الثالث والرابع بالتفصيل كل خاصية من الخصائص التي تدخل في تصميم العيّنة الصحيحة لاستقصاء الأسر المعيشية. يكون التركيز بوجه عام على الاستقصاءات الوطنية، وإن كان في الإمكان تطبيق جميع التقنيات الوارد وصفها هنا على استقصاءات كبيرة تُجرى على الصعيد دون الوطني، كالأستقصاءات التي تُقصر على واحدة أو أكثر من المناطق أو المحافظات أو المقاطعات أو المدن. ونظراً إلى الأهمية البالغة لأطر العيّنة في تحقيق ممارسة جيدة للمعاينة حُصص الفصل الرابع كله لهذا الموضوع.

## ٣ - ١ - ٢ قائمة مصطلحات المعاينة ومصطلحات ذات صلة بالموضوع

٦ - نبدأ بقائمة مصطلحات تُستخدم في هذا الفصل والفصل التالي (انظر الجدول ٣ - ١). لا يُقصد بهذه القائمة أن تعطي تعاريف رسمية لمصطلحات المعاينة، التي بعضُها مصطلحات رياضية. وإنما هي، بدلاً من ذلك، تصف استخدام المصطلحات في سياق هذا الدليل وتركز طبعاً على تطبيقات استقصاءات الأسر المعيشية.

الجدول ٣ - ١

## قائمة مصطلحات المعاينة ومصطلحات ذات صلة بالموضوع

المصطلح	الاستخدام
أثر التصميم (def)	نسبة تباين تصميم العينات المركب إلى معاينة عشوائية بسيطة لها نفس حجم العينة؛ يُشار إليه أحياناً باسم أثر التجميع، وإن كان أثر التجميع يشمل آثار التقسيم إلى طبقات وكذلك إلى مجموعات.
أخذ عينات من المناطق	اختيار وحدات مناطق جغرافية تشكّل إطار معاينة (يمكن أن يشمل قطاعات مناطق مختارة، معروفة بأنها تقسيمات فرعية من مناطق إدارية)
إطار (أطر) العينات	مجموعة مواد تُختار منها بالفعل عينة، كقائمة أو مجموعة مناطق، ومن هنا تُجمَع وحدات سكانية
الالتماس	طريقة لـ "تغطية" منطقة جغرافية بحثاً عن مساكن و/أو أسر معيشية تُطبَّق عادة في عمليات تُجرى لتحديث إطار عينات
الترابط بين عناصر الطبقة الواحدة	مُعامل العلاقة بين عناصر الطبقة الواحدة يقيس تشابُه العناصر في تجانسها
التزجيج الذاتي	تصميم عينة يكون لكل حالة فيها نفس الوزن من الاستقصاء
تصميم عينات مركب	يشير إلى استخدام مراحل متعددة ووضع عينات استقصاء الأسر المعيشية في مجموعات وفي طبقات، خلافاً للمعاينة العشوائية
تقسيم الأجزاء إلى أجزاء فرعية (قطع كبيرة)	عملية ميدانية في العادة تقسّم فيها المجموعة التي هي أكبر ممّا كان متوقّفاً إلى مجموعات فرعية لتقليل أعباء وضع القوائم
التقسيم إلى مجموعات؛ مُقسّم إلى مجموعات	يشير إلى ميل وحدات العينات — أشخاص أو أسر معيشية — إلى أن تكون لها خصائص مشتركة
التقسيم الضمني إلى طبقات	وسيلة التقسيم إلى طبقات بواسطة الفرز الجغرافي لإطار العينات، مقروناً بالأخذ المنهجي لعينات تكون الاحتمالية فيها متناسبة مع الحجم
حجم العينة	عدد الأسر المعيشية المختارة أو عدد الأشخاص المختارين
حجم المجموعة	(متوسط) عدد وحدات العينات — الأشخاص أو الأسر المعيشية — في المجموعة
خطأ غير أخطاء المعاينة	تحيزٌ في تقدير الاستقصاء ناشئ عن خطأ في التصميم والتنفيذ؛ يشير إلى دقة التقدير أو صحته خلافاً لموثوقيته أو إحكامه
خطأ المعاينة (الخطأ المعياري)	خطأ عشوائي في تقدير الاستقصاء نتيجة لحقيقة أن الاستقصاء شمل عينة واحدة ولم يشمل جميع السكان؛ الجذر التربيعي لمتغير المعاينة
الخطأ المعياري النسبي (مُعامل التباين)	الخطأ المعياري كنسبة مئوية من تقدير الاستقصاء، بعبارة أخرى خطأ معياري مقسوم على التقدير
الدقة (الصحة)	انظر الخطأ من غير أخطاء المعاينة
السكان المستهدفون	تعريف السكان المراد تغطيتهم بالاستقصاء؛ ويُعرفون أيضاً باسم عالم التغطية
العد السريع	يشير إلى تحديث العملية عندما تُعدّ المساكن تقريباً لإعطاء القياس الراهن للحجم؛ انظر أيضاً الالتماس
العينة الرئيسية	عينة "ممتازة" يُقصدُ بها أن تُستعمل في استقصاءات متعددة و/أو في جولات متعددة من الاستقصاء نفسه، ويكون ذلك عادةً في إطار زمني مدته عشر سنين

## الجدول ٣ - ١

## قائمة مصطلحات المعاينة ومصطلحات ذات صلة بالموضوع (تابع)

المصطلح	الاستخدام
العيّنة العشوائية البسيطة (SRS)	عيّنة عشوائية بسيطة (نادراً ما تستخدم في استقصاءات الأسر المعيشية)
الفرع	قسم فرعي مرسومٌ حدودُه وخريطته من مجموعة أكبر حجماً
الكسر من العيّنات (f)	نسبة حجم العيّنة إلى مجموع عدد الوحدات السكانية
متغير المعاينة	مربع الخطأ المعياري أو خطأ المعاينة
المجال	الوحدة الجغرافية التي يلزم وضع تقديرات مستقلة لها
مجموعة غير مترابطة	مجموعة عيّنات تتكوّن من أسر معيشية متفرقة جغرافياً
مجموعة مترابطة	مجموعة العيّنات المؤلفة من أسر معيشية متجاورة جغرافياً
مرحلة الاختيار الصوري	مرحلة صورية للاختيار، يُقصدُ بها تبسيط وظيفة الدليل في تعيين المناطق الفرعية التي ستوجد فيها مجموعات العيّنات في نهاية المطاف
مستوى الثقة	يصف درجة الثقة الإحصائية التي يُحصَلُ بها على الدقة أو هامش الخطأ حول تقدير الاستقصاء، وتعتبر نسبة ٩٥٪ هي المستوى القياسي بوجه عام
مُعامل الترجيح	عكس احتمالية الاختيار؛ عامل تضخّم يطبق على البيانات الخام؛ يُعرف أيضاً باسم مُعامل الترجيح في التصميم
معاينة احتمالية	منهجية اختيار تعرف بها كل وحدة سكانية (شخص، أسرة معيشية، إلخ.) أن لها فرصة غير صفرية لأن تُشمل بالعيّنة
المعاينة بناءً على الاتجاه	تصميم عيّنات لتقدير التغير الحاصل من فترة زمنية إلى أخرى
المعاينة الطباقية	تقنية لتنظيم إطار عيّنات في مجموعة فرعية تكون متجانسة داخلياً ومتنافرة خارجياً لضمان كون اختيار العيّنات موزعاً توزيعاً صحيحاً بين المجموعات الفرعية السكانية الهامة
المعاينة على فترات متعاقبة؛ يُعرف أيضاً بالمعاينة المزدوجة أو المعاينة بعد التقسيم إلى طبقات	اختيار العيّنة في فترتين زمنيتين (بوجه عام)، وتكون عيّنة الفترة الثانية عادةً عيّنة فرعية من عيّنة الفترة الأولى؛ ينبغي عدم الخلط بينها وبين عيّنة الاتجاه (انظر أدناه)
المعاينة على مراحل	يعني الطريقة التي تُختار بها عيّنة المناطق الإدارية والأسر المعيشية/الأشخاص في مرحلتين متتاليتين لتحديد المواقع الجغرافية التي يُجرى فيها الاستقصاء
معاينة غير احتمالية	انظر في الفرع ٣ - ٢ - ٢ أوصافاً نصّية لعيّنات هذه الطريقة: اختيار بحسب نظام حصص؛ اختيار بحكم قضائي، اختيار هادف، اختيار بحسب المصلحة، اختيار عشوائي أثناء المشي
معاينة القوائم	اختيار الوحدات التي تشكّل الإطار من قائمة
معاينة المجموعات	عملية أخذ عيّنات تنطوي فيها المرحلة قبل الأخيرة على وحدة معرفة جغرافياً كمجموعة تعداد
المعاينة مع تساوي الاحتمالات (EPSEM)	معاينة مع تساوي الاحتمالات
المعاينة مع تناسُب الاحتمال مع الحجم	اختيار وحدات المرحلة الأولى (الثانية، إلخ.) تُختار كل وحدة منها مع تناسُب احتمالها مع مقياس حجمها؛ انظر أيضاً المعاينة مع تناسُب الاحتمال مع الحجم المقدّر في النص
المعاينة المنهجية	اختيار العيّنات من قائمة باستخدام نقطة بداية عشوائية وفترة فاصلة بين الأرقام المختارة، محددة سلفاً وتطبق بصورة متتابعة
مقياس الحجم	في المعاينة متعددة المراحل، عدُّ أو تقدير حجم (عدد الأشخاص، مثلاً) كل وحدة في مرحلة معيّنة
المقدّر	المقدّر لتصميم عيّنات محدد هو طريقة تقدير المَعْلَمَات السكانية من بيانات العيّنات، فالتوسط الحسابي للعيّنات مُقدّر
الموثوقية (الدقة، هامش الخطأ)	يشير إلى درجة من خطأ معاينة مرتبط بتقدير استقصاء معين
وحدة المعاينة الأولى	وحدة إدارية معرفة جغرافياً اختيرت في المرحلة الأولى من مراحل المعاينة

## ٣-١-٣ ملاحظات

٧ - تستخدم الملاحظات المعيارية في هذا الفصل والفصول اللاحقة (انظر الجدول ٣ - ٢ أدناه). تشير الحروف المكتوبة بالحرف الكبير، بوجه عام، إلى القيم السكانية، والحروف المكتوبة بالحرف الصغير إلى ملاحظات على العينات. فبالعبارة،  $\bar{Y}$  مثلاً تشير إلى القيم السكانية، بينما تُستخدم العبارة،  $\bar{y}$  بوجه عام، لبيان قيم العينات. ويتضح ممّا ورد أعلاه أن الحرف  $N$  يشير إلى حجم السكان، بينما يشير الحرف  $n$  إلى حجم العينة. ومن الأهمية بمكان أن نلاحظ أن المَعْلَمَات السكانية يُعبّرُ عنها إما بحروف لاتينية مكتوبة بالحرف الكبير وإما بحروف يونانية. ففي الجدول ٣ - ٢ يبين الحرف  $\bar{Y}$  والحرف  $\sigma$ ، مثلاً، المتوسط السكاني والانحراف المعياري، على التوالي. وتمثل أجهزة تقدير المَعْلَمَات السكانية في الملاحظات على العينة بالحرف الكبير بإضافة الرمز  $\wedge$  إلى المَعْلَمَات السكانية.

الجدول ٣ - ٢

## ملاحظات مختارة تُستخدَمُ قيماً للسكان وخصائص للعينات

ملاحظة لـ				
العينة				
الخاصية	السكان	تقديرات مع الرمز " $\wedge$ " فوق الملاحظة	استخدام حرف صغير في كتابة الملاحظة	العينة
الوحدات	$N$	$n$	$n$	
ملاحظات	$\hat{Y}_1 \hat{Y}_2 \dots \hat{Y}_p \dots \hat{Y}_n$	$\hat{Y}_1 \hat{Y}_2 \dots \hat{Y}_p \dots \hat{Y}_n$	$Y_1, Y_2, \dots, Y_p, \dots, Y_n$	
متوسط القيمة	$\bar{Y}$	$\hat{Y}$	$\bar{y}$	
النسبة	$P$	$\hat{P}$	$p$	
المَعْلَمَات والمقدّر	$\theta$	$\hat{\theta}$	$\hat{\theta}$	
تباين $y$	$\sigma^2(y)$	$\hat{\sigma}^2(y)$	$s^2(y)$	
الانحراف المعياري لـ $y$	$\sigma(y)$	$\hat{\sigma}(y)$	$s(y)$	
الترابط بين عناصر الطبقة الواحدة	$\delta$	$\hat{\delta}$	$\hat{\delta}$	
النسبة	$R$	$\hat{R}$	$r$	
حاصل الجمع	$\sum_{i=1}^N$	$\sum_{i=1}^n$	$\sum_{i=1}^n$	

## ٣ - ٢ العيّنات الاحتمالية مقابل الطرق الأخرى للمعاينة في استقصاءات الأسر المعيشية

٨ - مع أن مناقشة نظرية الاحتمالية تقع خارج نطاق هذا الدليل، فمن الأهمية بمكان توضيح السبب الذي لأجله تؤدي طرق الاحتمالية دوراً لا يُستغنى عنه في المعاينة لاستقصاء الأسر المعيشية. يرد في هذا الفرع تعريف ووصف موجزان للمعاينة الاحتمالية، وأسباب أهميتها. ونذكر طرق أخرى، كالمعاينة لأسباب قضائية أو هادفة، والمعاينة العشوائية أثناء المشي، والمعاينة بنظام الحصص، والمعاينة بداعي الملاءمة، لا تفي بشروط الاحتمالية، وتُبحث أسباب عدم التوصية باستخدامها في استقصاءات الأسر المعيشية.

### ٣ - ٢ - ١ المعاينة الاحتمالية

٩ - المعاينة الاحتمالية في سياق استقصاء الأسر المعيشية يضمّ الوسائل التي تُختار بها عناصر السكان المستهدفين — الوحدات الجغرافية، والأسر المعيشية، والأفراد — لشمولها بالاستقصاء. تقتضي المعاينة الاحتمالية ما يلي: (أ) كل عنصر يوجد لديه فرصة حسابية معروفة لاختياره (ب) فرصة أن تكون أكثر من صفر، (ج) فرصة لأن يُحسب مقداره بالأرقام. ومن الأهمية بمكان ملاحظة أنه لا يلزم أن تكون فرص اختيار كل العناصر متساوية، وإنما يمكن أن تتباين وفقاً لأهداف الاستقصاء.

١٠ - إنّ الطبيعة الرياضية للمعاينة الاحتمالية هي التي تسمح بأخذ تقديرات العيّنات الاحتمالية القائمة على أساس علمي من الاستقصاء. وأهم من ذلك أن هذا هو أساس استنتاج أن تقديرات العيّنات تمثل مجموع السكان الذين أخذت منهم العيّنات. وإن استنتاج أنه يمكن تقدير أخطاء المعاينة من البيانات التي جُمعت من حالات العيّنات منتج ثانوي هام جداً للمعاينة الاحتمالية في الاستقصاءات. وليس أيّ من هذه الجوانب خصائص لطرق المعاينة غير الاحتمالية. وبسبب هذه الجوانب، يوصى بقوة بأن تُستخدم المعاينة الاحتمالية دائماً في استقصاءات الأسر المعيشية، حتى عندما تكون تكاليف الاستقصاء أعلى من تكاليف الطرق غير العلمية وغير الاحتمالية.

### ٣ - ٢ - ١ - ١ المعاينة الاحتمالية على مراحل

١١ - يجب استخدام المعاينة الاحتمالية، كما ذكر ضمناً أعلاه، في كل مرحلة من مراحل عملية اختيار العيّنات بغية تلبية المتطلبات. فالمرحلة الأولى من الاختيار، مثلاً، تنطوي بوجه عام على اختيار وحدات معينة جغرافياً كالقرى. وتنطوي المرحلة الأخيرة على اختيار أسر معيشية محددة أو أشخاص محددين لمقابلتهم. وينبغي أن تستخدم هاتان المرحلتان وأيّ مراحل موجودة بينهما طرقاً احتمالية لأخذ عيّنات صحيحة. يرد أدناه مثال توضيحي مبسّط.

#### مثال

افترض أن عيّنة عشوائية بسيطة تضم ١٠ قرى من ١٠٠ قرية في منطقة ريفية. وافترض أيضاً أنه وُضعت قائمة بأسماء الأسر المعيشية في كل قرية من قرى العيّنة. وأجري اختيار منهجي من القائمة بواقع أسرة واحدة من كل خمس أسر لإجراء مقابلة معها من أجل الاستقصاء، بغض النظر عن عدد الأسر الواردة أسماؤها في قائمة كل قرية. هذا تصميم عيّنات احتمالية، حيث يجري الاختيار في مرحلتين مع كون

الاحتمال في المرحلة الأولى  $100/10$ ، وفي الثانية  $5/1$ . ويكون الاحتمال الإجمالي لاختيار أسرة معيشية معيّنة للاستقصاء هو  $50/1$ ، أي  $100/10$  مضروباً بـ  $5/1$ .

١٢ - مع أن المثال المذكور أعلاه ليس ذا كفاءة عالية، فهو مع ذلك يوضح كيف تستخدم المرحلتان كلاتهما المعاينة الاحتمالية. وبسبب هذا، يمكن تقدير نتائج الاستقصاء بطريقة غير متحيزة بتطبيق احتمالات الاختيار تطبيقاً صحيحاً على مرحلة تحليل البيانات من عملية الاستقصاء (انظر مناقشة مُعامل الترجيح في الاستقصاء، الواردة في الفصل السادس).

### ٣ - ٢ - ١ - ٢ حساب الاحتمال

١٣ - يوضح المثال الوارد أعلاه أيضاً كيف تمت تلبية المطلبين الآخرين للمعاينة الاحتمالية. أولاً، أُعطيت كل أسرة معيشية في القرية فرصة غير صفيرية لاختيارها. خلافاً لذلك، لو استثنيت قرية واحدة أو أكثر من الاختيار لأي سبب كان، كدواعي القلق الأمنية، مثلاً، لكانت فرصة اختيار هذه القرى صفراً، وكان ذلك سيعني انتهاك الطبيعة الاحتمالية للعينة. واختيرت الأسر المعيشية في المثال الوارد أعلاه مع احتمال غير صفيري. غير أنه لو استثنيت بعض هذه الأسر عن قصد لسبب كعدم إمكانية الوصول، مثلاً، لكانت احتمالياتها صفراً وكان تنفيذ العينة قد عاد إلى تصميم عدم الاحتمال. يتناول الفرع ٣ - ٢ - ١ - ٣ طرق تناول الوضع عندما تُستثنى مناطق من الاستقصاء.

١٤ - ثانياً، يمكن حساب احتمال اختيار القرى والأسر المعيشية على حد سواء بناءً على المعلومات المتاحة. في حالة اختيار القرى كان حجم العينة  $(10)$  وحجم السكان  $(100)$  معلومين، وهذان هما المعلمان اللذان يحددان الاحتمالية،  $100/10$ . وفي حالة الأسر المعيشية، كان حساب الاحتمالية مختلفاً إلى حد ما لأننا لم نعرف قبل الاستقصاء كم عدد الأسر التي سنختار في كل قرية من قرى العينة. وإنما طُلب منا أن نختار أسرة من كل خمس أسر. فإذا كان يوجد ما مجموعه ١٠٠ أسرة في القرية ألف و ٧٥ أسرة في القرية باء فإننا سنختار ٢٠ أسرة و ١٥ أسرة على التوالي. ومع ذلك يبقى احتمال اختيار الأسرة  $5/1$  بغض النظر عن حجم السكان وحجم العينة  $(20/100 = 5/1)$ ، وكذلك  $(75/15)$ .

١٥ - بالإشارة إلى التوضيح الوارد أعلاه، كان يمكن أن يُحسب احتمال الاختيار في المرحلة الثانية كاختبار شامل بعد إتمام الاستقصاء. فعندما يكون الحرفان  $m_i$  و  $M_i$  معروفين، حيث  $m_i$  و  $M_i$  يمثلان، على التوالي، عدد أسر العينة ومجموع عدد الأسر المعيشية في القرية رقم  $i$ ، يكون الاحتمال مساوياً لـ  $m_i/M_i$ . وستكون هناك ١٠ احتمالات من هذا القبيل — واحد لكل قرية من قرى العينة. غير أن هذه النسبة ستظل، كما لاحظنا،  $5/1$  للتصميم المحدد. لذلك لن يلزم الحصول على أعداد الأسر المشمولة بالعينة ومجموع الأسر لمجرد حساب احتمالات المرحلة الثانية. ولأغراض مراقبة الجودة، سيكون من المفيد، مع ذلك، الحصول على هذه الأعداد لضمان تطبيق نسبة العينة  $5/1$  تطبيقاً دقيقاً.

### ٣ - ٢ - ١ - ٣ عندما يكون السكان المستهدفون غير معرفين جيداً

١٦ - تُنتهك أحياناً شروط المعاينة الاحتمالية بسبب كون معايير تعريف السكان المستهدفين المراد شمولهم فضفاضة. فربما يكون السكان المستهدفون، مثلاً، هم جميع الأسر المعيشية في البلد بأسره. ومع ذلك، عندما يُصمم الاستقصاء و/أو ينفذ، غالباً ما يتبين أن مجموعات فرعية معيّنة من السكان، مثل أسر البدو الرُّحَّل، وسكان القوارب، وسكان مناطق بأسرها، لا يمكن الوصول إليها بسبب وعورة التضاريس أو تُستثنى عن قصد. وفي حالات

أخرى، تُستثنى مجموعة سكانية يُراد بها أن تغطي مجموعة سكانية خاصة ومقيدة، كالنساء اللاتي تزوجن في أي وقت من حياتهن، أو الشباب الذين هم دون سن الخامسة والعشرين، (تُستثنى) مجموعات فرعية هامة لأسباب مختلفة. فمجموعة السكان المستهدفة، التي يُراد بها أن تغطي الشباب دون سن الخامسة والعشرين، مثلاً، ربما تستبعد من يكونون في الجيش أو في السجن أو محتجزين في أي ظرف آخر.

١٧ - حيثما كانت مجموعة السكان المستهدفة فعلاً، التي يشملها الاستفتاء، مختلفة عن المجموعة المقصودة، يجب أن يبذل فريق الاستقصاء عناية تامة لإعادة تعريف السكان المستهدفين بدقة أكبر. هذه مسألة هامة لا لتوضيح نتائج الاستقصاء للمستخدمين فقط، وإنما أيضاً لتلبية شروط المعاينة الاحتمالية. وينبغي وصف المثال المذكور أعلاه، مثال الشباب الذين هم دون سن الخامسة والعشرين — وصفاً أكثر دقة وإعادة تعريفهم بأنهم مدنيون غير مقيمين في مؤسسات وهم دون سن الخامسة والعشرين. وإلا، ينبغي توسيع تغطية الاستقصاء لتشمل المجموعات الفرعية المستثناة.

١٨ - هكذا نرى أن من الأهمية بمكان تعريف السكان المستهدفين بعناية فائقة لتغطية السكان الذين سيُعطون فعلاً الفرصة لاختيارهم للاستقصاء. وفي الحالات التي تُستثنى فيها مجموعات فرعية قصداً، من الضروري طبعاً تطبيق طرق احتمالية على السكان الذين يشكلون فعلاً إطار الاستقصاء. علاوةً على ذلك، لا بد من أن يصف مديرو الاستقصاء بوضوح للمستخدمين، عندما تُنشر النتائج، ما هي القطاعات السكانية التي شملها الاستقصاء، وما هي القطاعات التي استثناها.

### ٣ - ٢ - ٢ طُرُق المعاينة غير الاحتمالية

١٩ - لا توجد نظرية إحصائية، كهذه للمعاينة الاحتمالية، يُستَرشدُ بها في استخدام عيّنات غير احتمالية. فلا يمكن تقدير العيّنات إلا بتقييم موضوعي فقط. لذلك، يعني عدم استخدام تقنيات الاحتمال أن تقديرات الاستقصاء ستكون متحيزة. علاوةً على ذلك، لن يُعلم مقدار هذه التحيزات واتجاهها في الغالب إلى الانتقاص من التقدير أو الزيادة فيه. ويمكن تقدير الدقة في تقدير العيّنات، كما ذكر أعلاه، أي تقدير الأخطاء المعيارية، عندما تُستخدم المعاينة الاحتمالية. هذا أمر ضروري لكي يتمكن المستخدمون من قياس موثوقية تقديرات الاستقصاء وبناء فترات ثقة حول هذه التقديرات. ويمكن أيضاً أن تُجرى تقديرات متحيزة مع المعاينة الاحتمالية في أحوال معينة، عندما يُستصوب، مثلاً، جعل توزيع السكان المشمولين بالاستقصاء يتفوق مع سُبُل مراقبة أخرى (انظر الفصل السادس للاطلاع على مزيد من مناقشة هذه النقطة).

٢٠ - على الرغم من أوجه النقص الموجودة في العيّنات غير الاحتمالية، تُستخدم هذه العيّنات مراراً وتكراراً في أحوال وأوضاع مختلفة. ويستند التبرير الذي يقدمه مستخدموها بوجه عام إلى التكاليف أو الملاءمة أو حتى خشية فريق الاستقصاء من أن يقصّر المعاينة "العشوائية" عن تمثيل السكان المستهدفين تمثيلاً صحيحاً. وفي سياق استقصاءات الأسر المعيشية، سنناقش بإيجاز أنواعاً مختلفة من العيّنات غير الاحتمالية، ويكون معظم مناقشتنا بتقديم أمثلة وذكر بعض الأسباب التي تدعو إلى عدم استعمالها.

### ٣ - ٢ - ٢ - ١ المعاينة بناءً على أساس حكم شخصي

٢١ - المعاينة بناءً على أساس حكم شخصي طريقة تعتمد على "خبراء" ليختاروا عناصر العيّنة. ويدّعي مؤيدو هذه الطريقة أنها تزيد احتمال اختيار عيّنات "رديئة" أو شاذة، عند استعمال تقنيات الاختيار العشوائي، مثل معاينة تُسفر، مثلاً، عن وقوع جميع عناصر العيّنة، لسوء الحظ، في المنطقة الشمالية الغربية.

### مثال

من الأمثلة على أخذ عينات بناءً على أساس حكم شخصي يطبق على استقصاء للأسر المعيشية ما يحدث عندما تختار مجموعة من الخبراء عن قصد مقاطعات جغرافية لكي تُستخدم كعناصر في المرحلة الأولى من مراحل الاختيار في خطة معاينة ويستند قرارهم إلى آرائهم هم أنفسهم في أي المقاطعات هي مقاطعات نمطية أو أكثر تمثيلاً بمعنى ما أو في سياق ما.

٢٢ - الصعوبة الرئيسية في هذا النوع من المعاينة هو ذاتية تقرير ما الذي يشكل مجموعة تمثيلية من المقاطعات. من السخرية أن هذا الاختيار يتوقف أيضاً إلى حد كبير على اختيار الخبراء أنفسهم. أما في اختيار العينات الاحتمالية، خلافاً لذلك، فتُصنف المقاطعات في طبقات ويُستخدم في تصنيفها عند اللزوم أي معايير يريد فريق التصميم أن يفرضها. لاحظ أن معايير التقسيم إلى طبقات يمكن أن يكون ذاتياً، وإن كانت توجد مبادئ توجيهية لتطبيق معايير أكثر موضوعية (انظر الفرع ٣ - ٤ المتعلق بالتقسيم إلى طبقات). ثم تُختار من كل طبقة عينة احتمالية من المقاطعات (بأية طريقة من مجموعة طرق متنوعة). لاحظ أن التصنيف الطبقي يخفّض إلى حد كبير احتمال اختيار عينة شاذة كالتالي أُشير إليها أعلاه. هذا هو السبب الذي اخترع لأجله التصنيف الطبقي. ففي العينة المختارة من طبقات تكون أمام كل مقاطعة فرصة معروفة غير صفرية للاختيار، غير متحيزة وغير متأثرة برأي شخصي (حتى عندما تكون الطبقات نفسها مُعرّفة تعريفاً ذاتياً). وإن المعاينة بناءً على أساس حكم شخصي، من جهة أخرى، لا تشمل آلية تضمن لكل مقاطعة فرصة غير صفرية للمشاركة، ولا آلية لحساب مدى احتمال اختيار المقاطعات التي يقع عليها الاختيار في نهاية المطاف.

### ٣ - ٢ - ٢ - ٢ المعاينة أثناء المشي في مسار عشوائي وبموجب نظام حصص

٢٣ - يوجد نوع آخر من أنواع المعاينة غير الاحتمالية يُستخدم على نطاق واسع، وهو ما يُسمّى بإجراء معاينة عشوائية أثناء المشي في مسار عشوائي في المرحلة الأخيرة من مراحل استقصاء الأسر المعيشية. غالباً ما تُستخدم هذه الطريقة حتى وإن كانت عناصر العينات التي أخذت في المراحل السابقة قد أخذت بطرق احتمالية مشروعة. يبين التوضيح الوارد أدناه نوعاً من أنواع الإجراء الذي تؤخذ به العينات، وهو عبارة عن مزيج من المعاينة العشوائية أثناء المشي في مسار عشوائي والمعاينة بموجب نظام حصص. وهذه طريقة أخرى غير احتمالية يُعطى المقابلون بموجبها حصصاً من أنواع معينة من الأشخاص المراد مقابلتهم.

### مثال

توضيحاً لهذه الطريقة، يؤمر المقابلون ببدء المقابلة عند نقطة جغرافية مختارة عشوائياً كقرية مثلاً، ويسيروا في مسار محدد لاختيار الأسر المعيشية المراد مقابلتها. وربما يستدعي ذلك إما اختيار كل  $n^{\text{th}}$  أسرة [يحدد عدد معينة يفصل بين كل أسرة والأسرة التي تليها في المقابلة]، وإما استعراض كل أسرة يمر بها في مساره للتحقق من وجود مجموعة سكانية خاصة مستهدفة، كالأطفال الذين هم دون سن الخامسة، مثلاً. في الحالة الأخيرة، تُقابل كل أسرة مؤهلة للمقابلة من أجل الاستقصاء حتى يصل المقابل إلى نهاية حصة مقررة سلفاً من عدد الأسر المراد مقابلتها.

٢٤ - غالباً ما تكون هذه الطريقة مبررة كطريقة لتجنب نفقات مرحلة مسبقة تكون باهظة التكاليف وتستغرق وقتاً طويلاً، هدفها وضع قائمة بكل الأسر المعيشية في منطقة العينة — القرية أو المجموعة أو القطاع — قبل اختيار الأسر التي ستجرى معها مقابلات. وهو مبرر أيضاً بحجة تجنب عدم الاستجابة، لأن المقابل، رجلاً كان

أو امرأة، يواصل سيره بعدّ الأسر التي لا تستجيب حتى يحصل على العدد الكافي من الأسر المستجيبة للمء الحصّة المطلوب منه ملؤها. علاوةً على ذلك، يدّعي مؤيدو هذه الطريقة أنها غير متحيزة ما دامت نقطة بداية المسار الذي يأخذه المقابل محددة عشوائياً. ويدّعون أيضاً أنه يمكن تحديد احتمالات الاختيار تحديداً صحيحاً باعتباره حاصل قسمة عدد الأسر المختارة على مجموع عدد الأسر التي في القرية، على افتراض أن هذا العدد معروف أو يمكن تقديره إلى حد قريب من الصحة.

٢٥ - نظراً إلى الشروط التي وُضعت في الفقرة السابقة مباشرة، يمكن من الناحية النظرية الحصول على عيّنة احتمالية. غير أنه يُشكّك، من الناحية العملية، في إمكانية تحقيق ذلك فعلاً. ويفشل هذا النهج في العادة بسبب (أ) سلوك المقابل و(ب) معاملة الأسرة غير المستجيبة، بما فيها الأسر التي يُحتمل عدم استجابتها. وقد أظهرت دراسات لا حصر لها أنه عندما يُعطى المقابلون حق التحكم باختيار العيّنات في الميدان يسفر ذلك عن عيّنات متحيزة. فمتوسط عدد أفراد الأسرة المشمولة بالعيّنة (عدد الأشخاص)، مثلاً، يكون في العادة أقلّ منه في أسر السكان ككل<sup>١</sup>. فمن الطبيعة البشرية الأساسية أن يتجنب المقابل أسرة يفهم عنها أو يتصور أنها صعبة من أيّ وجه كان. لهذا السبب يكون من الأسهل تجنب أسرة يوجد لديها كلب يهدده، أو أسرة بيئتها محاطة بسور له أبواب مغلقة ويصعب الوصول إليها، ويفضل أسرة مجاورة لها لا توجد مشاكل من هذا النوع في الوصول إليها.

٢٦ - بالاستعاضة عن الأسر غير المستجيبة بأسر مستجيبة تكون المعاينة متحيزة إلى الأسر المتعاونة والتي يمكن الوصول إليها بسهولة. من الواضح أن ثمة فروقاً في خصائص الأسر المعيشية، وهذا يتوقف على مدى استعدادها وتوافرها للمشاركة في الاستقصاء. وفي نهج المعاينة بموجب نظام الحصص من الأرجح أن يكون تمثيل الأشخاص الذين يصعب الاتصال بهم أو ليس لديهم استعداد للمشاركة أقلّ من تمثيل الأشخاص في حالة اختيار العيّنات الاحتمالية. ففي النهج الأخير يُطلب من المقابلين بوجه عام أن يتصلوا عدة مرات بالأسرة التي يكون أفرادها خارج البيت مؤقتاً لدى اتصالهم بها. يضاف إلى ذلك أن المقابلين في حالة الاستقصاء القائم على الاحتمالية يكونون في العادة مدربين على بذل جهد إضافي لإقناع الأسرة المترددة بالموافقة على مقابلتها.

### ٣ - ٢ - ٢ - ٣ المعاينة بالملاءمة

٢٧ - تُستخدم المعاينة بالملاءمة على نطاق واسع بسبب سهولة تنفيذها. ومع أن المعاينة بالملائمة لا تطبق غالباً في استقصاءات الأسر المعيشية يمكن تقديم إيضاحات كثيرة لاستخدامها، مثلاً، في إجراء استقصاء للشباب من طلاب المدرسة في عيّنات مختارة عن قصد من مدارس يسهل الوصول إليها، وهي معروفة بأنها متعاونة، أي ملائمة. وثمة تطبيق آخر شائع في الوقت الراهن، وهو الاستطلاع الفوري للرأي، الذي ينظّم على مواقع شبكة الإنترنت، حيث يُسأل الأشخاص الذين يفتحون الموقع عن آرائهم في مواضيع مختلفة. ربما كان من البديهي أن تكون العيّنات المأخوذة من هذا القبيل متحيزة أصلاً وينبغي ألا تُستخدم للوصول إلى استنتاجات بشأن السكان بوجه عام.

## ٣ - ٣ تحديد حجم العيّنة لاستقصاءات الأسر المعيشية

٢٨ - وُضع هذا الفرع من الفصل بتفصيل كبير بسبب أهمية حجم العيّنة لعملية الاستقصاء بكاملها وتكاليفها. فالأهمية لا تُقاس من حيث عدد الأسر المعيشية التي تقابل فقط، وإنما تقاس أيضاً من حيث عدد المناطق

<sup>١</sup> الممارسة المتبعة الآن في كثير من منظمات الاستقصاء لضمان تسمية الأسر التي تُختار للعيّنة هي أن يُجرى اختيار العيّنات في المكتب حيث يكون من الأسهل السيطرة عليه بالإشراف. علاوةً على ذلك، ينبغي أن يختار العيّنة شخص ليس له ضلع في إعداد قائمة الأسر قبل اختيار العيّنات أو ليست له معرفة بالظرف السائد في الموقع.

الجغرافية التي تؤخذ لها عيّنات من وحدات المعاينة الأولى، وعدد المقابلين الذين يُستأجرون، وما هو حجم عبء العمل الذي يتحمّله كل مقابل، إلخ. العوامل والمعلّمات التي يجب أن تؤخذ في الاعتبار لدى تحديد حجم العيّنة كثيرة لكنها تدور بصورة رئيسية حول أهداف قياس الاستقصاء. وسنبحث تحديد حجم العيّنات من حيث التقديرات الرئيسية التي يُراد إجراؤها، والسكان المستهدفون، وعدد الأسر المعيشية التي ينبغي أخذ عيّنات منها للوصول إلى السكان المراد استهدافهم، ومستوى الدقة والثقة المرغوب، ومجالات التقدير، وهل المراد قياس المستوى أم قياس التغيّر، وأثر التقسيم إلى مجموعات، ومراعاة عدم الاستجابة، والميزانية المتاحة. ومن الواضح أن حجم العيّنة جانب محوري يتحكم بالتصميم الكلي للعيّنة.

### ٣ - ٣ - ١ أحجام تقديرات الاستقصاء

٢٩ - في استقصاءات الأسر المعيشية، سواء أكانت ذات غرض عام أم مكرسة لموضوع معيّن، كالصحة أو النشاط الاقتصادي، مثلاً، يحتاج كل تقدير (يُشار إليه غالباً بكلمة مؤشر) يُراد الحصول عليه من الاستقصاء إلى حجم مختلف للعيّنة ليتسنى القياس قياساً موثقاً. ويتوقف حجم العيّنة على حجم التقدير، بمعنى نسبه المتوقعة إلى مجموع السكان. فتقدير نسبة الأسر المعيشية التي تحصل على الماء الآمن تقديراً موثقاً، مثلاً، يحتاج إلى حجم عيّنة مختلف عن الحجم اللازم لتقدير نسبة الأشخاص البالغين العاطلين عن العمل في الوقت الراهن.

٣٠ - تستند التعبيرات عن حساب أحجام العيّنات إلى بيانات احتمالية، بأنه يجب احتواء المعلّمة الحقيقية للسكان في فترة يوجد فيها احتمال معيّن (مستوى الثقة). ويتوقف طول الفترة الفاصلة (أو الدقة) على متغير السكان المشار إليه في الجدول ٣ - ٢؛ وعلى درجة الثقة، وعلى حجم العيّنة. ويمكن القول بوجه عام أنه كلما كان مقدار تباين خواص السكان أكبر أو مستوى الثقة المرغوب أعلى، ازداد طول الفترة الفاصلة. ومن جهة أخرى، سينقص طول الفترة كلما ازدادت أحجام العيّنات أمثلة على فترات الثقة في الفقرة ٢٢ من الفصل السابع. ويمثل التعبير التالي فترة للثقة بشأن متوسط سكاني مقداره  $\bar{Y}$ ، مع أخذ طريقة تقدير المتوسط السكاني  $\hat{Y}$  بناءً على عيّنة عشوائية بسيطة من غير بديل، حجمها  $n$ .

$$(٣ - ٥) \quad P \left[ \hat{Y} - z_{1-\alpha} \sqrt{\left(1 - \frac{n}{N}\right) \frac{s^2(y)}{n}} \leq \bar{Y} \leq \hat{Y} + z_{1-\alpha} \sqrt{\left(1 - \frac{n}{N}\right) \frac{s^2(y)}{n}} \right] = (1 - \alpha) 100\%$$

حيث يكون  $1 - \alpha$  مُعامل الثقة للفترة.

٣١ - الاستقصاء نفسه يمكن أن يكون له، عملياً، حجم واحد. ولحساب حجم العيّنة يجب أن يُختار واحد من بين تقديرات كثيرة لقياسه في الاستقصاء. فإذا كان التقدير الرئيسي، مثلاً، هو نسبة البطالة، سيتوقف حساب حجم العيّنة على هذه النسبة<sup>٢</sup>. وعندما توجد مؤشرات رئيسية كثيرة يطبّق أحياناً اتفاقاً لحساب حجم العيّنة اللازمة لكل مؤشر ثم يُستخدم المؤشر الذي يأتي بأكبر عيّنة. هذا بوجه عام هو المؤشر الذي تكون المجموعة السكانية الأساسية المتعلقة به أصغر "المجموعات الفرعية السكانية المستهدفة"، من حيث نسبتها إلى مجموع

<sup>٢</sup> من قبيل التناقض الظاهري إلى حد ما أنه إذا أُريدَ حساب حجم العيّنة فمعادلته تتطلب معرفة القيمة التقريبية للتقدير المراد قياسه. غير أن القيمة يمكن أن "تُخمّن تخميناً" بطرق عدة منها، مثلاً، استخدام بيانات من تعداد أو استقصاء مماثل، أو من بلد مجاور، أو استقصاء تجريبي وما أشبه ذلك.

السكان. وينبغي طبعاً أن تؤخذ الدقة المرغوبة في الحساب (انظر أدناه). وعندما يستند حجم العيّنة إلى هذا التقدير سيقاس كل واحد من التقديرات الرئيسية الأخرى بنفس الدرجة من الوثوقية أو بدرجة أعلى منها.

٣٢ - يمكن، بدلاً من ذلك، أن يستند حجم العيّنة إلى نسبة صغيرة نسبياً من المجموعة السكانية المستهدفة بدلاً من تحديد مؤشر بعينه. وربما يكون هذا، على الأرجح، أفضل نهج في استقصاء الأسر المعيشية ذي الغرض العام، الذي يركّز على عدة مواضيع لا يرتبط بعضها ببعض، وفي هذه الحالة ربما يكون من غير العملي ومن غير الحكمة أن يستند حجم العيّنة إلى مؤشر يتعلق بموضوع منفرد. لذلك، ربما يقرر مديرو الاستقصاء أن يحددوا حجم العيّنة بناءً على أساس القدرة على قياس خاصية يمتلكها ٥ في المائة (أو ١٠ في المائة) من السكان قياساً موثوقاً، ويتوقف الخيار الدقيق في النهاية على اعتبارات الميزانية.

### ٣ - ٣ - ٢ المجموعة السكانية المستهدفة

٣٣ - يتوقف حجم العيّنة أيضاً على المجموعة السكانية المستهدفة التي سيشملها الاستقصاء. وهذه المجموعات كالمؤشرات، كثيراً ما توجد عدة مجموعات سكانية مستهدفة في استقصاءات الأسر المعيشية. فاستقصاء الصحة، مثلاً، يمكن أن يستهدف (أ) الأسر المعيشية، لتقييم إمكانية الوصول إلى الماء الصالح للشرب والمرافق الصحية، بينما يستهدف أيضاً (ب) جميع الأشخاص، لتقدير الأحوال الصحية المزمنة والشديدة، و(ج) النساء اللاتي تتراوح أعمارهن بين ١٤ و ٤٩ سنة لتقدير مؤشرات الصحة التناسلية و(د) الأطفال الذين هم دون سن الخامسة لقياس أطوالهم وأوزانهم الجسمية.

٣٤ - لذلك، يجب أن يأخذ حساب حجم العيّنة في الاعتبار كل المجموعات السكانية المستهدفة. وكما سبق ذكره، تكون لاستقصاءات الأسر المعيشية في أحيان عديدة مجموعات سكانية مستهدفة متعددة وكلها موضع اهتمام على قدم المساواة فيما يتعلق بأهداف القياس في الاستقصاء. وهنا أيضاً من المجدّي أن نركّز على أصغر مجموعة مستهدفة في تحديد حجم العيّنة. فإذا كان الأطفال الذين هم دون سن الخامسة، مثلاً، مجموعة مستهدفة هامة للاستقصاء فيجب أن يستند حجم العيّنة إلى هذه المجموعة. وقد يقرر فريق إدارة الاستقصاء، إذ يستخدم المفهوم الوارد وصفه في الفقرة ٣٢، أن يحسب حجم العيّنة لتقدير الخاصية التي يحملها ١٠ في المائة من الأطفال الذين هم دون سن الخامسة. وسيكون حجم العيّنة الناتج عن هذا القياس أكبر كثيراً من الحجم اللازم لفئة مستهدفة تتكوّن من كل الأشخاص أو كل الأسر المعيشية.

### ٣ - ٣ - ٣ الدقة والثقة الإحصائية

٣٥ - اقترح أعلاه أن تكون التقديرات موثوقة، لا سيما تقديرات المؤشرات الرئيسية. فتحديد حجم العيّنة يتوقف بصورة هامة جداً على درجة الدقة المرغوبة في المؤشرات. وكلما كانت تقديرات الاستقصاء أدق وأكثر موثوقية يجب أن يكون حجم العيّنة أكبر — وأكبر بمضاعفات. فإذا لزم مضاعفة الوثوقية، مثلاً، ربما يلزم أن يضاعف حجم العيّنة أربعة أضعاف. ومن البديهي أنه يجب أن يكون مديرو الاستقصاء على معرفة بالأثر الذي تتركه المتطلبات الصارمة للدقة على حجم العيّنة، ومن ثم على تكلفة الاستقصاء. وعلى العكس من ذلك، يجب أن يحرصوا على عدم اختيار حجم للعينة صغير جداً بحيث يجعل المؤشرات الرئيسية غير موثوقة إلى حد كبير لإجراء تحليل فيه كثير من الأخبار أو تخطيط ذي معنى.

٣٦ - كذلك، يزداد حجم العيّنة كلما زادت درجة الثقة الإحصائية المرغوبة للمحافظة على درجة الدقة المعلنة. وتؤخذ نسبة مستوى الثقة البالغة ٩٥ في المائة بصورة عالمية تقريباً باعتبار أنها المستوى القياسي المعترف به، ويُحسب حجم العيّنة اللازم لتحقيق تلك النسبة بناءً على ذلك (يرجى الرجوع إلى الفقرة ٣٠ أعلاه).

٣٧ - إذا أخذت المؤشرات في الحسبان، ألفينا تقليدياً جرى العرف باستخدامه في كثير من الاستقصاءات المصممة بناءً على تصوّر جيد باعتباره مطلب الدقة، وهو اعتماد هامش خطأ نسبي مقداره ١٠ في المائة عند مستوى الثقة البالغ ٩٥ في المائة في حالة المؤشرات الرئيسية المراد تقديرها، وهذا يعني أن الخطأ المعياري في المؤشر الرئيسي ينبغي ألا يزيد عن ٥ في المائة من التقدير نفسه. وهذا يُحسب باعتباره  $(٢ * ٠,٠٥ * x)$ ، حيث  $x$  هو تقدير الاستقصاء). فإذا كانت النسبة التقديرية للأشخاص الذين هم في القوى العاملة ٦٥ في المائة، مثلاً، ينبغي ألا يكون هامش الخطأ المعياري فيها أكثر من ٣,٢٥ في المائة، وهذا الرقم هو حاصل ضرب ٠,٦٥ بـ ٠,٠٥ ويمثل ضعف الرقم ٠,٣٢٥، أو ٠,٠٦٥، الهامش النسبي للخطأ عندما يكون مستوى الدقة بنسبة ٩٥ في المائة. وفي الفقرة ٣٠ أعلاه، مثلاً، توجد عندنا المعادلة التالية:

(٦ - ٣)

$$\sqrt{\left(1 - \frac{n}{N}\right) \frac{s^2(y)}{n}} = 0,05 \times \hat{Y}$$

٣٨ - بذلك يكون حجم العيّنة اللازم لتحقيق معيار هامش الخطأ النسبي البالغ ١٠ في المائة ربع الحجم اللازم عندما يُحدّد هامش الخطأ النسبي بـ ٥ في المائة. ويعتبر هامش الخطأ النسبي البالغ ٢٠ في المائة بوجه عام الحدّ الأقصى المسموح به للمؤشرات الهامة (وإن كنا لا نوصي به). وذلك لأن فترة الثقة المحيطة بالتقديرات التي تسمح بخطأ أكبر من ذلك أوسع من أن تحقق نتائج ذات معنى لتلبية معظم احتياجات التحليل أو احتياجات السياسات. ونوصي بوجه العموم بأخطاء نسبية تتراوح بين ٥ و ١٠ في المائة للمؤشرات الرئيسية، إن سمحت الميزانية بذلك.

### ٣ - ٣ - ٤ مجموعات التحليل: المجالات

٣٩ - يوجد عامل هام آخر له أثر كبير على حجم العيّنة، وهو عدد المجالات. تُعرّف المجالات بوجه عام بأنها المجموعات الفرعية التحليلية التي يُراد الحصول على بيانات بشأنها موثوقة بالتساوي. يُزاد حجم العيّنة بعامل يساوي عدد المجالات المطلوبة تقريباً<sup>٢</sup>. غير أن هذا يصح أيضاً إذا وُجدَ تباين مماثل في كل واحد من المجالات (انظر الحاشية رقم ٣ للحصول على مزيد من الإيضاحات). ذلك لأن حجم العيّنة عند مستوى معيّن من الدقة لا يتوقف على حجم المجموعة السكانية ذاتها إلا إذا كانت نسبتها عالية — كأن تكون ٥ في المائة أو أكثر — من حجم السكان (وهذا نادراً ما يحدث في استقصاءات الأسر المعيشية). وهكذا يكون حجم العيّنة اللازم لإقليم واحد (إذا كان الاستقصاء محصوراً بإقليم واحد) هو نفس الحجم اللازم للبلد بأسره. هذه نقطة بالغة الأهمية وكثيراً ما يُسيء فهمها العاملون في الاستقصاء، الذين يظنون خطأً أنه كلما زاد حجم المجموعة السكانية يجب أن يزداد حجم العيّنة.

٤٠ - إذن عندما تكون البيانات المطلوبة على المستوى الوطني فقط، يوجد مجال واحد فقط، وحجم العيّنة المطلوب ينطبق على العيّنة الممثلة للبلد بأسره. أما إذا تقرر أنه ينبغي الحصول على نتائج موثوقة بالتساوي

<sup>٢</sup> هكذا تكون الحالة عندما تطلب نفس الدرجة من الموثوقية لجميع المجالات.

للمناطق الحضرية وللناطق الريفية، كلاً على حدة، فعندها يُحسب حجم العيّنة لكل مجال منهما، سعياً إلى توليد نتائج موثوقة. ويجب، بوجه عام، حساب حجم العيّنة لكل مجال من المجالات ذات الصلة بطريقة تبين أنه إذا كانت المجالات  $D_1, D_2, \dots, D_k$ ، ستكون أحجام العيّنات بالضرورة  $n_1, n_2, \dots, n_k$ ، وهذا يتوقف على تباين الخاصية ذات الصلة في كل واحد من المجالات وكذلك على المستوى المنشئ للثقة والدقة. ولذلك سيكون مجموع حجم العيّنات للاستقصاء  $n = n_1 + n_2 + \dots + n_k$ .

### ٣ - ٤ - ١ الإفراط في المعاينة لتقديرات المجالات

٤١ - ثمة أثر هام لمطلب التساوي في الموثوقية بين جميع المجالات، وهو أنه يجب استخدام نسب معاينة غير متناسبة. فإذا لم يكن التقسيم بالناصفة ٥٠ - ٥٠، كما يُرجَّح أن تكون الحالة في المجالين الحضري والريفي، ربما يكون من الضروري تعمد الإفراط في المعاينة من المجال الحضري، مثلاً، في معظم البلدان للحصول على الموثوقية بالتساوي. غير أنه ينبغي توكيد أن الإفراط في المعاينة في مجال من مجالات الدراسة في استقصاء وطني تمليه بصورة رئيسية الحاجة إلى الحصول على تقديرات تتفق مع مستوى معين للثقة.

٤٢ - من المهم أن نلاحظ نتيجتين من نتائج الإفراط في المعاينة من المجموعات الفرعية سواء أكان ذلك من المجالات أم من الطبقات. الأولى، أن الإفراط يستدعي استخدام معاملات ترجيح تعويضية في الاستقصاء لإنتاج تقديرات على الصعيد الوطني. والثانية، وهي الأهم، أن التقديرات الوطنية ستكون، إلى حد ما، أقل موثوقية مما تكون عليه إذا وزعت العيّنات بالتناسب بين المجموعات الفرعية.

### ٣ - ٤ - ٢ اختيار المجالات

٤٣ - المناطق الفرعية الجغرافية مهمة طبعاً ويوجد دائماً ضغط لمعاملتها كمجالات لأغراض التقدير. ففي الاستقصاء الذي يُجرى على الصعيد الوطني، مثلاً، يؤدّ المستخدمون في الدوائر غالباً أن يحصلوا على بيانات لا لكل منطقة كبيرة فقط، وإنما في الغالب لكل إقليم أيضاً. ومن الواضح أنه يجب أن يُنظر بعناية في عدد المجالات وأن تُختار مجالات التقدير التي تشكّل هذه المجالات بحكمة. وإن الاستراتيجية الممكنة عملياً هي التي تقرّر ما هي مجموعات التقدير التي لا تتطلب، رغم أهميتها، موثوقية على قدم المساواة في قياس الاستقصاء. وبدلاً من ذلك، تُعامل مجموعات التقدير، في التحليل، كقوائم جدولة رئيسية، لا كمجالات. عندئذ تكون أحجام العيّنات لكل فئة أصغر كثيراً مما تكون عليه إذا عوملت كمجالات؛ ومن ثمّ تكون موثوقيتها كذلك أقل. غير أنه يجب ملاحظة أن الإفراط في المعاينة من مجال معين ربما يأتي بدافع الحاجة إلى الحصول على تقديرات، في ذلك المجال، على مستوى معين من الثقة والدقة مستقل عن المستويات المنشأة على المستوى الوطني.

### مثال

يشير المثال التالي إلى كيفية المعاينة وكيف يكون أثر ذلك على الموثوقية إذا عوملت المناطق الحضرية والمناطق الريفية كمجموعات جدولة لا كمجالات. افترض أن توزيع السكان كان بنسبة ٦٠ في المائة في المناطق الريفية و٤٠ في المائة في المناطق الحضرية. فإذا تقرر أن حجم العيّنة المحسوب سيكون ٨٠٠٠ أسرة معيشية، لتلبية دقة محددة، يجب إذن أخذ عيّنات بحجم ١٦٠٠٠ أسرة معيشية إذا كانت المناطق الحضرية والريفية مجالين مختلفين — ٨٠٠٠ أسرة من كل قطاع. وإذا عوملت، بدلاً من ذلك، كمجموعات جدولة، تُختار العيّنة الوطنية المؤلفة من ٨٠٠٠ أسرة، بالتناسب بين المناطق الريفية والحضرية، بعدد ٨٠٠ أسرة من المناطق الريفية و٣٢٠٠ أسرة من المناطق الحضرية. وافترض، علاوة على ذلك أن الخطأ

المعياري المتوقع بنسبة ١٠ في المائة للخاصية، بناءً على عينة مقدارها ٨٠٠٠ أسرة، يساوي ٠,٧ في المائة. فهذا هو الخطأ المعياري الذي ينطبق على العينة الوطنية (أو على العينتين الحضرية والريفية كل على حدة، إذا كانت العينة المأخوذة من كل واحد من المجالين ٨٠٠٠ أسرة). في حالة العينة الوطنية البالغة ٨٠٠٠ أسرة معيشية، المختارة بالتناسب من المناطق الحضرية والريفية، يكون الخطأ المعياري بالنسبة إلى المناطق الريفية نحو ٠,٩ في المائة، محسوباً كحاصل الجذر التربيعي لنسبة أحجام العينات، والخطأ المعياري في التقدير الوطني، أو

$$\cdot \left( \sqrt{\frac{8\,000}{4\,800}} \right)^{*0.7}$$

وفي المناطق الحضرية، يكون الخطأ المعياري نحو ١,١ في المائة، أو

$$\cdot \left( \sqrt{\frac{4\,800}{3\,200}} \right)^{*0.7}$$

وثمة طريقة أخرى لتقييم الأثر، وهي حقيقة أن الأخطاء المعيارية في جميع التقديرات الريفية ستكون أعلى بنحو ٢٩ في المائة

$$\cdot \left( \sqrt{\frac{8\,000}{4\,800}} \right)$$

منها في التقديرات الوطنية؛ وستكون في المناطق الحضرية أعلى بنحو ٥٨ في المائة

$$\cdot \left( \sqrt{\frac{8\,000}{3\,200}} \right)$$

٤٤ - لاحظ أن الجملة الأخيرة من المثال تنطبق بغض النظر عن مقدار الخطأ المعياري على الصعيد الوطني. بعبارة أخرى، تنطبق على كل تقدير مُجَدُولٍ في الاستقصاء. لذلك يمكن تحليل الأثر على الوثوقية، قبل المعاينة، لمختلف المجموعات الفرعية التي يمكن اعتبارها مجالات. وبهذه الطريقة، سيكون لدى فريق الاستقصاء المعلومات اللازمة لمساعدته على أن يقرر إن كان ينبغي أن تُعامل المجالات المحتملة مُعاملة مجموعات الجدولة. هذا يعني، كما ورد ضمناً أعلاه، أنه سيُستخدَمُ تخصيصُ العينات بالتناسب، لا بالتساوي. فإذا حُطِّطَ استقصاء وطني لبلد نسبة سكان المناطق الحضرية فيه ٢٠ في المائة فقط، سيكون حجم العينة المأخوذة من المناطق الحضرية ٢٠ في المائة فقط من مجموع حجم العينة. وبذلك يكون مقدار خطأ العينة من التقديرات الحضرية ضعف (الجذر التربيعي لـ  $0,8/0,2n$ ) خطأ العينة من التقديرات الريفية، ويكون أكبر بنحو ضعفين وربع ضعف خطأ العينة من التقديرات الوطنية (الجذر التربيعي لـ  $n/0,2n$ ). في هذه الحالة، يمكن أن يقرر مديرو الاستقصاء أن من الضروري الإفراط في المعاينة من القطاع الحضري،<sup>٤</sup> وبذلك يخلقون، عملياً، مجالين منفصلين أحدهما حضري والآخر ريفي.

٤٥ - يمكن، كذلك، إجراء تحليل للعلاقة بين الأخطاء المعيارية والمجالات مقابل مجموعات الجدولة للإرشاد في عملية اتخاذ القرارات بشأن ما إذا كان ينبغي اتخاذ المناطق أو وحدات جغرافية أخرى دون الوطنية كمجالات، وفي هذه الحالة كم يكون عددها. وحيث إن من الضروري أخذ عينات متساوية الحجم من المجالات، فإن استخدام ١٠ مناطق يتطلب ١٠ أضعاف حجم العينة الوطنية؛ لكن هذا العدد سينخفض إلى النصف إذا رُئِيَ من المناسب تعيين ٥ مناطق فقط باعتبارها كافية لتلبية احتياجات السياسة العامة. وكذلك، إذا عوملت المناطق مُعاملة

<sup>٤</sup> يُتَّخَذُ هذا القرار إذا كانت الأخطاء المعيارية النسبية المتوقعة للمؤشرات الحضرية الرئيسية (أو أيٍّ منها)، مثلاً، أكبر من ٧,٥ في المائة (يكون مستوى الثقة البالغ ٩٥ في المائة بنسبة ١٥ في المائة، وهي النسبة المقترحة في هذا الدليل باعتبارها الحد الأعلى المسموح به).

مجموعات الجدولة، بدلاً من ذلك، فستوزع العيّنة الوطنية بينها بالتناسب. وفي هذه الحالة، يكون مقدار الأخطاء المعيارية في المنطقة المتوسطة أكبر منه في حالة التقديرات الوطنية بـ ٣,٢ أضعاف، إذا كان في البلد ١٠ مناطق، لكن أكبر منه بضعفين فقط إذا كان في البلد ٥ مناطق.

### ٣ - ٣ - ٥ آثار التقسيم إلى مجموعات

٤٦ - يبحث هذا الفرع كيف يتأثر تحديد حجم العيّنة (يوجد بحث أكثر تفصيلاً لمعاينة المجموعات في الفرع ٣ - ٥). درجة تجميع العيّنات في استقصاء الأسر المعيشية تؤثر في موثوقية التقديرات أو دقتها، ومن ثم في حجم العيّنات. تنشأ آثار التقسيم في مجموعات في استقصاءات الأسر المعيشية من (أ) وحدات العيّنات قبل الأخيرة، التي يُشار إليها بوجه العموم بكلمة "المجموعات"، وهذه قد تكون قُرى أو أحياءً في مدينة، (ب) الأسر التي تُستخدم كعيّنات، (ج) حجم المجموعات و/أو تباينها، (د) من أسلوب أخذ عيّنات الأسر المعيشية داخل المجموعات المختارة. ويمكن قياس التجميع وكذلك آثار التقسيم إلى طبقات قياساً رقمياً بقياس أثر التصميم (*deff*)، الذي يعبر عن مقدار زيادة حجم تباين المعاينة (مربع الخطأ المعياري) لعينة المجموعة الطباقية، مقارنة بالعيّنة العشوائية البسيطة المساوية لها في الحجم. ويميل التقسيم إلى طبقات إلى تخفيض تباين المعاينة لكن بدرجة صغيرة فقط. وخلافاً لذلك، يزيد التقسيم إلى مجموعات التباين زيادة كبيرة. لذلك يشير أثر التصميم بالدرجة الأولى إلى مقدار التقسيم إلى مجموعات الموجود في عيّنة الاستقصاء.

٤٧ - يتطلّب تصميم العيّنات بكفاءة أن تُستخدم المجموعات للتحكم بالتكاليف لكن مع إبقاء أثر التصميم منخفضاً قدر الإمكان لكي تكون النتائج قابلة للاستعمال وموثوقة. من سوء الحظ أن أثر التصميم لا يكون معروفاً قبل إجراء الاستقصاء، لكن يمكن تقديره بعد ذلك فقط من البيانات نفسها. وفي الحالات التي تكون قد أُجريت فيها استقصاءات سابقة أو استقصاءات مشابهة في بلدان أخرى، يمكن استخدام قيم أثر التصميم في تلك الاستقصاءات كبداية في معادلة الحساب لتقدير حجم العيّنة.

٤٨ - لإبقاء أثر العيّنة منخفضاً قدر الإمكان، ينبغي أن يتّبع تصميم العيّنة هذه المبادئ العامة (انظر أيضاً موجز المبادئ التوجيهية الوارد في نهاية هذا الفصل).

(أ) استخدم أكبر قدر ممكن عملياً من المجموعات؛

(ب) استخدم أصغر حجم للمجموعات من حيث عدد الأسر المعيشية الممكن؛

(ج) استخدم حجماً ثابتاً للمجموعات، لا حجماً متغيراً؛

(د) اختر عيّنة منهجية للأسر المعيشية في المرحلة الأخيرة، متفرقة جغرافياً، لا قطاعاً من أسر معيشية متجاورة جغرافياً.

٤٩ - وهكذا يفضّل، في عيّنة حجمها ١٢ ٠٠٠ أسرة معيشية، اختيار ٦٠٠ مجموعة تضم كل واحدة منها ٢٠ أسرة على اختيار ٤٠٠ مجموعة تضم كل واحدة منها ٣٠ أسرة معيشية. فأثر تصميم العيّنة أقل كثيراً في الحالة الأولى. علاوة على ذلك، ينخفض أثر التصميم إذا كان اختيار الأسر المعيشية منهجياً من جميع الأسر التي في المجموعة بدلاً من اختيارها من قطاعات فرعية متجاورة جغرافياً. عندما تُتبع هذه القياسات التقديرية يُرجّح أن يكون أثر التصميم أقل إلى حدّ معقول.

### ٣ - ٣ - ٦ تعديل حجم العيّنة لعدم الاستجابة المتوقع

- ٥٠ - جرت العادة في الاستقصاءات بأن يُزاد حجم العيّنة بمقدار مساوٍ لنسبة عدم الاستجابة المتوقعة. وهذا يضمن أن يكون عدد المقابلات المنجزة في الاستقصاء قريباً جداً من حجم العيّنة المستهدف.
- ٥١ - تتفاوت الدرجة المتوقعة لعدم الاستجابة في الاستقصاءات بحسب البلد ونوع الاستقصاء. في العملية الحسابية الواردة أدناه، سنسمح بنسبة متوقعة لعدم الاستجابة مقدارها ١٠ في المائة. وينبغي طبعاً أن تستخدم البلدان الرقم الذي يعكس خبرتها السابقة في إجراء الاستقصاءات بشكل أكثر دقة.

### ٣ - ٣ - ٧ حجم العيّنة من العينات الرئيسية

- ٥٢ - تُبحث العينات الرئيسية بالتفصيل في الفصل الرابع. ويركز هذا الفرع على حجم العيّنة في خطة العينات الرئيسية. ويمكن القول بإيجاز إن العيّنة الرئيسية عينة كبيرة من وحدات العينات الأولية في البلدان التي يوجد فيها برامج استقصاء متكاملة ومستمرة. ويقصد بالعيّنة الكبيرة أن تقدم من حالات العينات المسجلة في "البنك" ما يكفي لدعم استقصاءات متعددة على مدى بضع سنوات دون أن يلزم إجراء مقابلات متكررة مع المستجيبين.
- ٥٣ - عندما تُجرى استقصاءات كثيرة، ومن ثم تُستوعب مواضع فنية كثيرة في العيّنة الرئيسية، توجد طبعاً مجموعات سكانية مستهدفة عديدة وتقديرات رئيسية يجب خدمتها. وفي هذا الصدد، تنشئ معظم البلدان عينات تستند أحجامها إلى اعتبارين اثنين. الأول، اعتبار يتعلق بالميزانية، وهذا لا يحتاج إلى دليل. والثاني هو أحجام العينات المتوقعة للاستقصاءات الفردية، التي يمكن استخدامها على مدى الفترة الزمنية المشمولة بالعيّنة الرئيسية، والتي غالباً ما تصل إلى ١٠ سنين بين عمليات تعداد السكان. وبذلك تكون أحجام العينات الممكن تحديدها للعينات الرئيسية كبيرة جداً، وقد تصل إلى ٥٠ ٠٠٠ أسرة معيشية أو أكثر. وتُصاغ خطط الاستفادة من "بنك" الأسر المعيشية بأسره.

#### مثال

افترض أن العيّنة الرئيسية في البلد ألف تضم ٥٠ ٠٠٠ أسرة معيشية. ويُقصد بالعيّنة الرئيسية أن تُستخدم في إجراء ثلاثة استقصاءات حُطّط لها مسبقاً، ويُحتمل استخدامها في استقصاءين اثنين محتملين لم يُحطّط لهما بعد. سيُجرى واحد من الاستقصاءات لإيرادات الأسر المعيشية ونفقاتها، وسيكرر ثلاث مرات خلال العقد — في السنوات الأولى والخامسة والثامنة. صمّم هذا الاستقصاء لاستقصاء ٨ ٠٠٠ أسرة معيشية في كل واحدة من السنوات الثلاث التي يعمل فيها. غير أنه في السنة الخامسة سوف يُستعاض بعينة مقدارها ٤ ٠٠٠ أسرة عن نصف العيّنة الأولى البالغ مقدارها ٨ ٠٠٠ أسرة، التي قوبلت في السنة الأولى. وكذلك في السنة الثامنة، سيُستعاض عن الـ ٤ ٠٠٠ أسرة الباقية من السنة الأولى بـ ٤ ٠٠٠ أسرة جديدة. وهكذا، ستستعمل ١٦ ٠٠٠ أسرة لاستقصاء الإيرادات والنفقات. والاستقصاء الثاني المخطط له هو استقصاء للصحة، ويُتوقع أن تستخدم فيه ١٠ ٠٠٠ أسرة معيشية، بينما يستعمل الاستقصاء الثالث، الذي سيجرى لمشاركة القوى العاملة، ١٢ ٠٠٠ أسرة معيشية. وبذلك تكون ٣٨ ٠٠٠ أسرة قد خُصّصت لهذه الاستقصاءات الثلاثة. وبناءً على ذلك، ما زالت ثمة ١٢ ٠٠٠ أسرة يمكن استخدامها في استقصاءات أخرى إذا لزم الأمر.

### ٣ - ٣ - ٨ تقدير التغيّر أو المستوى

٥٤ - في الاستقصاءات التي تُكرّر دورياً، يوجد هدف قياس رئيسي وهو تقدير التغيرات التي تحدث بين استقصاء وآخر. ففي المصطلحات الإحصائية يقدم تقدير الاستقصاء الذي يُحصّل عليه في المناسبة الأولى مستوى مؤشر معيّن، بينما الفرق بينه وبين تقدير المستوى في المناسبة الثانية هو التغيّر المقدّر. عند تقدير التغيّر، يلزم بوجه عام وجود حجم عينة أكبر كثيراً بالمقارنة مع الحجم اللازم لدى تقدير المستوى فقط، وذلك للتوصل إلى استنتاجات موثوق بها. غير أنه توجد تقنيات معيّنة للمعاينة من شأنها تخفيض حجم العينة (ومن ثمّ تخفيض التكلفة) لدى تقدير التغيّر (انظر الفرع ٣ - ٩ - ٢).

### ٣ - ٣ - ٩ ميزانية الاستقصاء

٥٥ - ربما يكون من البديهي أنه لا يمكن تجاهل ميزانية الاستقصاء لدى تحديد حجم مناسب للعيّنات في استقصاء الأسر المعيشية. ومع أن الميزانية ليست معلّمة تدخل في الحساب الرياضي لحجم العينة، فهي تأخذ مكاناً بارزاً على المستوى العملي.

٥٦ - إنه الخبير الإحصائي الذي هو أول من يحسب حجم العينة، إذ يأخذ في الحسبان كل واحد من المعلّات التي بُحّثت في هذا الفصل. غير أن الحقيقة في الغالب هي أن الحجم يثبت أنه أكبر مما تستطيع الميزانية تحمّله. وعندما يحدث ذلك يجب أن يلتمس فريق الاستقصاء أموالاً إضافية للاستقصاء، أو يقلل إمّا متطلبات الدقة وإمّا عدد المجالات.

٥٧ - إن من مسؤولية فنيّ العيّنات أن يساعد على إرشاد المشتركين في البحث بشأن التكلفة مقابل الدقة. وينبغي له أن يشرح الأشياء التي يمكن تعويض بعضها ببعضها الآخر، التي قد تنشأ عن الحد من عدد المجالات (تقليل الفائدة التي تعود على المستخدمين) أو تخفيض متطلبات الدقة (تقليل موثوقية المؤشرات الرئيسية)، كلما دعت الحاجة إلى تخفيض حجم العينة المناسب بسبب اعتبارات الميزانية. وينبغي أن يسير البحث على غرار الأمثلة على الدقة والمجالات التي دُكرت أعلاه. وينبغي على مختار العيّنات أن يزن بعناية حقيقة كون عدد من المجموعات تشكّل عاملاً رئيسياً في تحديد تكاليف الاستقصاء ودقة الاستقصاء، في إرشاداته لفريق الاستقصاء (يوصل هذا البحث في الفرع ٣ - ٥ - ٥).

### ٣ - ٣ - ١٠ حساب حجم العينة

٥٨ - نقدّم في هذا الفرع معادلة لحساب حجم العينة، آخذين في الحسبان المعلّات التي سبق بحثها. ولأننا نركز على استقصاءات الأسر المعيشية، يُحسب حجم العينة بأعداد الأسر المعيشية التي يجب اختيارها. وتقدم توضيحات بالرسوم أيضاً.

٥٩ - بوجه عام، عندما تُحسب النسبة  $p$ ، تكون معادلة التقدير ° لحجم العينة  $n_b$  هي

(٧ - ٣)

$$n_b = (z^2)(1-r)(f)(k)/(r)(p)(\bar{n})(e^2)$$

° يمكن أن تحتوي معادلة حجم العينة أيضاً على عامل، وهو المسمّى المضاعف المتناهي الصغر، الذي يجب أن يُؤخذ في الحسبان، حين يتبيّن أن حجم العينة المقيس يشكل نسبة كبيرة من حجم السكان. غير أن هذه الحالة نادرة ما تحدث في الاستقصاءات الكبيرة للأسر المعيشية من النوع الذي يُنظر فيه في هذا الدليل. بناءً على ذلك، يُفترض أن قيمة المضاعف المتناهي الصغر تساوي ١،٠، ولذلك يتم تجاهله في المعادلة ٣ - ٥.

حيث يكون  $n_b$  هو المَعْلَمَة المراد حسابها وهو حجم العَيِّنة من حيث عدد الأسر المعيشية المراد اختيارها؛ وتكون  $z$  الإحصائية التي تعرّف مستوى الثقة المرغوب؛ ويكون  $r$  تقديراً للمؤشر الرئيسي يراد أن يقيسه الاستقصاء؛ ويكون  $f$  أثر تصميم العَيِّنة الذي يفترض أنه ٢,٠ (القيمة الافتراضية)؛ ويكون  $k$  مضاعفاً مطلوباً لحساب النسبة المتوقعة لعدم الاستجابة؛ وتكون  $p$  النسبة التي تشكلها المجموعة السكانية المستهدفة من مجموع السكان والتي تستند إليها المَعْلَمَة  $r$ ؛ ويكون  $\bar{n}$  متوسط حجم الأسرة المعيشية (عدد الأشخاص في كل أسرة معيشية)؛ و  $e$  هو هامش الخطأ المراد بلوغه. وفيما يلي القيم الموصى بها لبعض المَعْلَمَات.

٦٠ - الإحصائية  $z$  المراد استخدامها ينبغي أن تكون ١,٩٦ لمستوى الثقة البالغ ٩٥ في المائة ( خلافاً لـ ١,٦٤٥ للمستوى البالغ ٩٠ في المائة). ويعتبر الرقم الأول بوجه عام المعيار السائد لإعطاء درجة الثقة المرغوبة في تقسيم هامش الخطأ في استقصاءات الأسر المعيشية. تُحدّد القيمة الفرضية لأثر تصميم العَيِّنة عادة بالرقم ٢,٠، ما لم توجد بيانات تجريبية مؤيدة من استقصاءات سابقة أو ذات صلة لاقتراح قيمة مختلفة. ينبغي اختيار مضاعف عدم الاستجابة،  $k$ ، لكي يعكس خبرة البلد نفسه في عدم الاستجابة — التي تكون عادة أقل من ١٠ في المائة في البلدان النامية. لذلك تعتبر القيمة البالغة ١,١ لـ  $k$  اختياراً محافظاً. يمكن في العادة حساب المَعْلَمَة  $p$  من نتائج أحدث عمليات التعداد. وغالباً ما تكون المَعْلَمَة  $\bar{n}$  نحو ٦,٠ في معظم البلدان النامية، لكن ينبغي استخدام القيمة المضبوطة، التي تكون متاحة في العادة من آخر تعداد. فيما يتعلق بهامش الخطأ،  $e$ ، يوصى بتحديد مستوى الدقة بـ ١٠ في المائة من  $r$ ؛ وبذلك يكون  $e = 0.10r$ . ويمكن الحصول على عَيِّنة أصغر حجماً مع هامش خطأ أقل تشدداً،  $e = 0.15r$ ، لكن نتائج الاستقصاء تكون بالطبع أقل موثوقية. ومن شأن الاستعاضة عن بعض القيم المختارة أن تعطينا:

(٣ - ٨)

$$n_b = (3,84) (1-r) (1,2) (1,1)/(r) (p) (6) (0,01)$$

وتخفضه المعادلة (٣ - ٢) كذلك إلى

(٣ - ٩)

$$n_b = (84,5) (1-r)/(r) (p)$$

٦١ - يمكن استخدام النص المخفض عندما تستخدم كل القيم الافتراضية الموصى بها بدلاً من القيم الأكثر دقة، المتاحة من خبرة البلد نفسه.

### مثال

تقرّر في البلد باء أن المؤشر الرئيسي للاستقصاء المراد قياسه هو نسبة البطالة، التي يُعتقد أنها نحو ١٠ في المائة من القوى العاملة المدنية. وتعرّف القوى العاملة المدنية بأنها المجموعة السكانية التي تبلغ أعمارها ١٤ سنة فما فوق، وهذه تشكّل ٦٥ في المائة من مجموع سكان البلد. في هذه الحالة يكون  $r = 0.1$  و  $p = 0.65$ . افترض أننا نود أن نقدر نسبة البطالة مع هامش خطأ مقداره ١٠ في المائة ومستوى ثقة مقداره ٩٥ في المائة؛ عندئذٍ  $e = 0.10r$  (أي خطأ معياري مقداره ٠,٠١، مثلاً) كما أوصي أعلاه. علاوة على ذلك، قيم نسبة عدم الاستجابة المتوقعة، وأثر التصميم، ومتوسط حجم الأسرة المعيشية هي الأرقام التي أوصينا بها. لذلك، يمكننا أن نستخدم المعادلة (٣ - ٩)، التي تعطينا عَيِّنة حجمها ١١٧٠ أسرة معيشية  $[(84,5*0,9)/(0,1*0,65)]$ . هذا حجم صغير إلى حد ما للعَيِّنة، والسبب الرئيسي في ذلك هو أن القاعدة السكانية تشكّل نسبة عالية جداً من مجموع السكان، وهي ٦٥ في المائة. تذكر أن حجم العَيِّنة المحسوب لمجال واحد فقط — في هذه الحالة هو المستوى الوطني. فإذا كانت أهداف القياس تشمل

الحصول على بيانات موثوقة للمناطق الحضرية والريفية بالتساوي، فعندئذٍ سيضاعف حجم العيّنة، على افتراض أن كل معلّمتي المعادلتين (٣ - ٨) و (٣ - ٩) تنطبق على كلتا الفئتين — المناطق الريفية والمناطق الحضرية. وبقدر ما يكون الاختلاف بين الفئتين (ربما يكون حجم الأسرة المعيشية الحضرية، مثلاً، مختلفاً عن حجم الأسرة الريفية، وكذلك نسبة عدم الاستجابة المتوقعة في المناطق الحضرية قد تختلف عنها في المناطق الريفية)، ينبغي استخدام القيم الأكثر دقة لحساب أحجام العيّنات للمناطق الحضرية والمناطق الريفية كلاً على حدة. وستكون النتائج مختلفة طبعاً.

٦٢ - المثال التالي يتناول مجموعة سكانية أساسية أصغر حجماً — الأطفال الذين هم دون سن الخامسة.

### مثال

في البلد جيم، تقرّر أن يكون المؤشر الرئيسي للاستقصاء نسبة الوفيات بين الأطفال الذين هم دون سن الخامسة، التي يُعتقد أنها تساوي ٥ في المائة. وفي هذه الحالة نجد أن  $r = 0.05$  وأن  $p$  تقدر بنحو 0,15 أو  $0,03*5$ . هنا أيضاً نود أن نقدّر نسبة الوفيات مع هامش خطأ نسبي مقداره ١٠ في المائة، عندئذٍ  $e = 0,10r$  (أو خطأ معياري مقداره ٠,٠٠٥) وتكون قيم عدم الاستجابة المتوقعة وأثر التصميم ومتوسط حجم الأسرة المعيشية هنا أيضاً الأرقام التي أوصينا بها. وتعطينا المعادلة (٣ - ٩) نحو ١٠٧٠٤ أسر معيشية  $(84,5*0,95)/(0,05*0,15)$ ، وهذا حجم عيّنة أكبر كثيراً من نظيره في المثال السابق. وهنا أيضاً يتصل السبب الرئيسي لذلك بحجم المجموعة السكانية الأساسية، أي الأطفال الذين هم دون سن الخامسة، الذين يشكّلون ١٥ في المائة من مجموع السكان. والمعلّمة المقدّرة  $r$  صغيرة أيضاً، وتجتمع هي والعامل الصغير  $p$  ليفرضا حجماً كبيراً للعيّنة.

٦٣ - المثال الأخير لحالة تكون المجموعة السكانية المستهدفة فيها هي مجموع السكان. في هذه الحالة،  $p = 1$  ويمكن تجاهله؛ غير أنه يظل في الإمكان استخدام المعادلتين (٣ - ٨) و (٣ - ٩) إذا استُخدمت القيم والمعلّمتان الموصى بها.

### مثال

تقرر أن يكون المؤشر الرئيسي في البلد دال نسبة الأشخاص الذي أصيبوا بحالة صحية حادة خلال الأسبوع السابق من مجموع السكان. ويعتقد أن هذه النسبة تتراوح بين ٥ و ١٠ في المائة، وفي هذه الحالة تستخدم القيمة الصغرى لأنها تعطي حجماً أكبر للعيّنة (النهج المحافظ). ففي هذه الحالة  $r = 0.05$  و  $p$  تساوي 1,0 طبعاً. هنا أيضاً نود تقدير نسبة الإصابات الحادة مع هامش خطأ نسبي مقداره ١٠ في المائة:  $e = 0,10r$  (أو خطأ معياري مقداره ٠,٠٠٥)، وهنا أيضاً تكون قيم نسبة عدم الاستجابة المتوقعة، وأثر التصميم، ومتوسط حجم الأسرة المعيشية هي نفس القيم التي أوصينا بها. وتعطي المعادلة (٣ - ٩) عيّنة حجمها أكثر قليلاً من ١٦٠٠ أسرة معيشية  $(84,5*0,95)/(0,05)$ .

٦٤ - كما دُكرَ آنفاً، ربما يتقرر حجم العيّنة في نهاية المطاف بحساب أحجام العيّنات من عدة مؤشرات رئيسية وإقامة القرار على أساس المؤشر الذي يعطي العيّنة الأكبر حجماً. بالإضافة إلى ذلك، ينبغي قبل التوصل إلى قرار نهائي النظر في عدد مجالات الاستقصاء وميزانية الاستقصاء.

٦ لما كان  $r$  ينطبق على السكان كلهم في هذه الحالة، فهو يساوي  $p$ ، وبذلك فإن  $e$  تساوي  $0,10p$ .

٦٥ - في حالات البلدان التي لا يتوفر لديها واحد أو أكثر من الافتراضات المذكورة أعلاه، يمكن إدخال تغييرات صغيرة على المعادلة (٣ - ٧) للحصول على أرقام دقيقة لحجم العينة. فقد يكون حجم الأسرة المعيشية، مثلاً، أكبر أو أصغر من ٦,٠؛ وربما يتوقع أن تكون نسبة عدم الاستجابة ٥ بدلاً من ١٠ في المائة؛ ويمكن أن تُحسب قيمة  $p$  لبلد معين بوجه عام حساباً أكثر دقة باستخدام أرقام التعداد.

٦٦ - غير أنه يوصى بعدم إحداث أي تغيير على القيمة الإحصائية لـ  $z$  البالغة ١,٩٦، التي تُستخدم بوجه العموم معياراً. وينبغي، لأغراض عملية، أن يُترك أثر التصميم  $f$ ، هو أيضاً، بقيمته البالغة ٢,٠، ما لم توح بغير ذلك بيانات تعداد سابق، كما ذُكر أعلاه. ويوصى أيضاً بأن يعرف  $e$  بأنه  $0,10r$ ، فيما عدا الحالات التي لا تستطيع فيها الميزانيات تحمّل حجم العينة الناتج عن هذه المعادلة. ففي تلك الحالة، يمكن زيادته إلى  $0,12r$  أو  $0,15r$ . غير أن هذه الزيادات على هامش الخطأ ستسفر عن أخطاء أكبر كثيراً في أخذ العينات.

### ٣ - ٤ التقسيم إلى طبقات

٦٧ - تقسيم السكان المراد استقصاؤهم إلى طبقات قبل اختيار العينات أسلوب شائع الاستعمال في تصميم استقصاءات الأسر المعيشية. فهو يصنف السكان في مجموعات سكانية فرعية — طبقات — على أساس معلومات مساعدة معروفة عن كل السكان. ثم تُختار عناصر العينات، مستقلةً بعضُها عن بعض، من كل طبقة، بطريقة تتفق مع أهداف القياس التي يضعها الاستقصاء.

### ٣ - ٤ - ١ التقسيم إلى طبقات وتخصيص العينات

٦٨ - في اختيار العينات من الطبقات، يتحكم في أحجام العينات المختارة من الطبقة الواحدة فنّي اختيار العينات، ولا يتقرر حجمها بصورة عشوائية بواسطة عملية المعاينة. فالسكان المقسمون إلى طبقات يمكن أن يكونوا وحدات معلومة  $n_s$  تؤخذ من كل منها عينات، ويشير الرمز  $n_s$  إلى العدد المرغوب من وحدات العينات في الطبقة ذات الرقم  $s$ . خلافاً لذلك، العينة غير المأخوذة من طبقات تعطي حجم عينة للمجموعة السكانية الفرعية ذات الرقم  $s$  تختلف إلى حد ما عن  $n_s$ .

#### مثال

افتراض أن تصميم العينة لاستقصاء ما يجب أن يتألف من طبقتين — حضرية وريفية. المعلومات الآتية من تعداد السكان متاحة لتصنيف جميع الوحدات الإدارية الجغرافية إما إلى وحدات حضرية وإما وحدات ريفية، مما يمكن من تصنيف السكان إلى طبقات بموجب هذا المعيار. وقد تقرر اختيار عينة متناسبة (خلافاً للعينة اللامتناسبة) في كل طبقة لأن السكان موزعون إلى ريفيين بنسبة ٦٠ في المائة وحضرين بنسبة ٤٠ في المائة. فإذا كان حجم العينة ٥٠٠٠ أسرة معيشية، سيضمن الاختيار المستقل للعينة أن تكون ٣٠٠٠ أسرة منها ريفية و٢٠٠٠ أسرة حضرية. وإذا اختيرت العينة عشوائياً فسيختلف عدد الأسر من كل فئة عن نسبة ٣٠٠٠ - ٢٠٠٠، وإن كان من المتوقع أن يكون انقسام الفئتين بهذه النسبة. ويمكن أن يكون تقسيم العينة غير المأخوذة من الطبقات، بمحض الصدفة غير المؤاتية ٣٢٠٠ أسرة ريفية و٢٨٠٠ أسرة حضرية.

٦٩ - لذلك يتمثل سبب واحد من أسباب التقسيم إلى طبقات في تخفيض فرصة أن يؤدي سوء الحظ إلى إيجاد عدد غير متناسب كبير (أو صغير) لوحدات العينة المختارة من مجموعة سكانية فرعية تعتبر هامة للتحليل. يجرى

التقسيم إلى طبقات لضمان تمثيل مجموعات سكانية فرعية هامة دون تحيز في عملية الاختيار. غير أن من الجدير بالملاحظة أن التمثيل الصحيح لا ينطوي على التناسب في المعاينة. وفي كثير من التطبيقات، يمكن أن تكون واحدة من الطبقات أو أكثر أيضاً مجال تقدير (بُحِثت أعلاه). في هذه الحالة، ربما يكون من الضروري اختيار عيّنات متساوية في الحجم من الطبقات المتأثرة، ومن ثم تنتج عندنا عيّنة غير متناسبة طبقياً. من ثمّ يكون التخصيص المتناسب وغير المتناسب لوحدة العيّنة بين الطبقات سمات تصميم شرعي في العيّنة المختارة من طبقات، ويتوقف الاختيار على أهداف القياس الواردة في الاستقصاء.

٧٠ - قد يوفر التقسيم إلى طبقات أيضاً، كما ذكرنا أعلاه في الجملة السابقة مباشرة، وسيلة لتخصيص العيّنة ضمناً، وهذا أسلوب أبسط وأكثر عملية من أسلوب التخصيص المثالي<sup>٧</sup>. بعبارة أخرى، مع وجود العيّنات المتناسبة بحسب الطبقة لا يلزم حساب عدد حالات العيّنات التي ستخصص لكل طبقة مسبقاً.

### مثال

افترض أن من أهداف تصميم العيّنات ضمان مخصصات دقيقة وتناسبية من أحجام العيّنات لكل واحد من الأقاليم العشرة التي يتكون منها البلد. فإذا كان الإقليم ألف، مثلاً، يحتوي على ١٢ في المائة من سكان الدولة، ينبغي أن تُختار ١٢ في المائة من مجموعات العيّنات من ذلك الإقليم شريطة أن يكون حجم المجموعة المتوقع ثابتاً. وافترض كذلك أن مجموع عدد المجموعات التي ستُختار من الدولة كلها ٤٠٠ مجموعة. يوجد أسلوب غالباً ما يُستعمل في بلدان كثيرة، وهو تخصيص ٤٨ (0,12 \* 400) مجموعة للإقليم ألف. غير أن هذا الإجراء غير ضروري إذا وجد التقسيم الصحيح إلى طبقات. ويجب، بدلاً من ذلك، مُعاملة كل إقليم كطبقة مستقلة في عملية اختيار العيّنات. وإن تطبيق عملية معاينة منهجية باحتمال متناسب مع الحجم، (انظر الجدول ٣ - ١)، ومع وجود فترة واحدة للمعاينة، سيسفر تلقائياً عن ٤٨ مجموعة متوقعة للإقليم ألف. سيُبحِث هذا النوع من التقسيم إلى طبقات وكذلك استخدامه في تبسيط مشاريع التخصيص في الفرع ٣ - ٤ - ٣.

### ٣ - ٤ - ٢ قواعد التقسيم إلى طبقات

٧١ - توجد قاعدتان أساسيتان ينبغي تطبيقهما لدى تقسيم السكان إلى طبقات. ويلزم دائماً التقيد بإحدى القاعدتين. وينبغي احترام القاعدة الأخرى في الأحوال العادية، وإن كان عدم احترامها لا يلحق ضرراً يُذكر بتصميم العيّنة. القاعدة المطلوبة هي أنه يجب اختيار وحدة عيّنات واحدة على الأقل من كل طبقة تُنشأ. والطبقات أساساً مجموعات فرعية من السكان مستقلة ويَجِبُ بعضُها بعضاً: يَجِبُ أن يوجد كل عنصر من عناصر السكان في واحدة من الطبقات دون غيرها. وبسبب هذه الخاصية، ينبغي أخذ عيّنة من كل طبقة وأن يُحسب منها تقدير غير متحيز لمتوسط السكان. وحيث إن كل طبقة يمكن نظرياً مُعاملتها معاملة مستقلة عن غيرها في تصميم العيّنة، لا يحتاج إنشاء الطبقات إلى تحقيقه باستخدام معايير موضوعية؛ وإنما يمكن أيضاً استخدام معايير ذاتية إذا وجدت رغبة في ذلك. وينبغي أن يكون المبدأ التوجيهي أن الوحدات التي تتكون منها الطبقة يجب أن تكون متشابهة، قدر الإمكان فيما يتعلق بمتغير الدراسة بغية تخفيض التباين داخل كل طبقة.

<sup>٧</sup> تُشير عبارة التخصيص المثالي إلى تخصيص الموارد على أساس وظائف التكلفة والتفاوتات المختلفة داخل الطبقة الواحدة (مقاسات التفاوت). وهي لم تناقش في هذا الدليل لأنها نادراً ما تُستخدم عملياً في البلدان النامية. وقد يعود ذلك إلى قلة الأرقام الثابتة للتكلفة في عمليات الاستقصاء. وربما يجد القارئ معلومات مفصلة عن التخصيص الأمثل في كثير من المراجع التي يوردها هذا الفصل.

٧٢ - القاعدة الثانية للتقسيم إلى طبقات هي أن كل طبقة تُنشأ يجب، في الأحوال المثالية، أن تكون مختلفة قدر الإمكان عن الطبقات الأخرى. لذلك، يجب أن يشكّل التباين بين الطبقات والتجانس داخل كل طبقة الخاصية الرئيسية التي ينبغي الاسترشاد بها في إنشاء الطبقات. لذلك، يمكن أن نرى بسهولة لماذا تُنشأ في الغالب المناطق الحضرية والمناطق الريفية كطبقتين من طبقات استقصاء الأسر المعيشية. فسكان المناطق الحضرية وسكان المناطق الريفية مختلفون بعضهم عن بعض بطرق كثيرة، كما ذكر أعلاه (نوع الوظيفة، مصدر الدخل ومقداره، متوسط حجم الأسرة، نسب الخصوبة، إلخ). بينما أفراد السكان داخل كل مجموعة فرعية متشابهون.

٧٣ - خاصية التباين مرشد مفيد في تحديد عدد الطبقات التي ينبغي إنشاؤها. وينبغي ألا تُنشأ طبقات أكثر من المجموعات السكانية الفرعية التي يمكن تعريفها لكل معيار معين يستخدم في تعريف الطبقات. فإذا كان البلد مقسماً إلى ثماني مناطق جغرافية للأغراض الإدارية وكانت اثنتان من هذه المناطق متشابهتين إلى حد كبير فيما يتعلق بموضوع الاستقصاء المقترح، يمكن تحقيق تصميم مناسب للعينات بإنشاء سبع طبقات (الجمع بين المنطقتين المتشابهتين). فلا يمكن تحقيق أيّ مكسب بإنشاء ٢٠ طبقة، مثلاً، إذا كانت ١٠ طبقات كافية لإنشاء المجموعات الفرعية المتغايرة نفسها.

٧٤ - من الأهمية بمكان ملاحظة أنه فيما يتعلق بالاختيار التناسبي تكون العينة الناتجة على الأقل معادلة في الدقة للعينة العشوائية البسيطة التي هي بنفس الحجم. لذلك ينتج التقسيم إلى طبقات مكاسب في دقة تقديرات الاستقصاء أو موثوقيتها، وأكثر ما تكون المكاسب عندما يبلغ التباين بين الطبقات أقصى مدى له. هذه الخاصية للعينات المأخوذة من الطبقات هي التي تضمن أنه حتى التقسيم الرديء إلى طبقات<sup>٨</sup> لا يضر بتقديرات الاستقصاء من حيث موثوقيتها.

٧٥ - ثمة نقطة أخرى هامة تتعلق بتقدير أخطاء العينات. فبينما تكفي عينة واحدة مختارة من كل طبقة لتلبية المتطلبات النظرية للمعاينة من الطبقات، ينبغي اختيار عينتين اثنتين كحد أدنى لكي تُستخدم نتائج اختيار العينات في حساب أخطاء العينات في تقديرات الاستقصاء.

٧٦ - ربما يكون من الضروري في بعض الأحيان استخدام أكبر عدد من المتغيرات للتقسيم إلى طبقات. ويجب في هذه الحالات أن نسترشد بالعوامل التالية: يُفضّل أن تكون متغيرات التقسيم إلى طبقات غير متصل بعضها ببعض لكنها متصلة بمتغير الاستقصاء؛ لا حاجة إلى الكمال، وإلى تشكيل الخلايا (يمكن الجمع بين الخلايا الصغيرة والأقل أهمية)؛ وبوجه عام، يمكن تحقيق مكاسب من استخدام التقسيمات العريضة لعدة متغيرات أكثر مما يمكن تحقيقه من التقسيمات الدقيقة لمتغير واحد.

### ٣ - ٤ - ٣ التقسيم الضمني إلى طبقات

٧٧ - يتحدد اختيار المعلومات المتاحة لإنشاء الطبقات، كما ذكر أعلاه، بواسطة أهداف القياس الواردة في الاستقصاء. وتوجد طريقة مفيدة بوجه خاص لاستقصاءات الأسر المعيشية التي هي كبيرة الحجم ومتعددة الأغراض في مضمونها، يُسمّى التقسيم الضمني إلى طبقات. وتكفي حقيقة أن معيارها الأساسي جغرافي، بوجه عام، لنشر العينة على الوجه الصحيح بين المجموعات الفرعية الهامة من السكان مثل المناطق الحضرية والمناطق الريفية، والمناطق الإدارية، والمجموعات السكانية الفرعية العرقية، والمجموعات الاقتصادية - الاجتماعية، إلخ.

<sup>٨</sup> يمكن أن تحدث رداءة التقسيم إلى طبقات عندما تُنشأ طبقات لا لزوم لها، أو عندما يُساء تصنيف بعض العناصر السكانية في الطبقات الخاطئة.

بسبب هذه الخاصية الجغرافية يكون التقسيم الضمني إلى طبقات ذا فائدة كبيرة حتى وإن كان موضوع الاستقصاء مركّزاً على موضوع واحد، سواء أكان القوى العاملة، أم النشاط الاقتصادي للأسرة المعيشية، أم مقياس الفقر، أم الصحة، أم الإيرادات والنفقات. يوصى توصيةً قوية باستخدام هذه الطريقة لهذه الأسباب وكذلك لسهولة تطبيقها.

٧٨ - يحتاج التقسيم الضمني إلى طبقات، لكي يُطبق تطبيقاً صحيحاً، إلى استخدام اختيار منهجي في المرحلة الأولى للمعاينة. تنفيذ هذا الإجراء بسيط ويستدعي أولاً ترتيب ملف وحدات المعاينة الأولية بتتابع جغرافي. ويُرجَّح أن يكون التتابع في بلدان كثيرة حضرياً بحسب الإقليم وداخل الإقليم بحسب المقاطعة، يليه التتابع الريفي بحسب الإقليم، وداخل الإقليم بحسب المقاطعة. والخطوة التالية هي اختيار وحدات المعاينة الأولية اختياراً منهجياً من الملف المُفرز. ويُجرى الاختيار المنهجي إما بأخذ العيّنات بناءً على تساوي الاحتمالات وإما، وهذا هو الأرجح، أخذ العيّنات بحيث يكون الاحتمال متناسباً مع الحجم.

٧٩ - ثمة ميزة هامة للتقسيم الضمني إلى طبقات، كما ذُكر آنفاً، وهي أنه يقضي على الحاجة إلى إنشاء طبقات جغرافية صريحة. وهذا بدوره يقضي على الحاجة إلى تخصيص العيّنة لتلك الطبقات، لا سيما عند استخدام العيّنات التناسبية. وتوجد ميزة أخرى هي البساطة المذكورة في الفقرة السابقة، لأن هذه الطريقة لا تحتاج إلى ما هو أكثر من فرز الملفات وتطبيق فترة (فترات) المعاينة. ويمكن أيضاً تطبيق المعاينة غير التناسبية بسهولة على المستوى الأول من مستويات الفرز الجغرافي. فإذا كانت المناطق الحضرية والمناطق الريفية، مثلاً، تشكل المستوى الأول، يكون تطبيق نسب العيّنات على المناطق الحضرية والريفية مسألة بسيطة لا تعقيد فيها. ويقدم الشكل ٣ - ١ مشروع تقسيم ضمني إلى طبقات مع المعاينة المنهجية. ويرد مزيد من البحث للمعاينة باحتمال متناسب مع الحجم في الفرع ٣ - ٦ أدناه.

### ٣ - ٥ معاينة المجموعات

٨٠ - وُضع مصطلح "معاينة المجموعات" أصلاً للإشارة إلى تصاميم عيّنات تؤخذ فيها عيّنات من جميع أعضاء المجموعة. وعُرِّفت المجموعات نفسها بأنها "المجموعات". فيمكن، مثلاً، اختيار عيّنة من المدارس في المرحلة الأولى، ومن الصفوف الدراسية في المرحلة الثانية. وإذا استُقصي جميع الطلاب في كل صف يكون لدينا عيّنة مجموعة للصفوف. في استقصاءات الأسر المعيشية، يتمثل مثال على الفكرة الأساسية، فكرة معاينة المجموعات باختيار أحياء من المدينة يُقابل فيها جميع القاطنين في الحي لأجل الاستقصاء. غير أنه في السنوات الأخيرة، استُخدمت "معاينة المجموعات" بوجه عام للإشارة بصورة أكثر عمومية إلى الاستقصاءات التي توجد فيها مرحلة معاينة قبل الأخيرة، تُختار فيها (وتُعرَّف) المجموعات كالحقير، مثلاً، ومناطق تعداد السكان، وأحياء المدينة. وفي مرحلة المعاينة الأخيرة تُستقصى عيّنة فرعية من الأسر المعيشية في كل مجموعة مختارة، بدلاً من استقصاء جميع الأسر. الاستخدام الأخير لهذا المصطلح هو المستعمل بوجه عام في هذا الدليل.

٨١ - سوف تستخدم المعاينة في استقصاءات الأسر المعيشية، في جميع الأحوال وبالضرورة، شكلاً ما من أشكال معاينة المجموعات، لضمان احتواء تكاليف الاستقصاء. فإجراء استقصاء لـ ١٠٠٠ أسرة، مثلاً، في ٥٠ موقعاً (بواقع ٢٠ أسرة في كل مجموعة) أرخص كثيراً، كما قلنا آنفاً، من استقصاء ١٠٠٠ أسرة مختارة عشوائياً من بين جميع السكان. لكن من المؤسف أن معاينة المجموعات تقلل الموثوقية بسبب الاحتمال الراجح أن يميل السكان القاطنون في نفس المجموعة إلى أن يكونوا متجانسين أو تكون لهم خصائص متشابهة إلى حد ما. ينبغي التعويض عما يُسمّى أثر تشكيل المجموعات في تصميم العيّنات بزيادة حجم العيّنة زيادة تناسبية.

الشكل ٣ - ١  
ترتيبات المناطق الإدارية للتقسيم الضمني إلى طبقات

الأقاليم الحضرية	
الإقليم رقم ٠١	
المقاطعة رقم ٠١	
منطقة تعداد السكان رقم ٠٠١	
منطقة تعداد السكان رقم ٠٠٢	
منطقة تعداد السكان رقم ٠٠٣	
منطقة تعداد السكان رقم ٠٠٤	
المقاطعة رقم ٠٢	
منطقة تعداد السكان رقم ٠٠٥	
منطقة تعداد السكان رقم ٠٠٦	
منطقة تعداد السكان رقم ٠٠٧	
الإقليم رقم ٠٢	
المقاطعة رقم ٠١	
منطقة تعداد السكان رقم ٠٠٨	
منطقة تعداد السكان رقم ٠٠٩	
المقاطعة رقم ٠٢	
منطقة تعداد السكان رقم ٠١٠	
منطقة تعداد السكان رقم ٠١١	
منطقة تعداد السكان رقم ٠١٢	
الإقليم رقم ٠٣، إلخ.	
الأقاليم الريفية	
الإقليم رقم ٠١	
المقاطعة رقم ٠١	
منطقة تعداد السكان رقم ١٠١	
منطقة تعداد السكان رقم ١٠٢	
منطقة تعداد السكان رقم ١٠٣	
منطقة تعداد السكان رقم ١٠٤	
المقاطعة رقم ٠٢	
منطقة تعداد السكان رقم ١٠٥	
منطقة تعداد السكان رقم ١٠٦	
منطقة تعداد السكان رقم ١٠٧	
الإقليم رقم ٠٢	
المقاطعة رقم ٠١	
منطقة تعداد السكان رقم ١٠٨	
منطقة تعداد السكان رقم ١٠٩	
منطقة تعداد السكان رقم ١١٠	
منطقة تعداد السكان رقم ١١١	
المقاطعة رقم ٠٢	
منطقة تعداد السكان رقم ١١٢	
منطقة تعداد السكان رقم ١١٣	
منطقة تعداد السكان رقم ١١٤	
الإقليم رقم ٠٣، إلخ.	

## ٣ - ٥ - ١ خصائص معاينة المجموعات

٨٢ - تختلف معاينة المجموعات اختلافاً كبيراً عن معاينة الطبقات بطريقتين<sup>٩</sup>. في الطريقة الأخيرة تُمثّل جميع الطبقات في العيّنة، لأنّ عيّنة من الوحدات تُختار من كل طبقة. أما في معاينة المجموعات فتُختار المجموعات نفسها؛ وبذلك تمثل المجموعات التي في العيّنة المجموعات التي ليست فيها. ويؤدي هذا الفرق المتميز الأول بين العيّنات المأخوذة من الطبقات وتلك المأخوذة من المجموعات إلى الطريقة الثانية التي تختلف فيها هذه العيّنات. وينبغي، كما ذُكر آنفاً، أن تُنشأ الطبقات، في الأحوال المثالية، على نحو تكون فيه متجانسة داخلياً ومتباينة خارجياً، فيما يتعلق بمتغيرات الاستقصاء المراد قياسها. وعكس ذلك هو الصحيح في المجموعات. وثمة ميزة أكبر، من حيث دقة العيّنات، أن تكون المجموعات متغايرة داخلياً قدر الإمكان.

٨٣ - ممّا يذكر أن حقيقة كون المجموعات، في استقصاءات الأسر المعيشية كلها تقريباً تُعرّف كوحدات جغرافية، كالقرى أو أجزاء القرى، تعني لسوء الحظ أنه لا تتحقق بوجه عام درجة عالية من التغير داخل المجموعة. فالواقع هو أن المجموعات المعرّفة جغرافياً يَرَجَّح أن تكون أكثر تجانساً منها تغيّراً في الداخل، فيما يتعلق بمتغيرات مثل نوع الوظيفة (كالزراعة، مثلاً)، ومستوى الإيرادات، وما أشبه ذلك. وإن درجة تجانس المجموعات من حيث متغير معيّن تقرر مدى "تجميع" العيّنة. وكلما ازدادت صبغة التجميع في العيّنة كلما قلّت موثوقيتها.

## ٣ - ٥ - ٢ أثر تصميم المجموعة

٨٤ - يُقاس أثر تجميع العيّنة جزئياً بأثر التصميم (*deff*). غير أن أثر التصميم يعكس أيضاً آثار التصنيف إلى طبقات. ولا بد لفريق تصميم العيّنات من أن يضمن أن تلتمس خطة أخذ العيّنات تحقيق توازن مثالي بين تقليل التكاليف وزيادة الدقة. ويتحقق هذا بتخفيض أثر التصميم أو التحكم به إلى أقصى حد ممكن. ولتقرير كيفية تقليل عنصر التجميع في أثر التصميم أو التحكم به من المفيد إلقاء نظرة على تعريفه الرياضي:

$$(١٠ - ٣)$$

$$deff = 1 + \delta(\bar{n} - 1)$$

حيث الرمز  $\delta$  هو مُعامل الارتباط داخل الطبقة (أو داخل المجموعة)، بمعنى أنه بقدر ما تكون وحدتان في مجموعة مكافئتين لوحدين مختارتين اختياراً عشوائياً من السكان، يُرَجَّح أن تكون لهما نفس القيمة؛ والرمز  $\bar{n}$  هو عدد وحدات المجموعة السكانية المستهدفة في المجموعة.

٨٥ - المعادلة (١٠ - ٣) ليست، بالمعنى الصحيح للكلمة، هي المعادلة لأثر التصميم، لأن التقسيم إلى طبقات موضع تجاهل، هو وكذلك عامل آخر يُدخَل عندما تكون المجموعات غير موحدة الحجم. ومع ذلك، ما دام عنصر التجميع هو العامل الأبرز في أثر التصميم، يمكن استخدامه كنموذج تقريبي مهمته هي أن يبين كيف يؤثر التجميع في تصميم العيّنة وما الذي يمكن فعله للتحكم به.

٨٦ - يمكن أن يُرى في التعبير الوارد أعلاه أن أثر التصميم هو دالّة مضاعفة لتغيرين اثنين هما مُعامل الارتباط داخل الفئة،  $\delta$ ، وحجم المجموعة،  $\bar{n}$ . لذلك يزداد أثر التصميم كلما ازداد الرمز  $\delta$  و  $\bar{n}$  كلاهما. وإن كانت المعاينة لا تستطيع ممارسة أيّ تحكّم بمُعامل الارتباط داخل الطبقة لأيّ متغير قيد النظر، فإنها تستطيع مع ذلك أن تعدل حجم العيّنة صعوداً أو نزولاً بتصميمها للعيّنة، وبذلك تتحكّم إلى حد كبير بأثر التصميم.

<sup>٩</sup> من الأهمية بمكان أن نلاحظ أن معاينة الطبقات ومعاينة المجموعات ليستا بديلين متنافستين في تصميم العيّنات، لأن الطريقتين تُستخدمان بدون استثناء في معاينة استقصاءات الأسر المعيشية.

## مثال

افترض أن مجموعة سكانية يوجد لها مُعامل ارتباط داخل الطبقة مقداره  $0,3$ ، وهذا صغير إلى حد ما في الأحوال الصحية المزمّنة. وافترض أيضاً أن مخططي العيّنات يناقشون إن كان عليهم أن يستخدموا مجموعات تتألف الواحدة منها من ١٠ أسر معيشية أو ٢٠ أسرة، وأن الحجم الإجمالي للعيّنة هو ٥٠٠٠ أسرة. وافترض كذلك، لمجرد تبسيط الشكل، أن كل الأسر المعيشية متماثلة في الحجم، خمسة أفراد في كل أسرة. عندئذ تكون قيمة  $r$  ٥٠ لـ ١٠ أسر معيشية و ١٠٠ لـ ٢٠ أسرة معيشية. وإذا أجرينا عملية تبديل بسيطة في المعادلة (٣ - ٤) يعطينا ذلك قيمة لأثر التصميم هي  $[0,3 + 1] (٤٩)$ ، أو  $٢,٥$  لتصميم مجموعة الـ ١٠ أسر معيشية، لكنها  $٤,٥$  لتصميم الـ ٢٠ أسرة معيشية. وبذلك يكون أثر التصميم أكبر بنسبة تبلغ نحو ٦٠ في المائة من حجم المجموعة الأكبر. عندئذ سيُضطر فريق الاستقصاء إلى أن يقرّر أيّ الخيارين أفضل لأخذ عيّنات مجموعات يبلغ عددها ضعف الـ (٥٠٠) باستخدام خيار العشر أسر معيشية لإبقاء الوثوقية في مستوى أكثر قبولاً أو أن يختار الخيار الأقل تكلفة البالغ ٢٥٠ أسرة مقابل زيادة متغير المعاينة زيادةً كبيرة. يمكن طبعاً النظر في خيارات أخرى تقع بين العشر أسر معيشية والعشرين أسرة معيشية.

٨٧ - توجد عدة طرق لتفسير أثر التصميم: فهو يُفسّر باعتباره العامل الذي يزيد به متغير العيّنات للتصميم الفعلي للعيّنة التي ستستخدم في الاستقصاء عن قيمة العيّنة العشوائية البسيطة التي هي من نفس الحجم؛ أو باعتباره مجرد قياس لمدى كون خطة المعاينة الفعلية أسوأ من العيّنة العشوائية البسيطة من حيث الدقة؛ أو باعتباره يجسّد عدد العيّنات الأخرى التي يجب اختيارها لتحقيق نفس الدرجة من الوثوقية التي تنتجها العيّنة العشوائية البسيطة. فآثر التصميم الذي تبلغ قيمته  $٢,٥$  مثلاً يعني أنه سيتعيّن اختيار ضعف عدد الحالات لبلوغ نفس مستوى الوثوقية الذي تحقّقه العيّنة العشوائية البسيطة ولذلك، يتّضح لنا أنّ وجود خطة عيّنات يكون أثر التصميم فيها أكبر كثيراً من  $٢,٥ - ٣,٥$  بالنسبة للمؤشرات الرئيسية هو أمر غير مستحب.

## ٣ - ٥ - ٣ حجم المجموعة

٨٨ - لُوَظ أن مصمّم العيّنات لا يستطيع التحكم بمعاملات الارتباط. علاوةً على ذلك، في حالة معظم متغيرات الاستقصاء لم يحاول بحث يذكر، هذا إن وُجد، من البحوث التجريبية تقدير قيمة مُعاملات الارتباط هذه. يمكن أن يتفاوت مُعامل الارتباط داخل الفئة، نظرياً، بين  $-١$  و  $+١$ ، مع أن من الصعب تصوّر متغيرات كثيرة متعلقة بالأسر المعيشية تكون فيها قيمة المُعامل سلبية. لذلك فإن الإمكانية الوحيدة التي توجد لدى مصمّم العيّنات لإبقاء أثر التصميم عند الحد الأدنى هي أن يحدّد على إبقاء أحجام المجموعات صغيرة بقدر ما تسمح به الميزانية. ويعرض الجدول ٣ - ٣ آثار التصميم لقيم متفاوتة لمُعامل الارتباط داخل الطبقة وحجم عيّنة ثابت.

الجدول ٣ - ٣

مقارنة عناصر التجميع لأثر التصميم بغية تنويع مُعاملات الارتباط داخل الطبقة  $\delta$  وأحجام المجموعات  $n$ 

$n$	$\delta$					
	٠,٥٠	٠,٣٥	٠,٢٠	٠,١٥	٠,١٠	٠,٠٥
٥	٣,٠٠	٢,٤٠	١,٨٠	١,٦٠	١,٤٠	١,٢٠
١٠	٥,٥٠	٤,١٥	٢,٨٠	٢,٣٥	١,٩٠	١,٤٥
٢٠	١٠,٥٠	٦,٦٥	٤,٨٠	٣,٨٥	٢,٩٠	١,٩٥
٣٠	١٥,٥٠	١١,١٥	٦,٨٠	٥,٣٥	٣,٩٠	٢,٤٥
٥٠	٢٥,٥٠	١٨,١٥	١٠,٨٠	٨,٣٥	٥,٩٠	٣,٤٥
٧٥	٣٨,٠٠	٢٦,٩٠	١٥,٨٠	١٢,١٠	٨,٤٠	٤,٧٠

٨٩ - يمكن أن يُرى بوضوح من الجدول ٣ - ٣ أن أحجام المجموعات التي تزيد عن ٢٠ تعطي أثراً للتصميم غير مقبول (أكثر من ٣,٠) ما لم يكن مُعامل الارتباط داخل الطبقة صغيراً جداً. ومن المهم أن نذكر لدى تقييم الأرقام الواردة في الجدول أن الرمز  $n$  يشير إلى عدد الوحدات من السكان المستهدفين، لا إلى عدد الأسر المعيشية. وفي هذا الصدد تكون قيمة الرمز  $n$  التي سُنستخدَم مساويةً لعدد الأسر المعيشية التي في المجموعة، مضروباً بمتوسط عدد الأشخاص في المجموعة المستهدفة. فإذا كانت المجموعة المستهدفة نساءً في الفئة العمرية ١٤ - ٤٩، مثلاً، توجد في العادة امرأة واحدة تقريباً من كل أسرة معيشية في هذه المجموعة، وفي هذه الحالة سيكون في كل مجموعة من الحجم  $b$  أسرة معيشية فيها نفس العدد تقريباً من النساء اللاتي هن في الفئة العمرية ١٤ - ٤٩. بعبارة أخرى، تكون قيمة  $n$  وقيمة  $b$  متساويتين تقريباً في حالة المجموعة المستهدفة، وينطبق الجدول ٣ - ٣ كما هو. وفي المثال الذي يليه عدد الأسر المعيشية والمجموعة السكانية المستهدفة غير متساويين.

### مثال

افترض أن المجموعة السكانية المستهدفة تضم جميع الأشخاص، كما يكون الحال في استقصاء صحي يُجرى لتقدير حالات صحية شديدة ومزمنة. وافترض كذلك أنه يُقصد بالاستقصاء استخدام مجموعات تضم الواحدة منها عشر أسر معيشية. تكون قيمة  $n$  في تلك الحالة ١٠ أضعاف متوسط حجم الأسرة المعيشية؛ فإذا كان متوسط الحجم ٥,٠ أشخاص تكون قيمة  $n$  هي ٥٠. وهكذا تكون ٥٠ هي قيمة  $n$  التي يجب أن تُستخدم في الجدول ٣ - ٣ لتقييم أثر التصميم المحتمل عليه. ويكشف الجدول ٣ - ٣ أن أثر التصميم كبير جداً إلا عندما يكون  $\delta$  نحو ٠,٠٢. ويوحى هذا بأن عيّنة مجموعة مصممة لاستخدام عيّنة لا يزيد حجمها عن ١٠ أسر معيشية لكل مجموعة، تعطي نتائج غير موثوق بها لخاصية كالحالات الصحية المعدية، لأن هذه الأخيرة يَرَجَّح أن تكون  $\delta$  فيها كبيراً.

٩٠ - يُوضَّح المثال سبب الأهمية البالغة التي تُعلَّق على أن يؤخذ حجم المجموعة في الحسبان لدى تصميم استقصاء للأسر المعيشية، لا سيما للمؤشرات الرئيسية المراد قياسها. علاوةً على ذلك، يجب ألا يغرب عن البال أن حجم المجموعة المعلن، في وصف تصميم العيّنة، سيشير بوجه عام إلى عدد الأسر المعيشية، بينما يجب أن يأخذ حجم العيّنة في الاعتبار، لأغراض تقييم آثار التصميم، المجموعة (أو المجموعات) السكانية المستهدفة، بدلاً من ذلك.

### ٣ - ٥ - ٤ حساب أثر التصميم (deff)

٩١ - يمكن حساب الآثار الفعلية للتصميم على متغيرات الاستقصاء، التي حددها المحللون، بعد إكمال الاستقصاء. وهذا يتطلب تقدير تباين العينات للمتغيرات المختارة (ستُبحث الطرق في الفصل السابع) ثم تُحسب لكل متغير نسبة تباينه عن نسبة تباين العينة البسيطة المختارة عشوائياً، التي يكون حجمها الإجمالي مساوياً للعينة. هذا الحساب عبارة عن تقدير لأثر التصميم "الكامل"، بما في ذلك آثار التقسيم إلى طبقات وكذلك تباين أحجام المجموعات، بدلاً من أحجام عناصر المجموعات.

٩٢ - يعطينا الجذر التربيعي لنسبة التباينات نسبة الأخطاء المعيارية، أو أثر التصميم، كما يُسمّى، ويُحسب هذا غالباً في الممارسة ويُقدّم في الوثائق التقنية للاستقصاءات، كالاستقصاءات الديمغرافية والصحية، مثلاً.

### ٣ - ٥ - ٥ عدد المجموعات

٩٣ - من الأهمية بمكان ألا يغرب عن بالنا أن أهمية حجم المجموعة تتجاوز أثره على دقة المعاينة ومسائل أيضاً تتعلق بالحجم الإجمالي للعينة، لأن حجم المجموعة يحدد عدد المواقع المختلفة التي ينبغي زيارتها في الاستقصاء. وهذا يؤثر طبعاً تأثيراً كبيراً في تكاليف الاستقصاء، ولذلك تُستخدم عينات المجموعات في المقام الأول. لذلك فإن عينة مقدارها ١٠٠٠٠ أسرة معيشية، إذا قُسمت إلى مجموعات حجم المجموعة منها ١٠ أسر، تحتاج إلى ١٠٠٠ مجموعة، بينما إذا كان حجم المجموعة ٢٠ أسرة، تحتاج إلى ٥٠٠ مجموعة. ومن المهم غاية الأهمية، كما أكدنا من قبل، أنه يجب أن يؤخذ عاملاً التكاليف والدقة كلاهما في الحسبان إذا ما أريد التوصل إلى قرار بشأن هذا الجانب من جوانب تصميم العينات.

### ٣ - ٦ - ٦ المعاينة على مراحل

٩٤ - على الصعيد النظري، تستدعي الخطة الكاملة للمعاينة في استقصاء الأسر المعيشية اختيار العينات أسر معيشية عددها،  $n$  للعينة، اختياراً عشوائياً من بين طبقات مُعرّفة تعريفاً ملائماً، تشكّل مجموع السكان البالغ عددهم  $N$  أسرة. العينة المأخوذة من الطبقات عشوائياً على هذا النحو ستعطي أكبر قدر ممكن من الدقة. غير أن استخدام عينة من هذا النوع سيكون باهظاً إلى حد يجعله غير عملي<sup>١١</sup>، كما لاحظنا من قبل في بحث الفوائد من حيث التكلفة التي تعود بها معاينة المجموعات.

### ٣ - ٦ - ١ فوائد المعاينة على مراحل

٩٥ - اختيار العينة على مراحل له فوائد عملية في عملية الاختيار نفسها. فهو يمكن مختار العينات من القيام، في خطوات متتابعة، بعزل المواقع الجغرافية لعمليات الاستقصاء — لا سيما وضع قوائم بالأسر المعيشية وإجراء المقابلات. وعندما يجب وضع قوائم لأن إطار المعاينة قديم، يمكن إدخال مرحلة اختيار للحد من حجم المنطقة المراد وضع القوائم بأسرها.

<sup>١١</sup> يُوجد استثناء واحد أو استثناءان وهما يخصان بلداناً صغيرة جداً من الناحية الجغرافية، مثل الكويت، حيث تكون تكاليف السفر، الذي يستدعيه اختيار عينة عشوائية من الأسر المعيشية، منخفضة جداً.

٩٦ - في حالة معاينة المجموعات بوجه عام، توجد مرحلتان اثنتان على الأقل في إجراء الاختيار — الأولى هي اختيار المجموعات، والثانية هي اختيار الأسر المعيشية. وتُعرّف المجموعات في استقصاءات الأسر المعيشية دائماً بأنها وحدات جغرافية من نوع ما. وإذا كانت تلك الوحدات صغيرة إلى الحد الكافي، من الناحية الجغرافية ومن حيث عدد السكان أيضاً، وحديثة، ووجدت قائمة كاملة ودقيقة بهذه المجموعات تُختار منها العيّنات، عندئذٍ تكفي مرحلتان اثنتان لتخطيط المعاينة. وإذا كانت أصغر وحدة جغرافية متوفرة كبيرة جداً بحيث لا يمكن استخدامها بكفاءة كمجموعة، سيكون من الضروري وضع ثلاث مراحل.

### مثال

افتراض أن بلداً ما يودُّ أن يُعرّف مجموعات كمناطق لتعداد السكان [تكتب اختصاراً بالإنجليزية EAs]، لأن هذه أصغر وحدة جغرافية موجودة إدارياً. إطار مناطق تعداد السكان (انظر الفصل الرابع للاطلاع على تفاصيل أكثر عن مناقشة الأطر) هنا كامل لأن البلد كله مقسم إلى مناطق تعداد السكان. وهو دقيق لأن كل أسرة معيشية تعيش، بحكم التعريف، في منطقة تعداد واحدة فقط لا غير. وهو، علاوةً على ذلك، حديثٌ إلى حدٍ معقول بمعنى أنه يستند إلى أحدث تعداد للسكان، شريطة ألا يكون قد حدث أيّ تغيير، بعد إجراء التعداد، في تعريف مناطق التعداد. افترض كذلك أن عمر التعداد سنتان اثنتان. لذلك تقرّر أن من الضروري وضع قائمة أحدث بالأسر المعيشية التي هي في مناطق التعداد المأخوذة منها العيّنة، بدلاً من استخدام قائمة أسر معيشية من تعداد مضى عليه سنتان. متوسط حجم منطقة التعداد ٢٠٠ أسرة معيشية، لكن يُراد أن يكون حجم المجموعة المرغوب فيها لإجراء المقابلات ١٥ أسرة لكل مجموعة. يرى فريق الاستقصاء أن تكلفة وضع قائمة من ٢٠٠ أسرة لكل ١٥ أسرة تدخل في نهاية المطاف في عيّنة (أي بنسبة ١٣ إلى ١) ستكون باهظة. عندئذٍ يقرر أخذ العيّنات أن ينفذ عملية ميدانية أرخص تكلفة، تُقسم بموجبها كل منطقة تعداد تؤخذ منها العيّنة إلى أربعة أقسام متساوية الحجم تقريباً يوجد في كل ربع ٥٠ أسرة معيشية. ثم تعدل خطة المعاينة لتستتبع اختيار ربع واحد، أو قطاع واحد، من كل منطقة تعداد تؤخذ منه عيّنة لإجراء عملية إعداد القائمة، وبذلك تقليل عبء العمل ثلاثة أرباع. يوجد لدينا في هذا التصميم ثلاث مراحل: المرحلة الأولى هي اختيار مناطق التعداد؛ والمرحلة الثانية هي اختيار قطاعات من منطقة التعداد؛ والمرحلة الثالثة هي اختيار الأسر.

### ٣ - ٦ - ٢ استخدام المراحل الصورية

٩٧ - ما يُسمّى بالمراحل الصورية هذه تُستخدم في اختيار العيّنات لتجنّب الاضطرار إلى أخذ عيّنة من ملف ضخم من الوحدات في المرحلة قبل النهائية. يمكن أن يحتوي الملف على وحدات كثيرة جداً وقد تكون ضخمة جداً بحيث يصبح من غير المعقول إدارتها باختيار يدويٍّ مُملٍّ. حتى وإن كان الملف مُعداً بالحاسوب، يمكن أن يكون مع ذلك كبيراً جداً إلى حدٍ تتعذر معه إدارة الملف بكفاءة لاختيار العيّنات<sup>١١</sup>. فالمرحلة الصورية تمكّن المرء من تضيق العالم الفرعي إلى أرقام أكثر قابلية للإدارة بالاستفادة من طبيعة التسلسل الهرمي للتقسيمات الإدارية الفرعية للبلد.

<sup>١١</sup> غير أن من الممكن أن يكون الملف الحاسوبي الكبير جداً أكثر قابلية للإدارة لأخذ العيّنات بطرق منها، مثلاً، تقسيم الملف إلى ملفات فرعية منفصلة لكل طبقة أو منطقة إدارية (كالمنطقة أو الإقليم، مثلاً).

٩٨ - لإجراء استقصاءات ريفية في بنغلاديش، مثلاً، تُعيّن القرى، في الغالب، بأنها تنتمي إلى المرحلة قبل الأخيرة من مراحل الاختيار. ومما يذكر أنه يوجد أكثر من ١٠٠ ٠٠٠ قرية في بنغلاديش، وهذا عدد أكبر كثيراً من أن يُدار بكفاءة لاختيار العينات. وإذا صُمِّمت خطة معاينة يُختار بها ٦٠٠ قرية في المرحلة قبل النهائية، مثلاً، ستُختار قرية واحدة فقط من كل ١٦٧ قرية تقريباً في بنغلاديش. لتخفيض أحجام الملفات لاختيار العينات، ربما يُقرَّر اختيار العينة في مراحل باستخدام التسلسل الهرمي للوحدات الجغرافية التي تنقسم إليها بنغلاديش — الثانات، والاتحادات، والقرى. ستمضي عملية اختيار العينات في خطوات باختيار ٦٠٠ ثانة في الخطوة الأولى، باستخدام احتمال متناسب مع أحجامها (تُناقش هذه الطريقة بالتفصيل في الفرع ٣ - ٧). وتكون الخطوة التالية اختيار اتحاد واحد فقط من عينة من الثانات، وهنا أيضاً يُستخدم الاحتمال متناسب مع حجم الاتحاد؛ وبذلك يصبح عندنا ٦٠٠ اتحاد في العينة. الخطوة الثالثة، هي اختيار قرية واحدة، باستخدام الاحتمال متناسب مع الحجم، من كل عينة من الاتحادات، وتسفر العملية هنا أيضاً عن اختيار ٦٠٠ قرية. وأخيراً، يتم اختيار عينات الأسر المعيشية من كل عينة من القرى وهذا يُنتج بوجه عام عينة منهجية تشمل جميع الأسر الموجودة في كل عينة من القرى.

٩٩ - منهجية اختيار العينات، الوارد وصفها أعلاه، هي في الواقع عملية ذات مرحلتين لاختيار عينات القرى والأسر المعيشية، مع أن مرحلتين صوريّتين قد استُخدمتا في البداية لاختيار الثانات والاتحادات التي ستُختار منها القرى. ومن الضروري في هذه الحالة توضيح الطابع الصوري للمرحلتين الأوليين رياضياً، بفحص الاحتمالات في كل مرحلة اختيار والاحتمال الإجمالي.

### ٣ - ٦ - ٢ - ١ المرحلة الأولى من مراحل الاختيار: الثانات

١٠٠ - تُختار الثانات ذات الاحتمال متناسب مع الحجم. ويُعطى الاحتمال في هذه المرحلة بـ

(٣ - ١١)

$$P_1 = \frac{(am_i)}{\sum m_i}$$

حيث  $P_1$  هي احتمال اختيار ثانة معينة؛ و  $a$  هي عدد الثانات المراد اختيارها (٦٠٠ في هذا الشكل)؛ و  $m_i$  هو عدد الأسر المعيشية الريفية<sup>١٢</sup> في الثانة الـ  $i$  وفقاً لإطار أخذ العينات المستخدم (أحدث تعداد سكان أجري، مثلاً).

١٠١ - العامل  $\sum m_i$  هو مجموع عدد الأسر الريفية في جميع الثانات في البلد. ويجب ملاحظة أن عدد الثانات المختارة فعلاً ربما يكون أقل من ٦٠٠. ويمكن أن يحدث هذا عندما تُختار ثانة أو أكثر مرتين، ويمكن أن يحدث هذا لأي ثانة يزيد مقاس حجمها عن فترة المعاينة. وتُحسب فترة معاينة الثانات بالمعادلة  $\sum m_i \div a$ . فإذا كانت فترة المعاينة ١٢ ٥٠٠، مثلاً، وكانت الثانة تحتوي على ١٣ ٨٠٠ أسرة معيشية فهي ستُختار تلقائياً مرة واحدة، ويكون أمامها فرصة ١٢ ٥٠٠ / ١٣ ٨٠٠ لأن تختار مرتين (بسط الكسر يساوي ١٣ ٨٠٠ - ١٢ ٥٠٠).

### ٣ - ٦ - ٢ - ٢ المرحلة الثانية من مراحل الاختيار: الاتحادات

١٠٢ - في المرحلة الثانية يُختار اتحاد واحد من كل عينة من عينات الثانات، وهنا أيضاً يتم الاختيار بتناسب الاحتمال مع الحجم. يتحقق هذا عملياً بوضع قائمة بجميع الاتحادات الموجودة في الثانة المختارة، وجمع مقاسات أحجامها،  $m_u$ ، واختيار رقم عشوائي بين ١ و  $m_u$ ، وهو مقاس حجم العينة المختارة من الثانات. وإن رقم الاتحاد

<sup>١٢</sup> هذا هو مقاس الحجم وربما يكون، بدلاً من ذلك، عدد سكان الثانة، شريطة أن يكون الرقم المستخدم ثابتاً في جميع مقاسات الحجم في كل مرحلة.

المشمول بهذه المجموعة، الذي تكون قيمته أصغر رقم مساوٍ للرقم العشوائي أو أكبر منه، هو الذي يُعرّف الاتحاد المختار (أو يُستخدم عرفاً كمكافئ لتعريف الاتحاد المختار). وإذا كانت ثالثة قد اختيرت أكثر من مرة في المرحلة الأولى، عندئذٍ يُختار منها نفس العدد من الاتحادات. ويعطى الاحتمال في المرحلة الثانية بالمعادلة

$$P_2 = (1) \binom{m_u}{m_t} / m_t \quad (١٢ - ٣)$$

حيث يكون  $P_2$  احتمال اختيار اتحاد معين من عيّنة الثانات؛ ويشير (١) إلى أن اتحاداً واحداً فقط قد اختير؛ ويكون  $m_u$  عدد الأسر في الاتحاد ذي الرقم  $u^{th}$  وفقاً للإطار.

### ٣ - ٦ - ٢ - ٣ المرحلة الثالثة من مراحل الاختيار: القرى

١٠٣ - تُختار في المرحلة الثالثة قرية واحدة من كل عيّنة من الاتحادات، يكون احتمال اختيارها متناسباً مع الحجم. ويعطى الاحتمال في المرحلة الثالثة كما يلي:

$$P_3 = (1) \binom{m_v}{m_u} / m_u \quad (١٣ - ٣)$$

حيث يكون  $P_3$  احتمال اختيار قرية معيّنة في عيّنة الاتحادات؛ ويشير (١) إلى أن قرية واحدة قد اختيرت؛ وأن  $m_v$  هو عدد الأسر المعيشية في القرية رقم  $v^{th}$  وفقاً للإطار.

### ٣ - ٦ - ٢ - ٤ المرحلة الرابعة من مراحل الاختيار: الأسر المعيشية

١٠٤ - نفترض في المرحلة الرابعة أن تكون القائمة الإطارية للأسر المعيشية متاحة لكل قرية مختارة، بحيث يتسنى اختيار عيّنة الأسر المعيشية اختياراً منهجياً من تلك القوائم. يُختار عدد محدد من الأسر المعيشية من كل عيّنة من عيّنات القرى؛ وهذا العدد هو حجم المجموعة المقرر سلفاً. ويعطى الاحتمال في المرحلة الرابعة كما يلي:

$$P_4 = (b) / m_v \quad (١٤ - ٣)$$

حيث  $P_4$  هو احتمال اختيار أسرة معيّنة من عيّنة القرى؛ و  $b$  هو العدد المحدد للأسر المعيشية المختارة من كل قرية.

### ٣ - ٦ - ٢ - ٥ الاحتمال الإجمالي للاختيار

١٠٥ - الاحتمال الإجمالي، الذي هو حاصل الاحتمالات في كل مرحلة، يُعطى كما يلي:

$$P = P_1 P_2 P_3 P_4 \quad (١٥ - ٣)$$

وبالتبديل يصبح لدينا

$$P = [(am_t) / \sum m_t] [(1)(m_u) / m_t] [(1)(m_v) / m_u] [b / m_v] = [(a)(b) / \sum m_t] \quad (١٦ - ٣)$$

١٠٦ - لاحظ أن العاملين  $P_2$  و  $P_3$  يلغي بعضهما بعضاً إلغاءً كلياً، مما يبين الطبيعة الصورية لعملية الاختيار على "أربع" مراحل. وهكذا، لا تعدو الثانات والاتحادات، وإن "اختبرت" مادياً، أن تكون وسيلة لتحديد مواقع القرى التي ستختار كعينات.

### ٣ - ٦ - ٣ التصميم على مرحلتين

١٠٧ - أولي اهتمام كثير في الآونة الأخيرة لاستخدام تصاميم اختيار العينات على مرحلتين في البلدان النامية. فهو التصميم المختار للمعاينة المستخدمة في استقصاءات المجموعات ذات المؤشرات المتعددة، الذي أجرته منظمة الأمم المتحدة للطفولة في أكثر من ١٠٠ بلد منذ أواسط التسعينات من القرن المنصرم. وهي تُستخدم بصورة رئيسية أيضاً في الاستقصاءات الديمغرافية والصحية.

١٠٨ - يتألف التصميم على مرحلتين في الأحوال النمطية ببساطة من اختيار عينة يكون احتمال اختيارها متناسباً مع الحجم في بضع مئات من الوحدات الجغرافية، تكون مقسمة تقسيماً مناسباً إلى طبقات، في المرحلة الأولى. وقد توضع قائمة حديثة بالأسر المعيشية في المرحلة الأولى من مراحل اختيار وحدات العينات، وهذا يتوقف على توفر المعلومات المتعلقة بعناوين و/أو مواقع الأسر المعيشية، وما إذا كانت هذه المعلومات حديثة أم لا. وتلي ذلك عينة منهجية لعدد محدد من الأسر المعيشية في المرحلة الثانية. وتُعرف الوحدات الجغرافية، المشار إليها بوجه عام بكلمة "المجموعات"، في العادة كالقرى أو مناطق تعداد السكان في المناطق الريفية أو الأحياء في مدينة في المناطق الحضرية.

١٠٩ - التصميم على مرحلتين، الوارد وصفه أعلاه، تصميم مُحَبَّب من عدة وجوه، لكن أكثر ما يحبُّه للناس هو بساطته. فمن الميزات المألوفة دائماً في تصميم العينات الاجتهاد في الحصول على البساطة لا على التعقيد بغية تخفيض احتمالات وقوع أخطاء من غير أخطاء المعاينة أثناء تنفيذ عملية أخذ العينات. التصميم على مرحلتين له جوانب مفيدة تجعله بالمقارنة بسيطاً ومرغوباً. من الأمثلة على ذلك:

- تصميم العينات، كما ورد في وصفه، فيه مُعامل ترجيح ذاتي (جميع الأسر التي في العينة تُختار بنفس الاحتمال)، أو ترجيح ذاتي تقريباً (انظر الفرعين ٣ - ٧ - ١ و ٣ - ٧ - ٢ للاطلاع على التمييز بين العينات المختارة باحتمال متناسب مع الحجم والعينات المختارة باحتمال متناسب مع الحجم المقدّر).
- المجموعات المعروفة بأنها مناطق تعداد أو أحياء في مدينة ذات حجم ملائم (ليست أكبر من اللازم) — في معظم البلدان، لا سيما إن كان يجب وضع قوائم جديدة بالأسر المعيشية قبل مرحلة الاختيار الأخيرة.
- جرت العادة بأن تكون قد وضعت خرائط لمناطق التعداد وأحياء المدينة ومعظم القرى، إما لأغراض عمليات التعداد وإما لأغراض أخرى، وتكون لها حدود مرسومة جيداً.

### ٣ - ٧ - ٣ المعاينة باحتمال متناسب مع الحجم وباحتمال متناسب مع الحجم المقدّر

١١٠ - عرض الفرع ٣ - ٥ مثلاً أبرز فيه الاحتمال المتناسب مع الحجم في اختيار المجموعات التي تختار منها العينات. ويبحث هذا الفرع اختيار العينات باحتمال متناسب مع الحجم بمزيد من التفصيل.

## ٣ - ٧ - ١ المعاينة باحتمال متناسب مع الحجم

١١١ - استعمال طريقة المعاينة باحتمال متناسب مع الحجم يسمح لأخذ العيّنات بممارسة تحكّم أكبر في حجم العيّنة التي تُختار في نهاية المطاف لاستقصاء المجموعات. وفي الحالات التي تكون فيها جميع المجموعات ذات حجم واحد، أو تكون ذات حجم واحد تقريباً، لا توجد ميزة لاستخدام طريقة المعاينة باحتمال متناسب مع الحجم. افترض، مثلاً، أن كل حي في المدينة تسكنه ١٠٠ أسرة معيشية وأريدت عيّنة من ١٠٠٠ أسرة منتشرة في عيّنة من ٥٠ حياً في المدينة. الخطة البديهية للمعاينة هنا هي اختيار عيّنة عشوائية من ٥٠ حياً، أي عيّنة متساوية الاحتمالات ثم اختيار واحدة من كل خمس أسر معيشية بالضبط من كل حي (وهنا أيضاً عيّنة متساوية الاحتمالات). وستكون النتيجة عيّنة تضم ٢٠ أسرة بالضبط من كل حي، أو ما مجموعه ١٠٠٠ أسرة. وتكون معادلة الاختيار في هذه الحالة كما يلي:

$$P=(50/M)(1/5)$$

حيث يكون العامل  $p$  احتمال اختيار أسرة؛ و  $(50/M)$  احتمال اختيار حي؛ و  $M$  مجموع عدد الأحياء في المدينة؛ و  $(1/5)$  احتمال اختيار أسرة من عيّنة من حي مشمول بالعيّنة.

١١٢ - العامل  $P$  يخفض إلى  $10/M$ . ولما كان العامل  $M$  ثابتاً، يكون الاحتمال الإجمالي لاختيار كل عيّنة من عيّنات الأسر المعيشية ١٠ مقسوماً على عدد الأحياء،  $M$ .

١١٣ - غير أن الأحياء أو الوحدات الجغرافية الأخرى، التي تستخدم كمجموعات لاستقصاء الأسر المعيشية، نادراً ما تكون في الأوضاع الحقيقية متشابهة في أحجامها إلى هذا الحد. ففي المثال الوارد أعلاه، ربما تتراوح أحجامها من ٢٥ إلى ٢٠٠. وأخذ العيّنة القائم على تساوي الاحتمالات يمكن أن يسفر عن اختيار "غير موفق" تكون فيه معظم العيّنات صغيرة الحجم أو كبيرة الحجم. وفي هذه الحالة، تكون النتيجة إما عيّنة إجمالية يختلف حجمها اختلافاً كبيراً عن العيّنة المرغوبة التي يبلغ حجمها ١٠٠٠ أسرة معيشية، التي بُحثت في المثال. يوجد أسلوب لتقليل احتمالات وجود عيّنة متفاوتة أحجامها تفاوتاً كبيراً هو لإنشاء طبقات تستند إلى حجم المجموعات واختيار عيّنة من كل طبقة. لا يوصى بهذه الطريقة بوجه عام لأنه ربما يخفض أو يعقد استخدام العوامل الأخرى للتصنيف إلى طبقات في تصميم العيّنة. وأخذ العيّنات بناءً على الاحتمال متناسب مع الحجم هو الطريقة المفضل لأنه يمكن من التحكم تحكماً أكبر بحجم العيّنات في نهاية المطاف دون إدخال الحاجة إلى التقسيم إلى طبقات بناءً على الحجم.

١١٤ - لتوضيح أخذ العيّنات مع الاحتمال متناسب مع الحجم نبدأ بمعادلة الاختيار المذكورة أعلاه لكن نعبر عنها بصورة أكثر رسمية لتصميم على مرحلتين<sup>١٣</sup> كما يلي:

(٣ - ١٧)

$$P(\alpha\beta) = P(\alpha)P(\beta|\alpha)$$

حيث العامل  $P(\alpha\beta)$  هو احتمال اختيار الأسرة المعيشية  $\beta$  في المجموعة  $\alpha$ ؛ والعامل  $P(\alpha)$  هو احتمال اختيار المجموعة  $\alpha$ ؛ والعامل  $P(\beta|\alpha)$  هو الاحتمال المشروط لاختيار الأسرة المعيشية  $\beta$  في المرحلة الثانية، على افتراض أن المجموعة  $\alpha$  قد اختيرت في المرحلة الأولى.

<sup>١٣</sup> انظر: (Kalton, 1983, pp. 38-47) لتطوير هذا المفهوم والاطلاع على بحث إضافي متعلق باختيار العيّنات على أساس الاحتمال متناسب مع الحجم.

١١٥ - لتثبيت حجم العينة الإجمالي من حيث عدد الأسر المعيشية، نحتاج إلى عينة متساوية الاحتمالات تضم  $n$  أسر معيشية من مجموعة سكانية تضم  $N$  أسر معيشية. وبذلك يكون المعدل الإجمالي للعينات  $n/N$  وهذا يساوي  $P(\alpha\beta)$  كما هو معرف أدناه. علاوة على ذلك، إذا كان عدد المجموعات المراد أخذ عينات منها محددًا بـ  $\alpha$  نحتاج في الوضع النموذجي إلى اختيار أسرة معيشية من كل مجموعة، بغض النظر عن أحجام المجموعات المختارة. وإذا عرفنا  $m_i$  بأنه حجم المجموعة الـ  $i$ ، فنحن نحتاج إلى  $P(\beta|\alpha)$  أن يكون مساويًا لـ  $b/m_i$ . ومن ثمَّ

$$P(\alpha\beta) = [P(\alpha)] [b/m_i]$$

ولما كان  $n = ab$  يكون لدينا

$$.ab/N = [P(\alpha)] [b/m_i]$$

وبحل المعادلة الأخيرة لـ  $P(\alpha)$ ، نحصل على

(١٨ - ٣)

$$P(\alpha) = (a)(m_i)/N$$

١١٦ - لاحظ أن  $N = \sum m_i$ ، بحيث يكون احتمال اختيار مجموعة لذلك متناسبًا مع حجمها. لذلك، فإن معادلة اختيار عينة يكون فيها الاحتمال متناسبًا مع الحجم في وحدات المرحلة الأولى التي تُختار فيها الوحدات النهائية مع ذلك اختيارًا متساوي الاحتمالات، تكون لذلك

(١٩ - ٣)

$$P(\alpha\beta) = [(a)(m_i)/\sum m_i] [b/m_i]$$

(٢٠ - ٣)

$$= [(ab)/\sum m_i]$$

١١٧ - تصميم العينة الذي نحققه بهذه الطريقة تصميم ذاتي الترتيب، كما يمكن أن يُرى من المعادلة (٣ - ١٩)، لأن جميع أرقام المعادلة ثابتة؛ تذكر أنه مع كون العامل  $m_i$  متغيرًا، يكون المجموع،  $\sum m_i$ ، ثابتًا ومساويًا لـ  $N$ . يعرض الشكل ٣ - ٢ أدناه مثالًا لكيفية اختيار عينة من المجموعات باستخدام الاحتمال المتناسب مع الحجم.

١١٨ - فيما يتعلق بالاختيار المادي للعينة، لاحظ أن فترة المعاينة،  $I$ ، في الشكل ٣ - ٢، تضاف تتابعياً إلى البدء العشوائي سبع مرات (أو  $a-1$  مرات، حيث  $a$  هو عدد المجموعات المراد اختيارها). وتكون أرقام الاختيار الناتجة عن ذلك ٣١١،٢ (الذي هو البداية العشوائية)، و٨٧٨،٨، و١٤٤٦،٤، و٢٠١٤، و٥٨١،٦٣، و١٤٩،٢، و٣٧١٦،٨، و٤٢٨٤،٤. والمجموعة التي تؤخذ منها العينة لأرقام الاختيار الثمانية هذه هي، في كل حالة، العينة التي يكون مقاس حجمها المتراكم أصغر قيمة مساوية لرقم الاختيار أو أكبر منه. وهكذا، تُختار المجموعة ٠٣ لأن الرقم ٣٧٧ هو أصغر رقم متراكم يساوي الرقم ٣١١،٢ أو يزيد عنه، وتُختار المجموعة ٢٦ لأن الرقم ٣٧٤٤ هو أصغر رقم تراكمي يساوي الرقم ٣٧١٦،٨ أو يزيد عنه.

١١٩ - مع أن التوضيح لا يُظهر هذا بشكل قاطع (لأن ثمانى مجموعات فقط قد اختيرت)، يميل أخذ العينات باحتمال متناسب مع الحجم يميل إلى اختيار مجموعات أكبر لا أصغر. ربما يكون هذا بديهياً، لأنه يمكن أن يُرى من المعادلة (٣ - ١٧) أن احتمال اختيار مجموعة متناسب مع حجمها؛ وهكذا يرجح أن يكون احتمال اختيار العينة التي تضم ٢٠٠ أسرة معيشية ضعف احتمال اختيار عينة تضم ١٠٠ أسرة معيشية. ومن ثم يجب ملاحظة أن نفس

المجموعة يمكن أن تُختار أكثر من مرة إذا كان مقياس حجمها يزيد عن فترة المعاينة،  $I$ . غير أن أيًا من المجموعات الواردة في الشكل لا تناسب هذا الشرط؛ لكن إذا حدث ذلك، فإن عدد الأسر المعيشية التي ستُختار من هذه المجموعة سيتضاعف إن كان الاختيار "مرتين" وسيكون ثلاثة أضعاف إن كان الاختيار ثلاث مرات، وهلم جرًا.

الشكل ٣ - ٢

مثال للاختيار المنهجي للمجموعات، باحتمال متناسب مع الحجم

اختيار العينة	العدد التراكمي	مقياس الحجم (عدد الأسر المعيشية)	المجموعة / عدد وحدات المعاينة الأولية
	٢١٥	٢١٥	٠٠١
	٢٨٨	٧٣	٠٠٢
٣١١,٢	٣٧٧	٨٩	٠٠٣
	٦٠٨	٢٣١	٠٠٤
	٧٢٨	١٢٠	٠٠٥
	٧٨٦	٥٨	٠٠٦
٨٧٨,٨	٨٨٥	٩٩	٠٠٧
	١٠٥٠	١٦٥	٠٠٨
	١٢٤٥	١٩٥	٠٠٩
١٤٤٦,٤	١٤٤٧	٢٠٢	٠١٠
	١٥٢٤	٧٧	٠١١
	١٥٨٣	٥٩	٠١٢
	١٨٢٨	٢٤٥	٠١٣
	١٩٩٩	١٧١	٠١٤
٢٠١٤,٠	٢٠٩٨	٩٩	٠١٥
	٢١٨٦	٨٨	٠١٦
	٢٣١٠	١٢٤	٠١٧
	٢٣٨٨	٧٨	٠١٨
	٢٤٧٧	٨٩	٠١٩
	٢٥٢٧	٦٠	٠٢٠
٢٥٨١,٦	٢٧٥٩	٢٢٢	٠٢١
	٢٨٩٦	١٣٧	٠٢٢
	٣٠٩٥	١٩٩	٠٢٣
٣١٤٩,٢	٣٣٠٥	٢١٠	٠٢٤
	٣٤٧٠	١٦٥	٠٢٥
٣٧١٦,٨	٣٧٤٤	٢٧٤	٠٢٦
	٣٩٥٣	٢٠٩	٠٢٧
	٤١٨٣	٢٣٠	٠٢٨
	٤٢٥٠	٦٧	٠٢٩
٤٢٨٤,٤	٤٣٢٢	٧٢	٠٣٠
	٤٤٣٠	١٠٨	٠٣١
	٤٥٤١	١١١	٠٣٢

**تعليمات بشأن العيّنات:** اختر ٨ وحدات معاينة أولية (مجموعات) من ٣٢ وحدة في العالم، مستخدماً احتمالاً متناسباً مع الحجم؛ لذلك تساوي فترة المعاينة ( $I$ ) ٥٤١/٨، أو ٥٦٧,٦، حيث العدد ٤٥٤١ هو مجموع المقاس التراكمي لأحجام جميع المجموعات، والعدد ٨ هو عدد المجموعات المراد اختيارها؛ البدء العشوائي هو عدد عشوائي يقع بين ٠,١ و ٥٦٧,٦ يختار من قائمة أعداد عشوائية؛ والبدء العشوائي في هذا الشكل يساوي ٣١١,٢.

### ٣ - ٧ - ٢ المعاينة باحتمال متناسب مع الحجم المقدّر

١٢٠ - منهجية أخذ العينات مع كون الاحتمال متناسباً مع الحجم الوارد وصفها في الفرع السابق، منهجية مثالية إلى حد ما وربما لا يمكن تحقيقها في التطبيقات العملية في معظم الحالات. وذلك لأن مقياس الحجم المستخدم لإثبات احتمال اختيار المجموعة في المرحلة الأولى غالباً ما يكون غير المقاس الفعلي للحجم حين تُختار عينة الأسر المعيشية في المرحلة الثانية.

١٢١ - مقياس الحجم المعتمد بوجه عام في استقصاءات الأسر المعيشية، لاختيار وحدات المعاينة الأولية أو المجموعات في المرحلة الأولى هو عدد الأسر المعيشية (أو المجموعة السكانية) الوارد من أحدث عملية تعداد. حتى لو كان التعداد حديثاً جداً، فإن العدد الفعلي للأسر المعيشية في وقت إجراء الاستقصاء يرجح أن يكون مختلفاً، حتى ولو برقم صغير. غير أن استثناءً يحدث حينما يُؤخذ اختيار الأسر المعيشية في المرحلة الثانية مباشرة من الإطار نفسه الذي استخدم لإثبات مقاسات الحجم (للاطلاع على مناقشة أطر المعاينة بمزيد من التفصيل، انظر الفصل الرابع).

#### مثال

افترض أن استقصاءً للأسر المعيشية أُجري بعد ثلاثة أشهر من الانتهاء من إجراء تعداد السكان. وقرر فريق الاستقصاء، بدلاً من وضع قوائم جديدة بالأسر المعيشية في المجموعات المختارة، أن يستخدم قائمة الأسر المستمدة من التعداد، في المرحلة الثانية من عملية أخذ عينات على مرحلتين، لأنه افترض، والافتراض معقول، أن قائمة التعداد حديثة ودقيقة لكل الأغراض العملية. في المرحلة الأولى، اختيرت عينة من القرى باستخدام أعداد الأسر المأخوذة من التعداد باعتباره مقياساً للحجم في كل قرية. وفي كل قرية اختيرت كعينة كان مقياس الحجم،  $m_i$ ، مماثلاً للعدد الفعلي للأسر المعيشية التي ستُختار منها العينة. وهكذا، إذا اختيرت القرية ألف وكان فيها ٢٣٥ أسرة معيشية، وفقاً لما جاء في التعداد، فإن القائمة التي ستُختار منها عينة الأسر للاستقصاء ستضم هي أيضاً ٢٣٥ أسرة.

١٢٢ - غير أن الاستقصاء، في كثير من تطبيقات استقصاءات الأسر المعيشية المستندة إلى أطر التعداد، كان يُجرى بعد بضعة أشهر وأحياناً بعد سنوات من إجراء التعداد (انظر الفصل الرابع للاطلاع على مزيد من البحث وتحديث أطر أخذ العينات). وغالباً ما يتقرر في هذه الظروف إجراء عملية ميدانية بغية إعداد قائمة جديدة بالأسر المعيشية في المجموعات التي اختيرت لشمولها بالعينة في المرحلة الأولى. وعندئذٍ تختار عينة من القوائم الجديدة للاستقصاء.

١٢٣ - مقياس الحجم،  $m_i$ ، المستخدم لاختيار المجموعة هو عدد الأسر المعيشية المأخوذ من عملية التعداد التي بُحثت في المثال الوارد أعلاه. غير أن القائمة الفعلية التي ستُختار منها عينة الأسر المعيشية ستكون مختلفة. وسيكون لها طبعاً مقياس للحجم مختلف إلى درجة ما، وذلك يتوقف على طول الفترة الزمنية التي انقضت بين إجراء التعداد وإعداد القائمة للاستقصاء. وستحدث الفروقات بسبب الهجرة إلى المجموعة أو منها، وبناء مساكن جديدة أو هدم مساكن قديمة، وإنشاء أسر معيشية جديدة عندما يحدث زواج (أحياناً تقيم الأسرة الجديدة في نفس وحدة السكن التي تسكن فيها الأسرة الوالدة)، وكذلك بسبب الوفيات. عندما تُختار العينة مع كون الاحتمال متناسباً مع الحجم المقدّر، سيكون احتمالها المأخوذ من معادلة الاختيار كما يلي:

(٣ - ٢١)

$$P(\alpha\beta) = \left[ \frac{a(m_i)}{\sum m_i} \right] \left[ \frac{b}{m_i} \right]$$

حيث يكون  $m_i$  عدد الأسر المعيشية بناءً على عملية إعداد القائمة وتُعرف الأحكام الأخرى كما عُرِّفت من قبل.

١٢٤ - نظراً إلى أنه يرجح أن تكون  $m_i$  و  $m_j$  مختلفتين في معظم عيّنات المجموعات إن لم يكن جميعها، ينبغي لحساب احتمال الاختيار (ومن ثم مُعامل الترجيح، أي عكس الاحتمال، مثلاً) أن يأخذ الفرق في الحسبان. وكما تبين المعادلة ٣ - ٢٠، سيكون لكل مجموعة مُعاملٍ ترجيحٍ مختلف، وهذا يستبعد الترجيح الذاتي في تصميم العيّنات.

١٢٥ - باستخدام مُعاملات الترجيح المضبوطة التي تعوّض الفروق بين التعداد ومقاسات الحجم في الاستقصاء، ستكون التقديرات الناتجة للاستقصاء غير متحيزة. ومن شأن عدم تعديل مُعاملات الترجيح بناءً على ذلك أن ينتج تقديرات متحيزة تزداد أحجامها بلا شك كلما طالت الفترة الفاصلة بين التعداد والاستقصاء. غير أنه ينبغي ملاحظة أنه عندما توجد اختلافات ضئيلة بين  $m_i$  و  $m_j$ ، تكون العيّنة ذاتية الترجيح تقريباً؛ وربما يكون من الحكمة، في بعض الظروف<sup>١٤</sup>، توليد تقديرات الاستقصاء دون مُعاملات ترجيح، لأن التحيزات ستكون غير ذات شأن. غير أن من الأساسي، قبل تقرير سبيل العمل هذا، فحص  $m_i$  و  $m_j$  مجموعةً فمجموعةً لتقييم الفوارق ومعرفة ما إذا كانت ضئيلة أم لا.

١٢٦ - توجد استراتيجية بديلة يمكن استخدامها لاختيار الأسر المعيشية في المرحلة الأخيرة حين تُستخدم المعاينة مع كون الاحتمال متناسباً مع الحجم المقدر — وهي استراتيجية تكون العيّنة فيها بالفعل ذاتية الترجيح. وهي تنطوي على اختيار الأسر المعيشية بنسب متفاوتة داخل كل مجموعة، ويتوقف ذلك على حجمها الفعلي (يُناقض هذا الموضوع في الفرع التالي).

### ٣ - ٨ الخيارات في المعاينة

١٢٧ - يناقش هذا الفرع بعضاً من الخيارات الكثيرة التي يمكن النظر فيها لدى تصميم عيّنة ملائمة لإجراء استقصاء للأسر المعيشية متعدد الأغراض، ويركز بالدرجة الأولى على استراتيجيات في المرحلتين الأخيرة وما قبل الأخيرة من الاختيار، لأن هاتين المرحلتين هما اللتان تتاح فيهما عدة بدائل. وهو يفحص خيار اختيار عيّنات المجموعات في المرحلة قبل الأخيرة مع تساوي الاحتمالات أو باحتمال متناسب مع الحجم، وكذلك أخذ عيّنات من الأسر المعيشية مع الخيار بين النسبة الثابتة والحجم الثابت في المرحلة الأخيرة، وإلى حد ما، يوجز الأجزاء السابقة فيما يتعلق بمسائل التحكم بحجم العيّنة، والتصميم ذاتي الترجيح والتصميم غير ذاتي الترجيح، وكذلك مسائل أخرى مثل أعباء عمل المقابلين. ويستعرض، بالإضافة إلى ذلك، تصاميم خاصة تُستخدم الآن على نطاق واسع، كالتصاميم المستخدمة في الاستقصاءات الديمغرافية والصحية واستقصاء اليونيسيف الذي أُجري في منتصف العقد للمجموعات المتعددة المؤشرات. تقدم هذه التصاميم خيارات إضافية تستحق أن يُنظر فيها، بما في ذلك استخدام المجموعات الصغيرة (التي تؤخذ كلها) وغير الصغيرة.

### ٣ - ٨ - ١ المعاينة مع تساوي الاحتمالات، والمعاينة باحتمال متناسب مع الحجم، والمعاينة مع الحجم الثابت ومع النسبة الثابتة

١٢٨ - في الجدول ٣ - ٤، تقدم التصاميم المحتملة إطاراً لبحث الإجراءات والشروط والميزات وأوجه القصور لمختلف خطط المعاينة.

<sup>١٤</sup> في الاستقصاءات التي تقتصر فيها التقديرات على النسب والمعدلات والنسب المئوية تكون هذه استراتيجية ملائمة؛ أما في الاستقصاءات التي يراد فيها أن تكون المجاميع مقدرة أو مطلقة فيجب استخدام معاملات الترجيح بغض النظر عما إذا كانت العيّنة ذاتية الترجيح أم لا.

الجدول ٣ - ٤

## الخطط البديلة للمعاينة: المرحلتان الأخيرتان من مراحل الاختيار

اختيار الوحدات قبل النهائية	الحجم الثابت للمجموعات (عدد الأسر المعيشية)	النسبة الثابتة للاختيار في كل مجموعة
الاحتمال المتناسب مع الحجم	الخططة ١	الخططة ٢ [لا يُوصى بها]
الاحتمال المتناسب مع الحجم المقدر	الخططة ٣	الخططة ٤ [لا يُوصى بها]
تساوي الاحتمالات	الخططة ٥	الخططة ٦

١٢٩ - ناقشنا كيف تكون المعاينة، باحتمال متناسب مع الحجم، لوحدة المعاينة الأولية أو المجموعات، كوسيلة للتحكم بحجم العينة في نهاية المطاف تحكماً أكثر دقة مما هو في حالة المعاينة مع تساوي الاحتمالات، وهذه هي ميزته الرئيسية، لا سيما إذا كانت المجموعتان واسعتي التباين في عدد الأسر المعيشية التي تضمها كل مجموعة. التحكم في حجم العينة هام ليس فقط لآثاره على الميزانية وإنما أيضاً لأنه يمكن مدير الاستقصاء من التخطيط بدقة لأعباء عمل المقابلين قبل بدء عمليات الاستقصاء. اختيار العينات مع تساوي الاحتمالات، من جهة أخرى، أسهل إجراءً من المعاينة باحتمال متناسب مع الحجم، ويكون له ما يبرره عندما تكون مقاسات أحجام العينات متساوية تقريباً أو تكون الفروق بينها ضئيلة. وينبغي، من ناحية عملية، استخدام المعاينة باحتمال متناسب مع الحجم المقدر، بدلاً من كونه متناسباً مع الحجم، حيثما كان المقاس الفعلي للحجم مختلفاً عن الحجم المعطى في الإطار.

١٣٠ - اختيار عدد ثابت من الأسر المعيشية في كل عينة من المجموعات له ميزتان هامتان جداً: الأولى أنه يمكن التحكم بحجم العينة بدقة؛ والثانية أن الطريقة تقدم لمدير الاستقصاء وسيلة لإعطاء المقابلين أعباء عمل مضبوطة وأن يساوي بين هذه الأعباء إن شاء. غير أن المعاينة بأحجام ثابتة معقدة إلى حد ما، لأنها تستدعي حساب فترات معاينة مختلفة لكل مجموعة. ومن شأن استخدام فترات معاينة مختلفة أن يؤدي إلى التشويش ويكون عرضة للخطأ. غير أنه توجد اختبارات متأصلة في التصميم لاختبار التحكم بالجودة، لأن عدد الأسر المعيشية المراد اختيارها معروف مسبقاً. ومع ذلك يمكن أن تسفر التعقيدات عن قلة في الكفاءة ناشئة من الوقت المفقود، نظراً إلى الاضطرار إلى تصحيح الأخطاء في الاختيار.

١٣١ - تستدعي المعاينة بحجم ثابت، بحكم التعريف، وضع قوائم بالأسر المعيشية، يمكن بناءً عليها تعيين وتعريف الأسر المعيشية المختارة. في معظم الأحيان، تكون القوائم حديثة جداً لأنها تُعد كجزء من العمليات الميدانية السابقة لإجراء الاستقصاء. ومن المفيد أن يُضمن اختيار عينات الأسر المعيشية في مكتب مركزي وأن يجريه شخص غير الشخص الذي وضع القائمة، لتقليل إمكانية حدوث تحيز في إجراء الاختيار إلى أدنى حد ممكن.

١٣٢ - يمكن، كبديل عن ذلك، أخذ عينات من الأسر المعيشية بنسبة ثابتة في كل مجموعة. في هذه الحالة يكون الاختيار أسهل وأقل عرضة للخطأ. وتوجد ميزة في الميدان، وهي أنه يمكن المعاينة في الوقت الذي يكون فيه المقابل يجوب أنحاء المجموعة سعياً إلى الحصول على قوائم حديثة بالأسر المعيشية. ويمكن إنجاز ذلك بتصميم استمارة القوائم على نحو يظهر فيه خطوطاً معينة مسبقاً لتعريف الأسر المدرجة في العينة. وكون وضع القوائم والمعاينة يمكن إجراؤهما في زيارة واحدة فيه ميزات بديهية من حيث التكلفة؛ غير أنه توجد بعض أوجه القصور الهامة للعملية.

١٣٣ - من أوجه القصور التي تعوق المعاينة بنسبة ثابتة أنها لا تتحكم بحجم العيّنات أو أعباء عمل المقابلين، ما لم يكن مقياس الحجم لكل مجموعة مساوياً تقريباً لأحجام المجموعات الأخرى. وثمة وجه آخر من أوجه القصور أكثر خطورة، وهو أنه عندما يُعهد إلى المقابلين بمهمة اختيار الأسر المعيشية لإدراجها في العيّنة، بمعنى تعيين الأسر التي ستوضع على قائمة العيّنة، غالباً ما يسفر ذلك عن تحيُّز. وقد أجريت دراسات لا حصر لها أثبتت أن الأسر التي تُختار أثناء كون المقابلين مسيطرين على الموقف تميل إلى أن تكون أسراً أصغر حجماً، مما يوحي بأن المقابلين ربما يختارون عن وعي أو عن غير وعي، أسراً يكون عدد المستجيبين فيها أقل لتخفيض أعباء عملهم.

١٣٤ - يتوقف كون التصميم ذاتي الترجيح أو غير ذاتي الترجيح على المزيح المعين لإجراءات المعاينة في كل مرحلة. وهكذا يكون التصميم على مرحلتين الذي يجمع بين معاينة المجموعات باحتمال متناسب مع الحجم ومعاينة الأسر المعيشية بناءً على الحجم الثابت ذاتي الترجيح، بينما التصميم الذي يجمع بين الاحتمال المتناسب مع الحجم والنسبة الثابتة غير ذاتي الترجيح. وتبين المناقشة التالية أياً من الخطط الواردة في الجدول ٣ - ٤ تكون ذاتية الترجيح.

### ٣ - ١ - ٨ - ١ الخطة ١: الاحتمال المتناسب مع الحجم، الحجم الثابت للمجموعة

الشروط:

- مقاسات حجم متباينة لعالم المجموعات
- اختيار أسر معيشية من نفس القوائم (مثال ذلك: قوائم الأسر المعيشية التي وضعها تعداد السكان) التي تُستخدم لمقاسات الحجم

الميزات:

- التحكم بمجموع حجم العيّنة، ومن ثم بالتكلفة
- التحكم بأعباء عمل المقابلين
- الترجيح الذاتي

أوجه القصور:

- تطبيق المعاينة باحتمال متناسب مع الحجم أصعب من تطبيق المعاينة مع تساوي الاحتمالات
- نسب اختيار مختلفة لاختيار الأسر المعيشية من كل مجموعة، مع احتمال وقوع أخطاء

### ٣ - ١ - ٨ - ٢ الخطة ٢: الاحتمال المتناسب مع الحجم، النسبة الثابتة

١٣٥ - لا توجد شروط معقولة يمكن استخدام هذا التصميم بموجبها. فإذا كانت المجموعات متباينة في الحجم، ستكون المعاينة باحتمال متناسب مع الحجم مقترنةً بحجم ثابت للمجموعات هي الخطة الصحيحة للاستخدام. وإذا كانت المجموعات ذات أحجام متساوية تقريباً تكون المعاينة بنسبة ثابتة ملائمة، لكن يجب اختيار المجموعات نفسها باستخدام المعاينة مع تساوي الاحتمالات.

## ٣ - ٨ - ١ - ٣ الخطة ٣: الاحتمال المتناسب مع الحجم، الحجم الثابت للمجموعة

## الشروط

- مقاسات حجم مختلفة لعالم المجموعات
- اختيار الأسر المعيشية من قوائم جديدة مع تحديث الأسر المنقولة من الإطار لإثبات مقاسات الحجم الأصلية

## الميزات

- التحكُّم بمجموع أحجام العيِّنات ومن ثم بالتكاليف
- التحكُّم بأعباء عمل المقابلين
- أدقُّ من الاحتمال المتناسب مع الحجم لإطار معيَّن بسبب كون قوائم الأسر المعيشية حديثة

## أوجه القصور

- تطبيق الاحتمال المتناسب مع الحجم المقدرُّ أصعب من تطبيق طريقة المعاينة مع تساوي الاحتمالات
- نسب اختيار مختلفة لاختيار أسر معيشية من كل مجموعة مع احتمال وقوع أخطاء
- ليس ذاتي الترجيح

## ٣ - ٨ - ١ - ٤ الخطة ٤: تساوي الاحتمالات، النسبة الثابتة

١٣٦ - لا توجد شروط معقولة لاستخدام الخطة ٤ وذلك للأسباب المعلنة أعلاه فيما يتعلق بالخطة ٢.

## ٣ - ٨ - ١ - ٥ الخطة ٥: تساوي الاحتمالات، الحجم الثابت للمجموعات

## الشروط

- مقاسات الحجم لعالم المجموعات متساوية تقريباً أو التباين بينها صغير جداً

## الميزات

- التحكُّم بمجموع أحجام العيِّنات ( لكن أقل من التحكُّم بموجب الخطة ١ إلى حد ما) ومن ثم بالتكلفة
- التحكُّم بأعباء عمل المقابلين لكن هنا أيضاً أقل منه بموجب الخطة ١
- تساوي الاحتمالات تطبيقه أسهل من تطبيق الاحتمال المتناسب مع الحجم أو الاحتمال المتناسب مع الحجم المقدر

## أوجه القصور

- نسب اختيار مختلفة لاختيار الأسر المعيشية من كل مجموعة مع احتمال وقوع أخطاء
- ليس ذاتي الترجيح

## ٣ - ٨ - ١ - ٦ الخطة ٦: تساوي الاحتمالات، النسبة الثابتة

## الشروط

- مقاسات الحجم لعالم المجموعات متساوية تقريباً

## الميزات

- ذاتي الترجيح
- اختيار العيّنات سهل جداً في كلتا المرحلتين

## أوجه القصور

- ضعف التحكم بمجموع أحجام العيّنات، مما يعود بعواقب تتعلق بالتكلفة والموثوقية، لا سيما إذا كانت مقاسات الحجم الراهنة مختلفة كثيراً عن مقاسات الحجم المنقولة من الإطار، وعواقب تتعلق بالموثوقية إذا كانت العيّنة أصغر مما كان مستهدفاً
- التحكم بأعباء عمل المقابلين ضئيل

## ٣ - ٨ - ٢ الاستقصاء الديمغرافي والصحي

١٣٧ - مع أن محور تركيز الاستقصاء الديمغرافي والصحي هو النساء اللاتي هن في سن الحمل، فإن تصميم عيّناته ملائم للاستقصاءات ذات الأغراض العامة.

١٣٨ - إن الاستقصاء الديمغرافي والصحي، الذي يُطبّق على نطاق واسع في البلدان النامية منذ سنة ١٩٨٤، يشجع على استخدام تصميم القطاع المعياري<sup>١٥</sup>، لكونه ملائماً وعملياً في دليله لاختيار العيّنات. يعرف القطاع المعياري من حيث الحجم، ويضم عادة ٥٠٠ شخص. يُخصّص لكل وحدة منطقة جغرافية في البلد تنتمي إلى إطار المعاينة مقاس للحجم يُحسب بتقسيم عدد سكانها على ٥٠٠ (أو أيّاً كان حجم القطاع المعياري الذي يقرّر استخدامه في البلد المعني). ويكون حاصل القسمة، مقرباً إلى أقرب عدد صحيح، عدد القطاعات المعيارية في وحدة المنطقة.

١٣٩ - تُختار عيّنة لوحدات المناطق باحتمال متناسب مع الحجم، باستخدام عدد القطاعات المعيارية مقاساً للحجم. ولما كانت وحدات المناطق المستخدمة لهذه المرحلة من مراحل المعاينة في العادة مناطق تعداد، أو أحياء في مدينة، أو قرى يكون مقاس الحجم لنسبة عالية منها مساوياً لواحد أو اثنين. ولأيّ وحدة منطقة مختارة يكون مقاس حجمها أكثر من واحد تنظم عملية رسم خارطة لها تُنشأ فيها قطاعات جغرافية، ويساوي عدد هذه القطاعات الجغرافية مقاس الحجم. لذلك، ترسم خارطة لوحدة المنطقة المدرجة في العيّنة والتي يبلغ مقاس حجمها ٣، بغية تقسيم الوحدة إلى ثلاثة قطاعات بأحجام متساوية تقريباً، بقدر ما تسمح به الحدود الطبيعية، من حيث عدد الأشخاص في كل قطاع (خلافاً لمساحتها الجغرافية).

١٤٠ - كل وحدة منطقة مقاس حجمها ١ تُدخل تلقائياً في العيّنة، وفي كل واحدة من الوحدات الأخرى يُختار قطاع اختياراً عشوائياً بالمعاينة باحتمال متناسب مع الحجم. ثم يُطاف بجميع القطاعات المشمولة بالعيّنة، بما

<sup>١٥</sup> استُخدم تصميم القطاع المعياري أيضاً في برنامج استقصاءات مشروع الدول العربية لتنمية الطفل في الثمانينات والتسعينات من القرن الماضي — انظر جامعة الدول العربية (١٩٩٠).

في ذلك المختارة تلقائياً للحصول على قوائم حديثة بالأسر المعيشية. ويُختار جزء معين (نسبة معينة) من الأسر المعيشية اختياراً منهجياً من كل عينة "مجموعة" مدرجة في العينة لكي تُجرى معها مقابلة في استقصاء ديمغرافي وصحي. ولأن جميع القطاعات ذات حجم واحد تقريباً، يسفر إجراء المعاينة عن عينة من القطاعات والأسر المعيشية اختيرت على أساس تساوي الاحتمالات وعلى مرحلتين.

١٤١ - تصميم القطاع المعياري للاستقصاء الديمغرافي والصحي قريب من الخطة ٦ الواردة أعلاه: اختيار مجموعات على أساس تساوي الاحتمالات، والاختيار على أساس نسبة ثابتة داخل المجموعات المدرجة في العينة (هنا أيضاً على أساس تساوي الاحتمالات). غير أنه بإجراء التقسيم إلى قطاعات معيارية، تُتجنب أوجه القصور الخطيرة المشار إليها أعلاه في الخطة ٦: يتم التحكم بالحجم الإجمالي للعينة بشكل يكاد يكون دقيقاً وكذلك التحكم بأعباء عمل المقابلين.

١٤٢ - ثمة ميزة هامة لتصميم القطاعات المعيارية وهي أن تخفيضاً كبيراً يحدث في عبء عمل إعداد القوائم في المرحلة قبل الأخيرة من مراحل الاختيار. فلكل وحدة منطقة تتألف من قطاعات عددها  $n$ ، يخفّض عبء العمل في إعداد القوائم إلى  $1/n$  (عندما يوجد قطاع واحد في الوحدة لا يحدث أي تخفيض). فإذا كانت وحدة منطقة معينة تحتوي على أربعة قطاعات، مثلاً، يكون عبء العمل في وضع القوائم ربع ما كان يمكن أن يكون عليه لو لزم إعداد القوائم في المنطقة بأسرها. وبذلك تُخفّض تكاليف إعداد العينات بسبب هذه الخاصية.

١٤٣ - بينما تكون تكاليف إعداد القوائم أقل، هناك ثمن إضافي يُدفع لهذا الانخفاض. فمن أوجه القصور التي تعوق تصميم القطاعات المعيارية أنه يجب إجراء عمليات وضع خرائط للقطاعات التي يكون مقاس حجمها أكثر من ١. ورسم الخرائط يمكن أن يكون متعباً ومكلفاً، ويتطلب تدريباً بعناية، وهو عرضة لوقوع أخطاء. وحقيقة أن أوجه القصور الطبيعية لا تكون في كثير من الأحيان معروفة جيداً تعوق ترسيم حدود القطاعات داخل وحدة المنطقة بشكل معقول. هذا القصور يجعل من الصعب على المقابلين الذين يزورون القطاع فيما بعد أن يحددوا موقع الأسرة المختارة بالضبط. غير أنه يمكن تخفيف حدة هذه المشكلة الأخيرة إلى حد ما، بإدراج اسم رب الأسرة في مرحلة وضع القوائم، وفي هذه الحالة يكون الضعف في رسم أوجه القصور مشكلة أقل إزعاجاً.

### ٣ - ٨ - ٣ التصميم المعدل للمجموعة: استقصاءات المجموعات المتعددة المؤشرات

١٤٤ - يشكو القائمون بإجراء الاستقصاء عموماً من النفقات اللازمة والوقت اللازم لوضع قوائم بالأسر المعيشية لدى أخذ عينات المجموعات التي تُختار في المرحلة قبل الأخيرة. وضع القوائم لازم بوجه عام في معظم الاستقصاءات — بما في ذلك طريقة تصميم القطاعات المعيارية في الاستقصاءات الديمغرافية والصحية، كما ذكر أعلاه — للحصول على قائمة حديثة بالأسر المعيشية تُختار منها الأسر المراد مقابلتها للاستقصاء. وهذا مطلب أساسي لا سيما عندما يكون قد مضى على وضع إطار المعاينة أكثر من سنة. ومما يُذكر أن عملية وضع القوائم تشمل تكاليف وعمليات استقصاء كبيرة يُغفل عنها في أحيان كثيرة لدى تخطيط الميزانية وتحديد مواعيد إجراءات الاستقصاء. ويجب القيام بزيارة إلى الميدان، منفصلة عن الزيارة اللازمة لإجراء المقابلات، لوضع القوائم. علاوة على ذلك، تتراوح نسبة الأسر المراد إدراجها في القوائم من ٥ على ١٠ أضعاف عدد الأسر التي ستُختار. افترض، مثلاً، أن خطة العينات تقتضي اختيار ٣٠٠ وحدة معاينة أولية وأن متوسط حجم المجموعة ٢٥ أسرة، تجد أن مجموع الأسر المعيشية المراد مقابلتها ٧ ٥٠٠ أسرة. وإذا كان متوسط ما تحتوي عليه وحدة المعاينة الأولية في المرحلة قبل النهائية ١٥٠ أسرة، يجب وضع قوائم بـ ٤٥ ٠٠٠ أسرة.

١٤٥ - استراتيجية المعاينة، التي استُخدمت في استقصاء المجموعات الذي وضعه البرنامج الموسع للتحصين (منظمة الصحة العالمية، ١٩٩١) وضعتها مراكز مكافحة الأمراض ومنظمة الصحة العالمية لأسباب منها تجنّب صرف النفقات والوقت في وضع القوائم. وقد استُخدم استقصاء العيّنات الذي وضعه البرنامج الموسع للتحصين، وقصد به تقدير تغطية الأطفال بالتحصين، في عشرات الدول النامية لفترة تزيد عن عقدين من الزمن. وثمة مسألة إحصائية هامة (تيرنر، وماغاناني، وشعيب، ١٩٩٦) تُعنى بمنهجية المعاينة. تستخدم منهجية استقصاء المجموعات عيّنة بالحصص في المرحلة الثانية من مراحل الاختيار، حتى وإن كانت وحدات المرحلة الأولى (القرى أو الأحياء) تُختار عادةً وفقاً لتعاليم المعاينة الاحتمالية. وطريقة اختيار العيّنة بالحصص، التي تُستخدم كثيراً، وإن كانت منها أشكال متباينة، تستتبع الشروع في المقابلات المرتبطة بالاستقصاء في موقع مركزي في القرية المختارة، ثم يمضي في اتجاه يقرّر عشوائياً، ويستمر في مقابلة الأسر المعيشية حتى تُستوفي الحصة المقررة. وتستمر زيارة الأسر المعيشية بموجب شكل من أشكال استقصاء المجموعات الذي وضعه البرنامج الموسع للتحصين حتى يوجد سبعة أطفال في الفئة العمرية المستهدفة. ومع أنه لا يوجد تحيز متعمد في استخدام هذه الأنواع من التقنيات، سجل كثير من الإحصائيين انتقادات مختلفة على مدى فترة طويلة من الزمن، من بينهم كالتون (١٩٨٧)، وسكوت (١٩٩٣)، وبينيت (١٩٩٣). الانتقاد الرئيسي هو أن هذه المنهجية لا تنتج عيّنة احتمالية (انظر الفرع ٣ - ٢، المتعلق بالمعاينة الاحتمالية مقابل طرق معاينة الأخرى للاطلاع على مناقشة السبب في جعل المعاينة الاحتمالية النهج الموصى به في استقصاءات الأسر المعيشية).

١٤٦ - ثمة شكل من أشكال استقصاء المجموعات الذي وضعه البرنامج الموسع للتحصين، وهو المسمّى بالتصميم المعدل لاستقصاء المجموعات، وُضِعَ لتلبية الحاجة إلى استراتيجية معيّنة تتجنب عمليات وضع القوائم، لكنها مع ذلك متجذرة في المعاينة الاحتمالية. وقد أُجريت تطبيقات مختلفة للتصميم المعدل لاستقصاء المجموعات، وكذلك تصاميم أخرى، في مختلف أنحاء العالم في استقصاءات للمجموعات ذات مؤشرات متعدّدة برعاية اليونيسيف، لرصد غايات وأهداف معيّنة لمؤتمر القمة العالمي المعني بالأطفال، متصلة بوضع الأطفال والنساء (صندوق الأمم المتحدة الدولي لتعليم الأطفال، ٢٠٠٠).

١٤٧ - التصميم المعدل لاستقصاء المجموعات استراتيجية معاينة عند الحد الأدنى. وهي معاينة تستخدم تصميماً بسيطاً ذا مرحلتين، ويُستخدم التقسيم بعناية إلى طبقات بالإضافة إلى طواف سريع في الوحدات وتقسيم المناطق إلى قطاعات. ولا توجد فيه عملية وضع قوائم. وفيما يلي الجوانب الأساسية لتصميم المعاينة بموجب التصميم المعدل لاستقصاء المجموعات:

- اختيار عيّنة لوحدات منطقة، كالقرى أو أحياء المدن، في المرحلة الأولى، باستخدام احتمال متناسب مع الحجم أو تساوي الاحتمالات، وذلك يتوقف على مدى تباين وحدات المعاينة الأولية فيما يتعلق بمقاسات أحجامها. يمكن استخدام مقاسات الحجم القديمة حتى وإن كان قد مضى على وضع إطار التعداد بضع سنوات، غير أنه يجب أن يشمل الإطار المجموعة السكانية التي هي موضع الاهتمام، سواء أكانت وطنية أم محلية.
- القيام بزيارات لكل وحدة منطقة معاينة لإجراء طواف سريع لطلب المعلومات وتقسيم المنطقة إلى قطاعات، باستخدام خرائط موجودة أو مسودات خرائط، مع تقرير عدد القطاعات مسبقاً وكون هذا العدد مساوياً لمقاس الحجم المحدد في التعداد مقسوماً على حجم المجموعة المرغوب (المتوقّع). وتكون القطاعات المنشأة متساوية تقريباً في حجم السكان.
- اختيار قطاع واحد من المنطقة مع تساوي الاحتمالات من كل وحدة معيّنة أولية.

- إجراء مقابلات مع جميع الأسر المعيشية في كل قطاع مختار.

١٤٨ - استخدام التقسيم إلى قطاعات دون وضع قوائم هو الميزة الرئيسية لتصميم مقياس الحجم. وهذا يختلف عن تصميم القطاع المعياري للاستقصاء الديمغرافي والصحي، الذي يقتضي وضع قائمة لكل قطاع. وتعوض عملية التقسيم إلى قطاعات أيضاً تعويضاً جزئياً عن استخدام إطار ربما يكون قد أبلاه الزمن. ومع أن في هذا التصميم ميزة إنتاج تقدير دون تحيز، يوجد فيه أيضاً قصور من حيث إنه أقل تحكماً بحجم العينة في نهاية المطاف، لأن القطاع المختار ربما يكون، بسبب النمو، قد أصبح أكبر حجماً بكثير مما يشير إليه الإطار.

١٤٩ - غير أن رسم الخرائط مطلوب لتصميم مقاسات الحجم، كما هو مطلوب لتصميم القطاعات المعياري للاستقصاء الديمغرافي والصحي، مع جميع ما ينطوي عليه ذلك من أوجه قصور، كما دُكر في طريقة الاستقصاء الديمغرافي والصحي. يضاف إلى ذلك أن إنشاء قطاعات صغيرة، يبلغ حجمها مقدار حجم المجموعة، وتكون حدودها مرسومة بدقة، يمكن أن يكون صعباً حين لا تكون أوجه القصور الطبيعية موجودة لمناطق صغيرة. ويوجد وجه قصور نهائي: فبقدر ما يكون القطاع الذي تجرى مقابلته مجموعة صغيرة — جميع الأسر متجاورة جغرافياً — يكون لهذا الوضع أثر أكبر على التصميم، بسبب الترابط الوثيق نسبياً داخل الطبقة الواحدة، من الأثر الذي يُنتج في حالة المجموعات غير الصغيرة للتصميم المعياري للقطاعات.

### ٣ - ٩ مواضع خاصة: العينات ذات المرحلتين والمعاينة لتحديد الاتجاهات

١٥٠ - يغطي هذا الفرع موضوعين خاصين حول تصميم العينات في استقصاءات الأسر المعيشية، وهما: (أ) المعاينة على مرحلتين، تُستخدم المرحلة الأولى منهما لإجراء مقابلة قصيرة لاستعراض الأفراد المقيمين في الأسر المعيشية بحثاً عن أشخاص ينتمون إلى المجموعة السكانية المستهدفة، وتستتبع المرحلة الثانية اختيار عينة من الأشخاص الذين يستوفون المعايير المطلوبة؛ و(ب) منهجية معاينة يكرّر فيها الاستقصاء بغية تقدير التغيير أو الاتجاه.

### ٣ - ٩ - ١ المعاينة على مرحلتين

١٥١ - يلزم نوع خاص من تصاميم العينات في استقصاءات الأسر المعيشية في الحالة التي تكون فيها المعلومات المتاحة غير كافية لإجراء اختيار بكفاءة لعينة من المجموعة السكانية المستهدفة والتي هي موضع اهتمام. تبرز هذه الحاجة بوجه عام عندما تكون المجموعة السكانية التي يستهدفها الاستقصاء مجموعة سكانية فرعية — وغالباً ما تكون نادرة — يوجد أفرادها في نسبة صغيرة من الأسر المعيشية. ومن الأمثلة عليهم أفراد مجموعة إثنية خاصة، والأيتام، والأشخاص الذين يزيد دخلهم أو يقل عن مستوى محدد. يمكن في أحيان كثيرة تقسيم أشخاص بعناية إلى طبقات لتعيين وحدات مناطق تتركز فيها، مثلاً، مجموعة إثنية معينة وهي موضع اهتمام، أو أشخاص ذوو دخل عالٍ؛ لكن إذا تفرق أفراد هذه المجموعة الإثنية بطريقة عشوائية إلى حد ما واختلطوا بالسكان في كل مكان، أو عندما تكون المجموعة المستهدفة — كالأيتام، مثلاً — مجموعة نادرة، عندئذ يكون التقسيم الطبقي استراتيجياً غير كافية، ويجب استخدام تقنيات أخرى لأخذ عينات منها.

١٥٢ - هناك تقنية واحدة تُستخدم في كثير من الأحيان، وهي المعاينة على مرحلتين، ويُشار إليها أيضاً بعبارة المعاينة بعد الطبقة، أو المعاينة المزدوجة. وهي تنطوي على أربع خطوات:

- (أ) اختيار عينة "كبيرة" من الأسر المعيشية؛
- (ب) إجراء مقابلة استعراضية قصيرة لتعيين الأسر المعيشية التي يقيم ضمن أفرادها السكان المستهدفون؛
- (ج) تقسيم العينة الكبيرة بعد التقسيم إلى طبقات إلى فئتين بناءً على المقابلة الاستعراضية؛
- (د) اختيار عينة فرعية من الأسر المعيشية من كل واحدة من الطبقتين لإجراء مقابلة ثانية أطول مع المجموعة المتسهدفة.

١٥٣ - الهدف من النهج ذي المرحلتين هو توفير التكاليف بإجراء مقابلة استعراضية صغيرة في العينة الكبيرة الأولى. وتليها مقابلة أخرى أكثر استفاضة في موعد لاحق، لكن مع الأسر المعيشية المؤهلة لذلك. لهذا السبب، غالباً ما تكون العينة الأولى في كثير من الأحيان عينةً اختيرت لغرض آخر وتلحق بها المقابلة الاستعراضية "كبند فرعي" ملحق بالاستقصاء الأصلي. هكذا يسمح الإجراء بتخصيص معظم الموارد للمعاينة في المرحلة الثانية وإجراء المقابلات بميزانية متوازنة فقط هي لازمة لمرحلة الاستعراض.

### مثال

افترض أن من المخطط إجراء استقصاء لـ ٨٠٠ طفل يتيم يقيمون في أسر معيشية للباقي على قيد الحياة من والديهم أو مع أقارب آخرين (خلفاً للأيتام الذين يعيشون في مؤسسة للأيتام). وافترض كذلك أنه يلزم أخذ عيّنات من ١٦ ٠٠٠ أسرة معيشية لمعرفة أماكن إقامة ٨٠٠ طفل يتيم — بمعدل طفل واحد تقريباً لكل ٢٠ أسرة. لأن نفقات تصميم وإدارة عينة من ١٦ ٠٠٠ أسرة معيشية تعتبر غير عملية لإجراء ٨٠٠ مقابلة مفصلة فقط، تقرر أن يُستفاد من استقصاء ذي أغراض عامة بشأن الصحة مخطط لإجرائه هو أيضاً. الاستقصاء الصحي مصمّم لعينة من ٢٠ ٠٠٠ أسرة معيشية. واتفق مديرا الاستقصاءين أن يُلحق بند فرعي بالاستقصاء الصحي يتألف من سؤال واحد هو: هل يوجد أي شخص يبلغ عمره ١٧ سنة أو أقل يعيش في هذه الأسرة وقد توفيت أمه أو أبوه أو كلاهما؟ يتوقع أن يعرف هذا السؤال الملحق أسراً معيشية تضم نحو ١ ٠٠٠ يتيم. يخطط مدير استقصاء الأيتام لأن يأخذ عندئذٍ عينة فرعية بواقع ٨٠ في المائة من تلك الأسر المعيشية لإجراء مقابلات تفصيلية.

١٥٤ - من شأن المثال الوارد أعلاه أن يوضّح أيضاً متى تكون المعاينة على مرحلتين استراتيجيةً ملائمة. لاحظ أن حجم العينة المستهدفة في التوضيح هو ٨٠٠ يتيم فقط، لكن حجم العينة من حيث عدد الأسر المعيشية اللازمة للعثور على هذا العدد من الأطفال يبلغ ١٦ ٠٠٠ أسرة. لذلك، من المرجح أن يستنتج الخبير الفني للمعاينة ومدير الاستقصاء أن المعاينة على مرحلتين هي التصميم الأكثر ما يكون عملياً والأكثر كفاءة لاستخدامه.

١٥٥ - تقسيم عينة المرحلة الأولى إلى فئتين بعد التقسيم إلى طبقات عملية هامة لسببين اثنين. سيكون السؤال أو الأسئلة الموجهة في المقابلة الاستعراضية موجزة في جميع الأحيان تقريباً، لأنها ملحقّة باستقصاء آخر من المؤكد أنه يتضمن من قبل مقابلة طويلة. ومن المستبعد أن يوافق مدير الاستقصاء الأصلي على مجموعة مفصلة جداً من الأسئلة الاستعراضية. ومن ثم يرجح أن يتبين أن بعض الأسر التي ذُكر في المثال الوارد أعلاه أن لديها أيتاماً لا يوجد لديها في الواقع أيتام. والعكس صحيح أيضاً. هذه الأخطاء التي تحدث في التقسيم الطبقي توحى بأنه يجب إنشاء طبقتين، واحدة للأسر التي كانت نتيجة الاستعراض بالنسبة إليها موجبة والأخرى للأسر التي كانت نتيجة الاستعراض بالنسبة إليها سالبة. وستؤخذ عيّنات من كل طبقة لإجراء المقابلة الكاملة معها بسبب أنه

ربما يكون قد حدث خطأ في التقسيم إلى حد ما. ستكون نسبة العينة في الطبقة التي رد أفرادها بـ "نعم" عالية جداً — قد تصل إلى ١٠٠ في المائة — ستكون أصغر كثيراً في الطبقة التي رد أفرادها بـ "لا".

### ٣ - ٩ - ٢ المعاينة لتقدير التغير أو الاتجاه

١٥٦ - تُصمَّم استقصاءات الأسر المعيشية في بلدان كثيرة لغرض مزدوج هو تقدير (أ) مؤشرات خط القاعدة (مستوياتها) في مناسبة إجراء الاستقصاء لأول مرة، و(ب) التغير في هذه المؤشرات في المرحلة الثانية واللاحقة من مراحل إجراء الاستقصاء. وعندما يتكرر الاستقصاء أكثر من مرة تقاس أيضاً الاتجاهات في المؤشرات. ومع تكرار الاستقصاء توجد آثار مختلفة على تصميم العينة لا تُدخَل عند إجراء استقصاء مرة واحدة لقطاعات متعددة. والمسائل التي هي مثار القلق بوجه خاص هي موثوقية تقدير التغير والمزيج الصحيح فيما يتعلق باستخدام نفس الأسر المعيشية أو أسر غيرها من مناسبة إلى التي تليها. ويتصل بالنقطة الأخيرة الخشية من وقوع تحيزات وعبء على المستجيبين إذا قوبلت الأسر نفسها مراراً وتكراراً.

١٥٧ - فحص مسألة الموثوقية يتطلب أيضاً مشاهدة عرض رياضي. نبدأ بالنظر إلى تباين التغير المقدَّر،  $d = p1 - p2$ ، يُعبَّر عنه كما يلي:

(٢٢ - ٣)

$$\sigma_d^2 = \sigma_{p1}^2 + \sigma_{p2}^2 - 2\sigma_{p1,p2} = \sigma_{p1}^2 + \sigma_{p2}^2 - 2\rho\sigma_{p1}\sigma_{p2}$$

حيث إن قيمة  $p$  هي النسبة الجاري تقديرها؛ و  $\sigma_d^2$  هو تباين الفرق؛ و  $\sigma_p^2$  هو تباين  $p$  في المناسبة الأولى أو الثانية، يشار إليه بـ ١ أو ٢؛  $\sigma_{p1,p2}$  هو التفاوت المشترك بين  $p1$  و  $p2$ ؛ و  $\rho$  هي العلاقة المشتركة بين القيم الملاحظة لـ  $p1$  و  $p2$  في مناسبتَي الاستقصاء.

وحيثما كان التغير المقدَّر صغيراً نسبياً، وغالباً ما يكون الأمر كذلك، يوجد لدينا

$$\sigma_{p1}^2 \approx \sigma_{p2}^2$$

ثمَّ،  $\sigma_d^2 = 2\sigma_p^2 - 2\rho\sigma_p^2$  (يمكننا أن ندع النصوص الفرعية، ١ و ٢). ومن ثمَّ:

(٢٣ - ٣)

$$\sigma_d^2 = 2\sigma_p^2(1 - \rho)$$

١٥٨ - لتقييم المعادلة (٢٢ - ٣)، نلاحظ أن تقديراً لـ  $\sigma_p^2$  لاستقصاء المجموعات هو تقدير عينة عشوائية بسيطة، مضرورياً بأثر تصميم العينة. أما الترابُط،  $\rho$ ، الذي هو أعلى ما يكون عندما تُستخدم العينة نفسها للأسر المعيشية، فربما يكون ٠,٨ أو حتى أعلى من ذلك. وفي هذه الحالة، يكون التقدير  $s_d^2$  من  $\sigma_d^2$  ويعطى كما يلي:

(٢٤ - ٣)

$$s_d^2 = 2[(pq)f/n](0.2) \quad \text{أو} \quad 0.4(pq)f/n$$

١٥٩ - إذا استُخدمت المجموعات نفسها لكن اختلفت الأسر المعيشية، فإن  $\rho$  تكون ما زالت إيجابية لكنها أصغر كثيراً — ربما تتراوح بين ٠,٢٥ و ٠,٣٥. وعندئذٍ كان يمكن أن يكون لدينا ( $\rho$  من ٠,٣):

(٢٥ - ٣)

$$s_d^2 = 2[(pq)f/n](0.7) \quad \text{أو} \quad 1.4(pq)f/n$$

١٦٠ - وأخيراً، مع عيّنة مستقلة تماماً في المناسبة الثانية، باستخدام مجموعات مختلفة وأسر معيشية مختلفة،  $p$  تساوي صفرًا ويكون لدينا:

(٣ - ٢٦)

$$s_d^2 = 2[(pq)f/n]$$

باستخدام قيمة نموذجية لأثر تصميم العيّنة مقدارها ٢,٠، تسفر المعادلة ٣ - ١٩ عن:

(٣ - ٢٧)

$$s_d^2 = 4[(pq)/n]$$

١٦١ - في حالة الاستقصاءات المكررة، باستخدام ترادف جزئي، مثلاً ٥٠ في المائة من نفس العيّنات/الأسر المعيشية و ٥٠ في المائة مجموعات/أسر معيشية جديدة، يجب أن تُضرب  $p$  بعامل  $F$  وهو يساوي نسبة العيّنة التي تراكب عيّنة أخرى. في هذه الحالة تصبح المعادلة ٣ - ١٦ ما يلي:

(٣ - ٢٨)

$$\sigma_d^2 = 2\sigma_p^2(1-F\rho)$$

١٦٢ - توجد نقاط طريفة يمكن إبدائها فيما ورد أعلاه. أولاً، التفاوت المقدر لتغيرٍ مقدر صغير نسبياً بين استقصاءين باستخدام العيّنة نفسها من الأسر المعيشية يساوي نحو ٤٠ في المائة فقط من تفاوت المستوى، سواء في المناسبة الأولى أم الثانية. وباستخدام المجموعات نفسها لكن مع أسر معيشية مختلفة ينتج تفاوتاً مقدراً للتغير يزيد بنسبة ٤٠ في المائة عن تفاوت المستوى. والعيّنات المستقلة تنتج تفاوتاً مقدراً يساوي ضعف تفاوت المستوى.

١٦٣ - لذلك توجد ميزات قوية لاستخدام نفس الأسر المعيشية في الاستقصاءات المكررة من حيث الموثوقية. وإن لم يحصل ذلك لا تزال هنالك تحسينات هامة جداً يمكن تحقيقها باستخدام إما (أ) جزء من الأسر المعيشية نفسها وإما (ب) نفس المجموعات مع أسر معيشية مختلفة. فكلتا الاستراتيجيتين تنتجان تقديرات مع تفاوت صغير، مقارنةً بأقل الخيارات جاذبية، وهو استخدام عيّنات مستقلة تماماً.

١٦٤ - فيما يتعلق بالخطأ من غير أخطاء المعاينة، تَكَرَّر حدوث أثرين سلبيين من المستجيبين — عدم الاستجابة أو إجابة مشروطة — كلما ازداد الاستخدام المكرر لنفس العيّنة من الأسر المعيشية. لا يقف الأمر عند كون المستجيبين يزدادون تردداً في التعاون، وبذلك يزداد الامتناع عن الإجابة في دورات الاستقصاء اللاحقة، لكنهم يتأثرون أيضاً بالشروط وربما تتدهور نوعية إجاباتهم ودقتها كلما تكررت المقابلات.

١٦٥ - ترتبط مع الشروط الظاهرة المعروفة باسم تحيز "الوقت في العيّنة"، حيث تقديرات الاستقصاء من المستجيبين الذين يقدمون تقاريرهم عن نفس الفترة الزمنية لكن بمستويات مختلفة من التعرّض للاستقصاء لها قيم متوقعة مختلفة. دُرست هذه الظاهرة دراسة مستفيضة وتبين أنها موجودة في الاستقصاءات المتعلقة بمواضيع كثيرة — القوى العاملة، والنفقات، والإيرادات، وضحايا الجرائم. وفي الولايات المتحدة، مثلاً، حيث يقابل المستجيبون في استقصاءات القوى العاملة ثمان مرات، كان تقدير البطالة من المستجيبين لأول مرة في الاستقصاء باستمرار يزيد بنحو ٧ في المائة عن متوسط ردود المستجيبين على مدى المقابلات الثماني كلها. وقد استمر هذا النمط على مدى عدد من السنين في الولايات المتحدة. ولجابهة هذا التحيز اقترح الخبراء، فيما اقترحوه، ما يلي:

- أن المقابلات لا يمكن أن تقدم نفس الحوافز للمستجيبين في المقابلات اللاحقة كما قدمت لهم في المرة الأولى.
- ربما يتعلم المستجيبون أن بعض الإجابات تثير أسئلة إضافية، لذلك يتجنبون إعطاء إجابات معيَّنة.
- ربما تشمل المقابلة الأولى أحداثاً خارج نطاق الفترة المرجعية، أما في المقابلات اللاحقة فالحدث "محدد".
- ربما يغير المستجيبون سلوكهم بسبب الاستقصاء.
- ربما لا يكون المستجيبون مجتهدين في تقديم إجابات دقيقة في المقابلات اللاحقة عندما يصبحون ضجرين من عملية الاستقصاء (كاسبرزيك، ١٩٨٩).

١٦٦ - ينبغي ملاحظة أن معظم الأسباب المذكورة أعلاه تنطبق على المقابلات المعادة لنفس الاستقصاء؛ لكن عندما تُستخدم نفس الأسر المعيشية في استقصاءات مختلفة يحدث نفس السلوك من المستجيبين.

١٦٧ - يمكن أن يُرى مما ذكر أعلاه أن ثمة آثاراً متنافسة مرتبطة مع استخدام:

(أ) نفس العينة من الأسر المعيشية في كل مناسبة؛

(ب) إشراك أسر معيشية بديلة في جزء من العينة؛

(ج) إدخال عينة جديدة من الأسر المعيشية في كل مرة يُجرى فيها الاستقصاء.

١٦٨ - بالمضي من (أ) إلى (ج)، يزداد خطأ المعاينة في تقديرات التغير بينما يميل الخطأ من غير أخطاء المعاينة إلى الانخفاض. أدنى ما يكون عليه خطأ المعاينة يحدث عندما تُستخدم نفس العينات من الأسر المعيشية في كل مناسبة لأن الترابط بين الملاحظات يكون في أعلى نقطة تصلها. وخلافاً لذلك، يزيد استخدام نفس الأسر المعيشية التحيز الآتي من غير المعاينة. ويحدث عكس ذلك عندما تستخدم عينة جديدة من الأسر المعيشية في كل مرة.

١٦٩ - الخيار (ب) هو الذي يُرى بوجه عام أنه يقدم حلاً وسطاً من حيث الموازنة بين خطأ المعاينة والتحيز الآتي من غير المعاينة. فإذا احتُفظ بجزء من العينة سنة بعد سنة يتحسن خطأ المعاينة في البند (ج) والخطأ الآتي من غير المعاينة يتحسن في البند (أ). عندما يُجرى استقصاء في مناسبتين اثنتين فقط، يُرجح أن يكون البند (أ) هو الخيار الأفضل. ويُستبعد أن تترك آثار المستجيبين أثراً ضاراً كثيراً على مجموع أخطاء الاستقصاء الكلي عندما تُستخدم عينة مرتين فقط. غير أن أفضل ما يخدم تكرار الاستقصاء ثلاث مرات أو أكثر هو البند (ب). والاستراتيجية الملائمة هي استبدال ٥٠ في المائة من العينة في كل مناسبة بالتناوب (انظر الفصل الرابع للاطلاع على أمثلة لتناوب العينات في العينات الرئيسية).

### ٣ - ١٠ عندما يحدث خطأ في التنفيذ

١٧٠ - يعرض هذا الجزء موجزاً للتدابير التي ينبغي اتخاذها عندما يواجه تنفيذ خطة المعاينة عقبات، وقد نوقش معظمها من قبل أو ألمح لها أعلاه. غير أن واحداً من المبادئ الهامة ركَّز عليه في هذا الفصل والفصل الذي يليه، وهو أنه بالتخطيط بعناية كبيرة جداً في الوقت الذي يتم تصوُّر تصميم العينة فيه يمكن استباق العقبات

التي تعترض سبيل التنفيذ. ومع ذلك يمكن أن تظهر مشاكل لم يكن يتنبأ بها أحد على الرغم من أفضل أنواع التخطيط.

### ٣ - ١٠ - ١ تعريف السكان المستهدفين وتغطيتهم

١٧١ - غالباً ما تحدث المشاكل لأيّ سبب من أسباب متنوعة عديدة، عندما يكون السكان المشمولون فعلاً بالاستقصاء غير السكان الذي قُصد استهدافهم.

#### مثال

انظر في استقصاء قُصدَ به شمول السكان المستهدفين النموذجيين من بين جميع سكان البلد. وغالباً ما يكون السكان المشمولون فعلاً (أي الذين تختار منهم العيّنة) أقل من المجموع لأيّ من الأسباب التالية:

- الأشخاص المقيمون في المؤسسات، كالمستشفيات، والسجون، والثكنات العسكرية، لا تؤخذ منهم عيّنات.
- الأشخاص المقيمون في مناطق جغرافية معيّنة ربما يُستبعدون من التغطية عن قصد. فقد تشمل تلك المناطق تضاريس يتعذر الوصول إليها، أو مناطق متأثرة بكوارث طبيعية، ومناطق أُعلن أن دخولها محظور بسبب اضطرابات مدنية أو حرب، أو مجمّعات أو معسكرات يُقيم فيها لاجئون أو عمال أجانب آخرون، وما أشبه ذلك.
- الأشخاص الذين ليس لديهم ترتيبات إقامة دائمة يُحكم بأنهم "خارج نطاق" الاستقصاء. وقد يشمل هؤلاء البدو الرُّحّل، أو سكان القوارب، أو العمال المؤقتين، إلخ.

١٧٢ - المشكلة فيما يتعلق بهذه المجموعات السكانية الفرعية هي أنها، فيما يتعلق بخطة المعاينة، لا تُعرّف عادةً قبل الاستقصاء بأنها مجموعات يجب استبعادها. وهكذا يعاني التنفيذ عندما تُختار في عملية اختيار العيّنات (أ) مجموعة يتبين فيما بعد أنها معسكر عمل أو سجن أو منام للعمال، بدلاً من كونها منطقة سكنية "تقليدية"، أو (ب) وحدة معاينة أولية يتبين أنها ذات تضاريس جبلية ويُعتقد أنه يتعذر الوصول إليها. غالباً ما يكون "الحل" المتبّع في هذه الحالة هو أن تُستبدل بها وحدة معاينة أولية أخرى. غير أن هذا الحل ينطوي على إجراء متحيز.

١٧٣ - الحل المقبول هو تَجَنُّب المشكلة في مرحلة تصميم العيّنة. وهذا يتحقق بتعريف المجموعة السكانية المستهدفة أولاً بعناية، وتحديد المجموعات السكانية الفرعية التي ينبغي استبعادها من التغطية، لا التي تشملها. ثانياً، ينبغي عندئذٍ تعديل إطار العيّنات لحذف أيّ منطقة جغرافية لا يُراد شمولها بالاستقصاء. وينطبق هذا كذلك على أي منطقة تعداد ذات غرض خاص — كمعسكرات العمل، مثلاً — ينبغي استبعادها. ثالثاً، ينبغي انتخاب العيّنة من الإطار المعدل.

١٧٤ - ينبغي أيضاً ألاّ يغرب عن البال أن الحل المقترح أعلاه يعمل على تعريف المجموعة السكانية المستهدفة بدقة أكثر. ومن الأهمية بمكان أنه يجب وصف المجموعة السكانية المستهدفة بالضبط في تقارير لكي يكون المستخدم على علم صحيح. يناقش الفصل الرابع أطر المعاينة بمزيد من التفصيل.

### ٣ - ١٠ - ٢ حجم العينة أكبر مما تتحمّله ميزانية الاستقصاء

١٧٥ - تحدث مشكلة أخرى عندما يكون حجم العينة المحسوب أكبر مما تستطيع ميزانية الاستقصاء تحمّله. وعندما يحدث ذلك يجب على فريق الاستقصاء إما أن يلتزم أموالاً إضافية للاستقصاء وإما أن يُعدّل أهدافه المحدّدة بالقياس، بتخفيض متطلبات الدقة أو بتخفيض عدد المجالات.

١٧٦ - من طرق تخفيض الدقة (زيادة خطأ المعاينة) لتخفيض التكلفة تخفيضاً كبيراً اختيار عدد أقل من وحدات المعاينة الأولية مع المحافظة على الحجم الإجمالي للعينة. فبدلاً من اختيار ٦٠٠ وحدة معاينة أولية، مثلاً، في كل وحدة منها ١٥ أسرة معيشية ( $n=9\ 000$ )، يمكن تعديل خطة المعاينة لاختيار ٤٠٠ وحدة معاينة أولية تضم الوحدة منها ٢٢ أو ٢٣ أسرة معيشية ( $n\approx 9,000$ ). وفيما يتعلق بالمجالات، يمكن أن يكون أحد الحلول الاكتفاء بأربع مناطق رئيسية في البلد بدلاً من ١٠ أقاليم، مثلاً.

### ٣ - ١٠ - ٣ حجم المجموعة أكبر أو أصغر مما كان متوقّعا

١٧٧ - ثمة مشكلة تحدث في كثير من الأحيان، وهي أن المجموعة المختارة كعينة ربما تكون أكبر كثيراً من مقياس حجمها نتيجة لبناء مساكن جديدة، مثلاً، لا سيما عندما يكون إطار العيّنات قديماً. وربما كان فريق الاستقصاء يتوقّع ١٢٥ أسرة معيشية في مجموعة معينة لكنه ربما يجد في مرحلة وضع القوائم ٤٠٠ أسرة بدلاً من ذلك. من الحلول المعقولة في هذه الحالة إعادة تقسيم المجموعة إلى قطاعات جغرافية فرعية بأحجام متساوية تقريباً في عدد السكان. ويجب أن يكون عدد القطاعات مساوياً للعدد الحالي للسكان مقسوماً على مقياس الحجم الأصلي، مقرباً إلى أقرب عدد صحيح. وسيكون هذا في مثالنا  $400/125$ ، أو ٣,٢، ويقرب إلى ٣ قطاعات. وستنشأ القطاعات بواسطة رسم الخرائط والعدّ السريع للمساكن (خلافاً لعدّ الأسر المعيشية). ثم يُختار قطاع واحد منها اختياراً عشوائياً ويدرج في القائمة.

١٧٨ - ربما تحدث المشكلة المعاكسة أيضاً. وربما تكون المجموعة أصغر كثيراً مما هو متوقع، نتيجة لهدم بيوت أو وقوع كارثة طبيعية أو أسباب أخرى. غالباً ما يميل الناس في هذه الحالات إلى استخدام مجموعة أخرى بدلاً من هذه المجموعة، لكن هذا الفعل يُدخل التحيز. ينبغي، بدلاً من ذلك، أخذ المجموعة كما هي. ومع أن هذا ربما يسفر في نهاية المطاف عن حجم عينة أصغر من الحجم المستهدف، ستكون الزيادة في خطأ المعاينة قليلة ما لم يتعلق الأمر بعدد كبير من المجموعات التي هي كذلك. فأخذ العينة الأصغر دون تعديل (أو تبديل) سيمكّن — مع ذلك — من إنتاج تقدير غير متحيز لأن المجموعة "تمثل" التغيّر الراهن للمجموعة السكانية الذي حصل منذ إنشاء الإطار.

### ٣ - ١٠ - ٤ تناوّل حالات عدم الاستجابة

١٧٩ - عدم الاستجابة، وإن كان يتصل بتنفيذ الاستقصاء أكثر مما يتصل بتنفيذ العيّنات، مسألة خطيرة يمكن أن تدمر تقديرات استقصاء الأسر المعيشية (انظر الفصلين السادس والثامن للاطلاع على مناقشة مفصلة لعدم الاستجابة). إذا سُمح بحدوث عدم الاستجابة في أكثر من ١٠ - ١٥ في المائة من العيّنات، فإن التحيز الناتج في التقديرات يمكن أن يجعلها موضع شك كبير. هنا أيضاً يوجد ميل في كثير من البلدان إلى "حل" مشكلة عدم الاستجابة باستخدام أسر تستجيب بدلاً من الأسر غير المستجيبة. هذه الطريقة نفسها متحيزة لأن الأسر الجديدة التي تُستخدم لتمثل الأسر المستجيبة فقط، لا الأسر غير المستجيبة. وإن خصائص المجموعتين الأخيرتين معروفة

بأنها مختلفة من حيث متغيرات هامة في الاستقصاء، لا سيما تلك المتصلة بالحالة الاقتصادية - الاجتماعية. أما الحل المفضل، وإن كان مع الأسف لا يمكن أن ينجح بنسبة ١٠٠ في المائة، فهو الحصول على ردود من الأسر التي لم تستجب بادئ الأمر. ويجب أن يتم ذلك بالتخطيط منذ البداية للعودة إلى الأسر غير المستجيبة في سلسلة من الاتصالات اللاحقة سعياً إلى الحصول على تعاونها (بالنسبة إلى الأسر الراضة) أو لعلها تكون في البيت (بالنسبة إلى الأسر الغائبة أو التي لم تكن موجودة لأسباب أخرى). ربما يلزم أن يعاود الاتصال خمس مرات، لكن يجب ألا يقل عن ٣ اتصالات.

### ٣ - ١١ موجز المبادئ التوجيهية

١٨٠ - يوجز هذا الفرع المبادئ التوجيهية الرئيسية التي يُراد استخلاصها من هذا الفصل. ومع أن بعض المبادئ التوجيهية يمكن أن ينطبق في جميع الظروف ("استعمل المعاينة الاحتمالية"، مثلاً)، هناك بعض المبادئ التي يكون الاستثناء منها ملائماً، وذلك يتوقف على الظروف الخاصة للبلد وموارده ومتطلباته. لهذا السبب، تُعرض المبادئ التوجيهية، في شكل قائمة، بروح "القاعدة التقريبية" أكثر منها كتوصيات ثابتة لا رجوع فيها. يجب أن يتعهد القائمون بالاستقصاء بالقيام بما يلي:

- استخدام تقنيات المعاينة الاحتمالية في كل مرحلة من مراحل الاختيار.
- الاجتهاد إلى أقصى حد ممكن لممارسة البساطة في تصميم العيّنات، لا إلى التعقيد.
- التماس تقنيات للاختيار تُنتج عيّنات تتسم بالترجيح الذاتي، أو ما يقرب من الترجيح الذاتي، في المجالات — أو بالإجمال، إذا لم يشتمل التصميم على مجالات.
- استخدام تصميم العيّنة على مرحلتين إن أمكن ذلك.
- حساب حجم العيّنة باستخدام معادلة مثل (٣ - ٥)، مع تعديل قيمة المَعْلَمَات الثابتة (كنسبة عدم الاستجابة المتوقعة، ومتوسط حجم الأسرة) حسب الاقتضاء لكي يعكس الوضع في البلد.
- استخدام قيمة لأثر التصميم مقدارها ٢,٠، باعتبارها القيمة الافتراضية لمعادلة حجم العيّنة ما لم تكن معلومات أفضل متاحة في البلد.
- إسناد حجم العيّنة إلى التقدير الرئيسي الذي يُعتقد بأنه يشمل أصغر نسبة مئوية من السكان من بين كل التقديرات الرئيسية التي سيشملها الاستقصاء.
- اختيار هامش خطأ، إذا سمحت الميزانية، أو مستوى دقة، لتقدير رئيسي (يزيد عن) يعادل ١٠ في المائة من التقدير، أو بعبارة أخرى، خطأ نسبي مقداره ١٠ في المائة، مع مستوى الثقة البالغ ٩٥ في المائة؛ وإلاً بخطأ نسبي مقداره ١٢ - ١٥ في المائة.
- تعريف وحدات الاختيار في المرحلة الأولى باعتبارها مناطق تعداد السكان، إذا كان ذلك مناسباً أو ملائماً.
- استغلال التقسيم الضمني إلى طبقات مقروناً بالمعاينة المنهجية باحتمال متناسب مع الحجم حيثما أمكن ذلك، لا سيما للتصاميم المتعددة الأغراض.

- الحد من عدد مجالات التقدير إلى أقل ما هو لازم على الإطلاق (لجعل حجم العينة عند مستوى يمكن إدارته).
- السعي إلى الحصول على أكبر عدد ممكن — بضع مئات — من المجموعات (أو وحدات المعاينة الأولية إذا كان الاختيار على مرحلتين): كلما كان أكثر فهو أفضل.
- استخدام المجموعات الصغيرة الحجم: ١٠ - ١٥ أسرة معيشية: كلما كانت أصغر فهي أفضل.
- استخدام حجم ثابت للمجموعة بدلاً من الأحجام المتباينة، بعبارة أخرى عدد ثابت من الأسر بدلاً من النسبة الثابتة.
- في حالة المجالات، التوجُّه نحو حد أدنى مقداره ٥٠ وحدة معاينة أولية لكل مجال.
- التخطيط لإجراء ثلاث مكالمات لاحقة، كحد أدنى، ويُفضَّل إجراء خمس مكالمات، لتحويل الأسر غير المستجيبة إلى أسر مستجيبة.
- في حالة المجموعات السكانية النادرة، يُنظر في نهج المعاينة على مرحلتين، المتمثل في تعليق سؤال "ملحق" على استقصاء كبير قائم، حُطِّط له من قبل، للبحث عن أشخاص مستهدفين؛ ويتابع بمقابلة مكثفة حول عينة فرعية.
- في حالة الاستقصاءات المصمَّمة لقياس التغيُّر، تُجرى مقابلة مع الأسرة المعيشية نفسها في المناسبتين إن كان المقصود إجراء مقابلتين اثنتين فقط؛ وإذا كان المقصود إجراء ثلاث مقابلات أو أكثر، يُستخدم مشروع فيه تراكُّب جزئي بالتناوب مع أسر جديدة تُدخل في العينة في كل مناسبة.

### المراجع ومواد أخرى للقراءة

- Bennett, S. (1993), The EPI cluster sampling method: a critical appraisal. Invited paper, International Statistical Institute Session, Florence, Italy.
- Cochran, W. (1977), *Sampling Techniques*, 3<sup>rd</sup> ed., New York: Wiley.
- Hansen, M., W. Hurwitz and W. Madow (1953). *Sample Survey Methods and Theory*, New York: Wiley.
- Hussmans, R., F. Mehran and V. Verma (1990). *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*. International Labour Office, Geneva: Chap.11, "Sample design".
- International Statistical Institute (1975). *Manual on Sample Design*. World Fertility Survey Basic Documentation. Voorburg, Netherlands.
- Kalton, G. (1983), *Introduction to Survey Sampling*. Beverly Hills, California: Sage. Publications.
- \_\_\_\_\_ (1987), An assessment of the WHO Simplified Cluster Sampling Method for estimating immunization coverage. Report to UNICEF, New York.
- \_\_\_\_\_ (1993). *Sampling Rare and Elusive Populations*. Department for Economic and Social Information and Policy Analysis, Statistical Division and National Household Survey Capability Programme, INT-92-P80-16E. New York: United Nations.
- Kasprzyk, D., and others, eds. (1989). *Panel Surveys*. New York: John Wiley & Sons, Chap. 1.

- Kish, L. (1965), *Survey Sampling*, New York: Wiley.
- Krewski, D., R. Platek, and J.N.K. Rao, eds. (1981). *Current Topics in Survey Sampling*. New York: Academic Press.
- Le, T. and V. Verma (1997). *An Analysis of Sample Designs and Sampling Errors of the Demographic and Health Surveys*, DHS Analytical Reports, No. 3. Calverton, Maryland: Macro International Inc.
- League of Arab States (1990), *Sampling Manual, Arab Maternal and Child Health Survey*, Basic Documentation 5. Cairo: Pan Arab Project for Child Development (PAPCHILD).
- Macro International Inc. (1996). *Sampling Manual*. DHS-III Basic Documentation, No. 6. Calverton, Maryland: Macro International Inc..
- Namboodiri, N., ed. (1978). *Survey Sampling and Measurement*. New York: Academic Press.
- Raj, D. (1972). *Design of Sample Surveys*, New York: McGraw-Hill.
- Scott, C. (1993), Discussant comments for session on “Inexpensive Survey Methods for Developing Countries”. Invited paper, International Statistical Institute Session, Florence, Italy.
- Som, R. (1966), *Practical Sampling Techniques*, 2<sup>nd</sup> ed., New York: Marcel Dekker, Inc.
- Turner, A., R. Magnani, and M. Shuaib, (1996), A not quite as quick but much cleaner alternative to the Expanded Programme on Immunization (EPI) Cluster Survey design. *International Journal of Epidemiology*, (Liverpool, United Kingdom) Vol. 25, No. 1.
- United Nations (1984). *Handbook of Household Surveys*, Revised Edition. Studies in Methods, No. 31. Sales No. E.83.XVII.13.
- \_\_\_\_\_ (1986). Sampling frames and sample designs for integrated household survey programmes, *National Household Survey Capability Programme*, New York: United Nations. Department of Technical Co-operation for Development and Statistical Office.
- \_\_\_\_\_ (2005). *Household Sample Surveys in Developing and Transition Countries*. Studies in Methods, No. 96. United Nations publication, Sales No. E.05.XVII.6.
- United Nations Children’s Fund (2000), *End-Decade Multiple Indicator Survey Manual*. Chap. 4; entitled “Designing and selecting the sample”, and appendix 7; entitled “*Sampling details*.” New York: UNICEF.
- United States Bureau of the Census (1978). *Current Population Survey Design and Methodology*, Technical Paper 40, Washington, D.C.: Bureau of the Census.
- Verma, V. (1991), *Sampling Methods*. Training Handbook, Tokyo: Statistical Institute for Asia and the Pacific.
- Waksberg, J. (1978), Sampling methods for random digit dialing, *Journal of the American Statistical Association*, vol. 73, pp. 40-46.
- World Bank (1999). *Core Welfare Indicators Questionnaire (CWIQ) Handbook*. Washington D.C.: World Bank, chap.4.
- Health Organization (1991). *Expanded Programme on Immunization, Training for Mid-level Managers: Coverage Survey*. WHO/EPI/MLM91.10. Geneva.

## الفصل الرابع أطر المعاينة والعينات الرئيسية

### ٤ - ١ أطر المعاينة في استقصاءات الأسر المعيشية

١ - غطى الفصل السابق، باستثناء أطر المعاينة، الجوانب المتعددة الأوجه لتصميم العينات وبعض الخيارات المتاحة لتصميم العينات في استقصاءات الأسر المعيشية. غير أن واحداً من أهم جوانب تصميم العينات في استقصاءات الأسر المعيشية هو إطاره، وقد حُصص فصل مستقل لهذا الموضوع.

٢ - لإطار المعاينة آثار هامة على تكلفة أيّ استقصاء، للأسر المعيشية أو غيرها، وعلى جودته. ففي استقصاءات الأسر المعيشية، تشكل أطر المعاينة المعبية مصدراً مشتركاً للأخطاء من غير أخطاء المعاينة، لا سيما التغطية الناقصة لمجموعات سكانية فرعية هامة. يحاول هذا الفصل أن يشرح بالتفصيل أفضل الممارسات في بناء الإطار واستخدامه، أخذاً في الحسبان مختلف مراحل المعاينة. يقسم الفصل إلى فرعين: الفرع الأول يغطي المسائل العامة المتعلقة بالأطر وتطويرها، مع التركيز على تصميم العينات المتعدد المراحل في استقصاءات الأسر المعيشية؛ ويناقش الفرع الثاني المسائل الخاصة التي تظهر لدى استعمال إطار عينة رئيسية.

### ٤ - ١ - ١ تعريف إطار المعاينة

٣ - التعريف العملي البسيط لإطار المعاينة هو: مجموعة مواد المصادر التي تُختار منها العينة. يشمل التعريف أيضاً الغرض من أطر المعاينة، وهو توفير وسيلة لاختيار الأعضاء الخاصين من المجموعة السكانية المستهدفة، الذين ستجرى معهم مقابلات في الاستقصاء. ويمكن أن يلزم أكثر من مجموعة واحدة من المواد. هكذا يكون الأمر بوجه عام في استقصاءات الأسر المعيشية بسبب طبيعتها المتعددة المراحل. في المراحل الأولى من الاختيار في استقصاءات الأسر المعيشية تؤخذ العينات في الأحوال النمطية من أطر مناطق جغرافية. وفي المرحلة الأخيرة، يمكن اختيار العينات إما من إطار منطقة وإما من إطار قائمة (انظر أدناه للاطلاع على مناقشة أطر مناطق وقائمت).

### ٤ - ١ - ١ - ١ إطار المعاينة والمجموعة السكانية المستهدفة

٤ - الاعتبار الهام في اتخاذ القرار بشأن الإطار المناسب (أو الأطر المناسبة) للاستعمال في استقصاءات الأسر المعيشية هو العلاقة بين المجموعة السكانية التي يستهدفها الاستقصاء ووحدة الاختيار. فوحدة الاختيار تحدد الإطار. وهي أيضاً تحدد احتمال الاختيار في المرحلة الأخيرة.

<sup>١</sup> يُحال القارئ إلى الفصل الثالث، الجدول ٣ - ١ للاطلاع على قائمة المصطلحات المتصلة بالمصطلحات المستخدمة في الفصلين الثالث والرابع.

## مثال

نقول توضيحاً إن فريق الاستقصاء، لدى إجراء استقصاء تكون المجموعة السكانية فيه الأطفال الرُّضّع، ربما ينظر في إطارين اثنين محتملين: أحدهما يمكن أن يكون مرافق طبية تسجل المواليد خلال الاثني عشر شهراً الماضية؛ والآخر أسر معيشية يشمل ساكنوها أطفالاً تقل أعمارهم عن ١٢ شهراً. في الحالة الأولى سيضم الإطار جزأين اثنين، منهما جزء واحد لكل مرحلة من مراحل الاختيار: أولاً، قائمة المستشفيات والعيادات التي يولد فيها الأطفال؛ ثانياً قائمة بجميع الأطفال الذين ولدوا في تلك المرافق خلال الاثني عشر شهراً الماضية. وستكون وحدات الاختيار المرافق الطبية في المرحلة الأولى، والأطفال في المرحلة الثانية. وهكذا تكون وحدة الاختيار والمجموعة السكانية المستهدفة مصطلحين مترادفين في مرحلة الاختيار الأخيرة. غير أن من المرجح، في الحالة الثانية، أن يُعرّف الإطار (في مرحلة لاحقة من مراحل الاختيار) كقائمة أسر معيشية في وحدات مناطق صغيرة، كالقرى أو أحياء المدينة. ولدى تطبيق خطة المعاينة تُختار الأسر المعيشية، وتُستعرض للتحقق من وجود أطفال فيها تتراوح أعمارهم من صفر إلى ١٢ شهراً. وفي هذه الحالة تكون الأسرة المعيشية وحدة الاختيار التي يستند إليها احتمال الاختيار. غير أن من الجدير بالملاحظة أن أفراد المجموعة السكانية المستهدفة لا يُعرفون ولا يُستقصدون في الواقع إلا بعد أن تُستعرض الأسر للتحقق من وجودهم أو عدمه. وهكذا تكون وحدة الاختيار والمجموعة السكانية المستهدفة مختلفين في حالة إطار الأسرة المعيشية.

٥ - وحدة الاختيار في استقصاءات الأسر المعيشية — التي هي موضوع هذا الدليل — والوحدة التي يدور حولها تصميم العيّنات، تقوم على أساس الأسرة المعيشية. ومع ذلك، ستكون المجموعة السكانية المستهدفة مختلفة، حتى في الاستقصاء ذي الأغراض العامة. وذلك يتوقف على أهداف القياس. وفيما عدا استقصاءات إيرادات الأسرة المعيشية ونفقاتها، ستكون المجموعة السكانية المستهدفة في العادة مجموعة سكانية غير الأسرة المعيشية نفسها. ومن الأمثلة على ذلك استقصاءات العمالة، حيث تكون المجموعة السكانية المستهدفة بوجه عام أشخاصاً تبلغ أعمارهم ١٠ سنوات (أو ١٤ سنة) أو أكثر، وهذا يستبعد الأطفال الصغار جميعهم؛ واستقصاءات الصحة التناسلية للمرأة، حيث تضم المجموعة السكانية المستهدفة النساء في الفئة العمرية ١٤ - ٤٩ (وغالباً ما يكنّ النساء اللاتي تزوجن فقط من هذه الفئة العمرية)، إلخ.

## ٤ - ١ - ٢ خصائص أطر المعاينة

٦ - يجب أن يشمل إطار المعاينة طبعاً، بالمعنى الإحصائي، كما نوقش أعلاه، المجموعة السكانية المستهدفة. فيما وراء ذلك، إطار المعاينة الكامل هو الإطار الذي يكون كاملاً ودقيقاً وحديثاً. هذه صفات مثالية لا يمكن تحقيقها في استقصاءات الأسر المعيشية. ومع ذلك من الأساسي أن نجاهد لتحقيقها إما في بناء إطار من لا شيء وإما في استخدام إطار موجود من قبل. ويمكن تقدير نوعية الإطار من حيث مدى اتصال صفاته المثالية بالمجموعة السكانية المستهدفة. تذكر من الفصل الثالث أن تعريفنا للمعاينة الاحتمالية — التي يوجد فيها لكل فرد في المجموعة السكانية المستهدفة فرصة معروفة غير صفرية لاختياره — مقياس مفيد للحكم على جودة الإطار.

٧ - ستكون نتائج الاستقصاء متحيزة بطرق مختلفة، وذلك يتوقف على مدى الفشل في تحقيق كل واحدة من الصفات المثالية، لكن التحيز يكون في اتجاه تقليل تقدير المجموعة السكانية المستهدفة.

## ٤ - ١ - ٢ - ١ الكمال

٨ - يعتبر الإطار المثالي كاملاً فيما يتعلق بالمجموعة السكانية المستهدفة إذا كان جميع أفرادها (العالم) مشمولين بالإطار. لذلك فإن تغطية المجموعة (المجموعات) السكانية المستهدفة جانب أساسي في الحكم على ملاءمة الإطار للاستقصاء. وإذا لم يكن ملائماً فإن على فريق الاستقصاء أن يقدر إن كان في الإمكان إصلاحه أو زيادة تطويره لكي يصبح أكثر كمالاً. في المثال السابق، يكون الأطفال المولودون في البيت أو في أماكن أخرى خارج المرافق الطبية غير مشمولين بالاستقصاء إذا استُخدمت المرافق الطبية باعتبارها الإطار الوحيد للمعاينة. لذلك، ستوجد في هذا المثال أعداد كبيرة من المجموعة السكانية المستهدفة تكون فرصة شمولها بالعينة صفرًا، ويُنتهك شرط العينة الاحتمالية. نتيجةً لذلك، سيكون عدد الأطفال الرُّضّع أقلّ تقديرًا بواسطة إطار المرفق الطبي. علاوةً على ذلك، يرجحُ أن تكون خصائص الأطفال الرُّضّع مختلفة تمامًا عن الأطفال المولودين في البيت. لذلك سوف يسفر إطار المرفق الطبي عن توزيعات متحيزة لمؤشرات هامة للأطفال أو للعناية بهم.

٩ - يمكن أن تشكل التغطية غير الكافية أيضاً مشكلة في استقصاءات الأسر المعيشية. فقد تكون ثمة خطة استقصاء وطني يقصد بها أن تغطي جميع السكان بواسطة استقصاء للأسر المعيشية. غير أنه توجد مجموعات سكانية فرعية مختلفة، كالأشخاص الذين يعيشون في مؤسسات، وأسر من البدو الرحّل، وساكني القوارب، لا تقيم في أسر معيشية. وفي هذه الحالة، من البديهي أنه لا يمكن تحقيق تغطية شاملة لجميع السكان بواسطة استقصاء للأسر المعيشية. وينبغي وضع أطر إضافية لتغطية هذه المجموعات التي تعيش خارج الأسر المعيشية، لإعطاء أفرادها احتمالاً غير صفري لشمولهم بالاستقصاء. وما لم يحدث ذلك ينبغي تعديل وصف السكان المستهدفين بالفعل ليتسنى تعريف السكان المشمولين به بعناية أكبر. وبهذه الطريقة يمكن إبلاغ المستخدمين بوضوح أيّ قطاعات السكان مستبعدة من التغطية.

## ٤ - ١ - ٢ - ٢ الدقة

١٠ - الدقة سمة هامة أخرى من سمات أطر المعاينة، وإن كان من المرجح أن تقع أخطاء في الأطر غير الأطر المستخدمة لاستقصاءات الأسر المعيشية. ويمكن أن يُقال إن الإطار دقيق إذا شُمِلَ كل فرد من أفراد المجموعة السكانية المستهدفة مرة واحدة فقط لا أكثر. انظر في قائمة بتلك المؤسسات التجارية التي توظف أكثر من ٥٠ شخصاً. يمكن أن تحدث أخطاء إذا (أ) كانت في القائمة أية مؤسسة توظف ٤٩ عاملاً أو أقل، أو (ب) كانت أية مؤسسة لديها ٥٠ عاملاً أو أكثر ولم يكن اسمها موجوداً في القائمة، أو (ج) كانت مؤسسة موجودة في القائمة في أكثر من موضع (ربما بأسماء مختلفة).

١١ - في استقصاءات الأسر المعيشية يكون وقوع أخطاء كهذه في الأطر أقل احتمالاً. غير أنه يمكن أن تظهر حالات يقع فيها، مثلاً: (أ) إطار يتألف من ملف حاسوبي لمناطق تعداد ولا توجد فيه بعض عناصره، أو (ب) إطار قائمات بالأسر المعيشية في قرية ولا توجد فيه بعض الأسر القاطنة على أطراف القرية، أو (ج) إطار قائمات بالأسر المعيشية في وحدة منطقة، ووردت أسماء بعض الأسر في أكثر من وحدة، أو (د) إطار قائمة قديم بالأسر المعيشية لا يشمل مساكن بُنيت حديثاً. هذه الحالة الأخيرة، حالة الإطار غير المحدّث، ستناقش بمزيد من التفصيل أدناه.

١٢ - عدم وجود مناطق تعداد [في القائمة] أو وجود أسر معيشية في قائمة وحدة منطقة، يشير طبعاً إلى أن الأسر المتأثرة لا توجد أمامها فرصة لاختيارها في العينة. هنا أيضاً يشكّل هذا انتهاكاً لواحد من شروط إنتاج عينة احتمالية حقيقية. وكذلك تكرار اسم الأسرة في القائمات ينتهك معيار الاحتمال، مالم تؤخذ في الاعتبار لكي يتسنى حساب احتمالات الاختيار. من سوء الحظ أن الأسماء غير الواردة في القائمات لا تُضاف في أغلب الأحيان، ولا تشطب

الأسماء المكررة منها. لذلك ربما لا يدري فني المعاينة بالحاجة إلى تصحيح الإطار قبل أخذ العيّنات منه. من جهة أخرى، يمكن القول إن نسبة صغيرة من الحالات التي يُغفل فيها عن وضع أسماء في الإطار أو تُكرّر أسماء فيه لا تسبّب في العادة أيّ تحيز ذي شأن — أو حتى ملحوظ — في تقديرات الاستقصاء.

#### ٤ - ١ - ٢ - ٣ الإطار الحديث

١٣ - الشيء المثالي، طبعاً، هو أن يكون الإطار حديثاً لكي تتوفر فيه الخاصيتان الأخريان: الكمال والدقة. فمن البديهي أن الإطار القديم يحتوي على أخطاء ويرجح أن يكون غير كامل، لا سيما في استقصاءات الأسر المعيشية. المثال الجوهري للإطار الذي عفا عليه الزمن هو تعداد للسكان أُجريّ قبل عدة سنوات. فالتعداد القديم لا يعكس بدقة المباني الجديدة أو المساكن التي هُدمت، ودخول لاجئين في المساكن أو خروجهم منها، والمواليد أو المتوفين. فأوْجُه القصور هذه تنتهك الطلب الذي ينص على أن أمام كل فرد في المجموعة السكانية المستهدفة فرصة معلومة لأن يُنتخَب.

#### مثال

افترض أن الإطار يتألف من مناطق تعداد للسكان معرّفة وفقاً لأحدث عمليات التعداد، التي أُجريت قبل ٤ سنين ولم يُجرَ أيّ تحديث للإطار. وافترض أيضاً أن عدة مناطق مستقطنين أُقيمت في ضواحي العاصمة، في مناطق تعداد كانت عندما أُجريّ التعداد إما خالية وإما شبه خالية من السكان. تصميم العيّنة، في هذه الحالة، لن يعطي الأسر المعيشية القاطنة في مناطق التعداد التي كانت من قبل خالية أية فرصة لشمولها، وبذلك ينتهك الشروط الاحتمالية للعيّنة. وفي مناطق التعداد التي كانت شبه خالية تظهر مشكلة خطيرة أخرى، حتى وإن لم تنتهك مناطق التعداد هذه فنياً متطلبات الاحتمال المشمولة بالمعاينة. فمما لا شك فيه أن العيّنة ستُختار باستخدام احتمال متناسب مع الحجم، ومقاس الحجم هو تعداد السكان أو عدد الأسر المعيشية. ولما كان عدد سكان هذه المناطق صغيراً جداً وقت إجراء التعداد، فإن أيّ نموّ عالٍ جديد في مناطق التعداد هذه تتاح له فرصة ضئيلة فقط في أن يُختارَ عندما يستخدم احتمال متناسب مع الحجم. نتيجةً لذلك، يمكن أن يكون في العيّنة تباينٌ معاينة كبيرٌ إلى حد غير مقبول.

#### ٤ - ١ - ٣ أُطر المناطق

١٤ - نناقش في هذا الفرع والذي يليه الفئتين من الأُطر اللتين تُستخدمان في المعاينة، سواء في استقصاءات الأسر المعيشية أم في تطبيقات أخرى. ومن الأهمية بمكان ملاحظة أنه في التصميم المتعدد المراحل يجب أن يُعتبر إطار كل مرحلة عنصراً مستقلاً. فالإطار المحدد يختلف في كل مرحلة عنه في الأخرى. ويرجح أن يستخدم تصميم العيّنة لاستقصاء الأسر المعيشية إطار منطقة (يُنَاقَشُ في هذا الفرع) للمراحل الأولى وإطار قائمة (يُنَاقَشُ في الفرع التالي) للمرحلة الأخيرة.

١٥ - في استقصاءات الأسر المعيشية، يشمل إطار معاينة المناطق الوحدات الجغرافية للبلد في تسلسل هرمي. وتحمل الوحدات بطاقات تعريف مختلفة إدارياً بين بلد وآخر، لكنها في الأحوال النمطية تشمل مصطلحات مرتبة ترتيباً تنازلياً، مثل إقليم أو محافظة؛ مقاطعة، ناحية، حي، قرية (في المناطق الريفية) أو حي (في المناطق الحضرية). لأغراض تعداد السكان، يُزاد تصنيف التقسيمات الفرعية الإدارية إلى كيانات، مثل مناطق رئيس

الفريق، ومناطق التعداد. وغالباً ما تكون منطقة التعداد في التعداد السكاني أصغر وحدة جغرافية تعرّف وترسم حدودها في البلد.

١٦ - لأغراض الاستقصاء، توجد أربع خصائص متميزة للوحدات الجغرافية وهي هامة لتصميم العينات:

(أ) أنها في العادة تغطي كامل مساحة البر في الدولة؛

(ب) حدودها مرسومة جيداً؛

(ج) توجد أرقام لعدد السكان متاحة لها؛

(د) رُسمت لها خرائط.

١٧ - تغطية جميع المساحة الجغرافية للدولة مسألة هامة، كما لاحظنا، لأن ذلك واحد من المعايير لتحقيق عينة احتمالية صحيحة. الحدود المرسمة جيداً، المرسومة لها خرائط، قيّمة جداً في تنفيذ العينة لأنها تحدد المواقع التي سيُجرى فيها العمل الميداني. وتساعد المعلومات الجيدة عن الحدود أيضاً المُقابل على تحديد مواقع الأسر المعيشية المدرجة في العينة التي تُختار في نهاية المطاف لإجراء المقابلات معها. وتلزم أرقام عدد السكان في تصميم العينات لتخصيص مقاسات الحجم ولحساب احتمالات الاختيار.

١٨ - نقطة البداية العادية في وضع إطار منطقة لاستقصاءات الأسر المعيشية هي تعداد لسكان البلد يستند إلى العوامل الأربعة الوارد ذكرها أعلاه. يضاف إلى ذلك أن منطقة التعداد وحدة جغرافية ذات حجم ملائم للاختيار في المراحل الأخيرة من المعاينة (المرحلة قبل الأخيرة في التصميم ذي المرحلتين). في معظم البلدان، تنشأ مناطق التعداد عن قصد لتحتوي على أعداد متساوية تقريباً من الأسر المعيشية — نحو ١٠٠ أسرة في أغلب الأحيان — بغية توفير أعباء عمل مقارنة للعاملين في إجراء التعداد.

١٩ - إطار المنطقة هو أيضاً قائمة، وفي هذا تناقض ظاهري — لأن على المرء أن يبدأ بقائمة بوحدات السكان الجغرافية الإدارية للمضي في المراحل الأولى من اختيار عينة لاستقصاء الأسر المعيشية. وهذا يحتاج إلى مناقشة أطر القوائم.

## ٤ - ١ - ٤ أطر القوائم

٢٠ - إطار القائمة للمعاينة هو، بكل بساطة، إطار مكوّن من قائمة بالوحدات السكانية المستهدفة. من الناحية النظرية يوجد إطار قائمة لاستقصاءات الأسر المعيشية لكل بلد، بعد إجراء تعداد للسكان فيها مباشرة. ويوفّر التعداد الجديد، من حيث المبدأ، قوائم مرتبة جغرافياً بكل أسرة معيشية — أو وحدة سكنية — في البلد.

٢١ - قائمة التعداد المنجزة حديثاً قائمة مثالية لإطار معاينة الأسر المعيشية لأنها أحدث وأكمل وأدق ما يمكن أن تكون عليه أيّ قائمة بالأسر المعيشية على الإطلاق. ونظراً إلى الترتيب الجغرافي لقائمة التعداد، يكون تقسيمها إلى طبقات لتوزيع العينة توزيعاً جغرافياً صحيحاً عملية بسيطة إلى حد ما. لذلك، عندما توجد حاجة إلى إجراء استقصاء بالعينة لمتابعة التعداد بغية الحصول على معلومات تكميلية أو معلومات أكثر تفصيلاً من المعلومات التي يستطيع التعداد ان يوفرها بكفاءة، تكون قائمة التعداد الجديدة بذلك مناسبة بشكل مثالي للاستعمال كإطار قائمة. غير أن من الأهمية بمكان الاعتراف بأن قائمة التعداد الجديدة لا تكون متاحة كإطار حديث إلا لفترة وجيزة. ومن البديهي أنه كلما طالت الفترة الفاصلة بين التعداد واستقصاء المتابعة، قلّت فائدة قوائم التعداد كمصدر للإطار.

٢٢ - توجد قوائم أخرى يمكن أن تعتبر أطرّ معاينة ملائمة لاستقصاءات الأسر المعيشية، وذلك يتوقف على جودتها، ومنها السجلات المدنية، وسجلات وصلات المنافع العامة بالمسكن. وتكون السجلات المدنية مرشحة للاستخدام في البلدان التي تحتفظ بسجلات لمواطنيها وعناوينهم. وربما تكون في بعض الحالات أكثر فائدة من إطار تعداد قائم على أساس المناطق، لأن من المرجح أن يُستكمل السجل باستمرار. وربما تكون وصلات المنافع العامة — الكهرباء في العادة — مفيدة كإطار معاينة عندما يكون تعداد سكان البلد قديماً جداً، لكنها يجب طبعاً أن يتم تقييمها لتقدير مشاكلها المحتملة وآثارها. وثمة مشكلة بديهية تؤدي إلى التقصير في التغطية عندما يتعذر حصول الأسر المعيشية على الطاقة الكهربائية؛ وثمة مشكلة أخرى تحتاج إلى الفرز وهي أن توجد نقاط وصل كهربائية تخدم أسراً معيشية متعددة.

٢٣ - يوجد إطار قائمة آخر يُستعمل على نطاق واسع في البلدان المتقدمة، وهو سجلّ المشتركين في الهواتف. تُجرى المعاينة بتقنيات الاتصال بالمشتركين بالهاتف اتصالاً عشوائياً على الأرقام دون الاطلاع على الأسماء لضمان حصول المشتركين الذين أرقامهم غير منشورة في الدليل على فرصة صحيحة لأن يُختاروا. غير أن المعاينة بواسطة الاتصال العشوائي بالأرقام لا يوصى به في البلدان التي تكون نسبة المشتركين في الهاتف فيها منخفضة.

٢٤ - في الاستقصاء التقليدي للأسر المعيشية نرى أن المرحلة الأخيرة من الاختيار، دون استثناء، تقوم على أساس مفهوم إطار القائمة. وقد ناقشنا من قبل كيف أن مرحلة التصميم قبل الأخيرة ربما تسفر عن عيّنة مجموعات تُجمَع فيها القوائم الجديدة للأسر المعيشية. ومن تلك القائمة تختار الأسر المعيشية التي ستُدْرَج في العيّنة. هكذا، يوجد لدينا إطار منطقة يُعرّف مجموعات العيّنة وإطار قائمة يُعرّف الأسر المعيشية التي تختار من المجموعات للعيّنة.

#### ٤ - ١ - ٥ الأطرّ المتعددة

٢٥ - ناقش الفصل الثالث المعاينة على مرحلتين في استقصاءات الأسر المعيشية. وهو ينطوي على استخدام تقنيات الاستعراض لتعريف مجموعة مستهدفة محددة في المرحلة الأولى، تليها مقابلة في المرحلة الثانية في معاينة فرعية للأسر التي تم تعريفها. وثمة تقنية معاينة أخرى ربما تحقق نفس النتيجة النهائية تقريباً، تستخدم أكثر من إطار معاينة واحد. يُستخدم فيها عادةً إطاران اثنان فقط، ويوضع لها تصميم إطار مزدوج؛ غير أنه يمكن أحياناً استخدام ثلاثة أطرّ وأكثر (تصميم الأطرّ المتعددة). مثال ذلك الإطار السكاني الذي يُعرّف بأنه مجموع عدة قوائم تُختار منها عيّنة مستقلة، وفي هذه الحالة يُصبح كل إطار فرعي طبقةً (انظر مناقشة التقسيم إلى طبقات في الفرع ٣ - ٤ - ٢ من الفصل الثالث، وفي الملحق الأول). غير أن المشكلة المشتركة بين كل هذه الأطرّ هي التكرار.

#### ٤ - ١ - ٥ - ١ الإطار المزدوج النمطي في استقصاءات الأسر المعيشية

٢٦ - لتبسيط العرض سوف نناقش تصاميم الإطار المزدوج، وإن كانت المبادئ مشابهة لمبادئ تصاميم الأطرّ المتعددة. تستدعي المنهجية بوجه عام الجمع بين إطار منطقة سكانية عام وإطار قائمة بالأشخاص الذين يُعرّف أنهم أفراد في مجموعة سكانية مستهدفة محددة قيد الدراسة. انظر، مثلاً، استقصاء يُقصد به دراسة خصائص الأشخاص العاطلين عن العمل. يمكن أن يقوم الاستقصاء على أساس إطار منطقة للأسر المعيشية، لكن يمكن تكملته بعيّنة من إطار قائمة بأشخاص عاطلين عن العمل في الوقت الراهن وهم مسجلون لدى وزارة الخدمات الاجتماعية. والهدف من عيّنة إطار مزدوج من هذا النوع هو زيادة حجم العيّنة بأشخاص لديهم احتمال عالٍ

جداً لأن يُشركوا في المجموعة السكانية المستهدفة. ويمكن أن يكون النهج بديلاً أرخص وأكثر كفاءة للعيّنة ذات المرحلتين. من الضروري أن يستخدم إطار الأسر المعيشية ذو الغرض العام لكي يأخذ في الحسبان أفراد المجموعة السكانية المستهدفة الذين لا توجد أسماؤهم في القائمة. وفي هذا المثال يكونون أشخاصاً غير مسجلين لدى وزارة الخدمات الاجتماعية.

٢٧ - بيد أنه تُفرض بضعة أوجه قصور عند استخدام تصاميم الإطار المزدوج. وأحد هذه الأوجه هو أن إطار القائمة يجب أن يكون حديثاً تقريباً. فإذا كانت نسبة كبيرة من الأشخاص المختارين من القائمة قد تغيرت أوضاعهم فهذا التغير يخرجهم من المجموعة السكانية المستهدفة، وبذلك يكون استخدام الإطار استخداماً تنقصه الكفاءة. وفي المثال الذي أماننا، توضّح حقيقة أن أيّ شخص عاطل عن العمل يصبح غير مؤهل، إن هو أصبح موظفاً عندما ينتهي الاستقصاء، لماذا يجب أن يكون إطار القائمة حديثاً جداً.

٢٨ - ثمة وجه قصور آخر تفرضه حقيقة أنه يرجّح أن تكون أماكن إقامة الأشخاص الواردة أسماؤهم في إطار القائمة قد تفرقت في سائر أنحاء المجتمع، مما يجعل مقابلتهم باهظة الكلفة بسبب السفر. وهذا يختلف طبعاً اختلافاً بيناً مع إطار الأسر المعيشية القائم على منطقة محددة، حيث يمكن اختيار العيّنة في مجموعات لتقليل تكاليف إجراء المقابلات.

٢٩ - ثمة مسألة خطيرة متصلة بتصاميم الإطار المزدوج، وهي مسألة التكرار. فالأشخاص المشمولون بإطار القائمة سيُشمولون بوجه عام بإطار المنطقة أيضاً. وهنا أيضاً في المثال الذي أماننا نجد أن الأشخاص العاطلين عن العمل الذين اختيروا من السجل أفراد في أسر معيشية. لذلك توجد أمامهم فرصة مكررة لأن يُختاروا عندما يُستخدم الإطاران كلاهما. ويمكن أن تعالج مسألة التكرار بطريقة تُعدّل فيها على الوجه الصحيح. غير أن هذا له آثار على مضمون استبيان الاستقصاء. ففي مثالنا هذا، كل شخص عاطل عن العمل تُجرى معه مقابلة يُسأل إن كان مسجلاً لدى الوزارة في قائمة الأشخاص العاطلين عن العمل. والذين يردّون بالإيجاب يلزم إجراء مزيد من العمل بشأنهم لمطابقة أسماؤهم مع إطار القائمة، وهذه العملية عرضة للخطأ ومحفوفة بالتعقيدات. فإذا نجحنا في العثور على اسم مطابق ينبغي تغيير مُعامل الترجيح في الاستقصاء الخاص بالشخص المتأثر به إلى  $(1/P_i + 1/P_j)$  لكي يعكس حقيقة أن لديه احتمال  $P_j$  لأن يُختار من إطار الأسر المعيشية وخيار  $P_i$ ، لأن يُختار من إطار القائمة. ومن الأهمية بمكان أن نلاحظ أن مقابلة الأسماء يجب أن تُجرى على أساس إطار القائمة كله لا مع مجرد الأشخاص الواردة أسماؤهم في الإطار وحدث أنهم اختيروا في العيّنة. هذا لأن الاحتمال (ومُعامل الترجيح) هو دالةٌ لفرصة الاختيار بغض النظر عما إذا كان الشخص قد اختير أم لا.

#### ٤ - ١ - ٥ - ٢ الأطر المتعددة لأنواع مختلفة من أماكن السكن

٣٠ - يحدث نوع آخر من معاينة الإطار المزدوج عندما تكون المجموعة السكانية المستهدفة في أنواع مختلفة من أماكن السكن لا يحدث بينها تراكب. مثال ذلك أن استقصاء الأيتام يرجّح أن يصمّم على نحو يشمل معه أيتاماً يعيشون في أيّ واحد من نوعين اثنين من ترتيبات الإسكان. الأول مؤسسات كدور الأيتام تشكّل إطاراً. والثاني أسر معيشية ينبغي أخذ عينات منها لتشمل الأيتام الذين يعيشون فيها مع أحد الوالدين يكون باقياً على قيد الحياة أو مع أقرباء آخرين أو مع غير الأقرباء. لذلك، يتكوّن تصميم الإطار المزدوج من إطار أسر معيشية وإطار مؤسسي، وهما طبعاً غير متراكبين.

٣١ - الهدف من التصميم الذي من هذا النوع هو شمول السكان المستهدفين شمولاً كافياً (أقرب ما يكون من ١٠٠ في المائة). وعندما تعيش أعداد كبيرة من السكان في أيّ واحد من نوعين اثنين من أماكن السكن، تحدث

تحيّزات كبيرة إن كانت العيّنة مقصورة على واحد فقط من الإطارين. فعيّنة الأيتام القائمة، مثلاً، على أساس المقيمين في أسر معيشية تعطينا تقديراً ناقصاً لمجموعة الأيتام لكنها تعطينا تقديراً متحيزاً من حيث خصائصهم. وتحدث تحيزات مشابهة في استقصاء يقوم حصراً على أساس الأطفال المقيمين في مؤسسات.

٣٢ - وجه القصور الذي نوقش أعلاه فيما يتعلق بالتكرار لا يتعلق بتصاميم من أطر مزدوجة غير متراكبة. وهي، لهذا السبب، أسهل إدارة بكثير.

#### ٤ - ١ - ٦ الإطار النمطي (الأطر النمطية) في تصاميم ذات مرحلتين

٣٣ - أكد الفصل الثالث القيمة العلمية لتصميم العيّنات ذي المرحلتين. ويناقش هذا الفرع الإطار الذي يُستخدم في الأحوال النمطية تصاميم ذات مرحلتين.

٣٤ - الوحدات الجغرافية، أو المجموعات، التي تُختار في المرحلة الأولى من مراحل الاختيار تُعرّف بأنها قري (أو أجزاء من قري) أو مناطق تعداد في المناطق الريفية، أو أحياء في مدن في المناطق الحضرية. لذلك يتألف الإطار من جميع الوحدات الجغرافية التي تشكّل عالم الدراسة أيّاً كان تعريفها — الأمة كلها، أو إقليم أو مجموعة أقاليم، أو العاصمة. تُجرى المعاينة بتجميع قائمة الوحدات، وبتفحصها للتأكد من اكتمالها، وتصنف القائمة في طبقات بطريقة ملائمة (غالباً ما يكون التصنيف جغرافياً) ثم تُختار (عادةً باحتمال متناسب مع الحجم) عيّنة منهجية من الوحدات.

٣٥ - إذا كان ملف المجموعات من العالم كبيراً جداً، قد يلزم وضع مراحل صورية متوسطة للاختيار، كما نوقش في الفصل السابق. وفي هذه الحالة، تعرّف وحدات الإطار تعريفاً مختلفاً لكل مرحلة من المراحل الصورية. في المثال السابق، المتعلق ببنغلاديش، عُرفت وحدات الإطار للمرحلتين الصورتين بثانات واتحادات.

٣٦ - أما وحدات الإطار في المرحلة الثانية، في تصميم ذي مرحلتين، فهي ببساطة الأسر المعيشية في مجموعات العيّنة في المرحلة الأولى. وعندما تؤخذ العيّنة من قائمة الأسر المعيشية، يكون الإطار، بحكم التعريف، إطار قائمة. ويمكن أيضاً أن تؤخذ العيّنات منها كقطاعات صغيرة أنشئت بتقسيم المجموعات إلى أجزاء جغرافية تكون جامعة مانعة ويَجِبُ بعضها بعضاً. وفي تلك الحالة يكون إطار المرحلة الثانية إطار منطقة.

#### ٤ - ١ - ٧ أطر المعاينة الرئيسية

٣٧ - هنا سنذكر بإيجاز مفهوم إطار المعاينة الرئيسي، الذي يناقشُ بقدر كبير من التفصيل في الفرع ٤ - ٢ أدناه.

٣٨ - لإطار المعاينة الرئيسي إطار يُستخدم لاختيار عيّنات إما لاستقصاءات متعددة، لكل منها مضمون مختلف، وإما لاستخدامه في حلقات مختلفة من استقصاء مستمر أو دوري. وفيما عدا التحديث الدوري عند الاقتضاء، يظل إطار المعاينة نفسه على حاله لا يتغير من استقصاء إلى آخر أو من حلقة إلى أخرى من الاستقصاء نفسه. وبدلاً من ذلك — وهذه خاصيته المتميزة — يُصمّم إطار المعاينة الرئيسي ويبنى ليكون إطاراً ثابتاً راسخاً لاختيار العيّنات الفرعية اللازمة لاستقصاءات خاصة أو دورات من الاستقصاء نفسه على مدى فترة ممتدة من الزمن.

## ٤ - ١ - ٨ المشاكل المشتركة بين الأطر والإصلاحات المقترحة

٣٩ - المشاكل التي تظهر في استقصاءات الأسر المعيشية من الأطر المعيبة تشمل تحيزاً من غير تحيزات المعاينة وتبايناً في المعاينة، على حد سواء. وكما قيل ضمناً من قبل، تقع المشاكل المشتركة عندما يكون إطار المعاينة قد عفا عليه الزمن، أو غير دقيق، أو غير كامل. في الأغلبية الساحقة من الاستقصاءات الوطنية العامة الأغراض، يكون الإطار الأساسي هو أحدث تعداد للسكان، وهو الإطار المشار إليه في هذا الفرع. وغالباً ما تحدث مشاكل البلى وعدم الدقة وعدم الاكتمال مجتمعة في الأطر القائمة على أساس تعداد السكان. وهي تميل إلى التعاضم كلما ازدادت الفترة الفاصلة بين التعداد والاستقصاء.

٤٠ - دُكر أن الإطار يجب أن يكون حديثاً جداً لكي يعكس الوضع الراهن للسكان. فالإطار الذي يقوم على تعداد أُجري قبل خمس سنوات، مثلاً، لا يأخذ في الحسبان بشكل كافٍ نمو السكان والهجرة. حتى إطار التعداد الحديث يمكن أن يكون غير كامل ويسبب مشاكل إزاء استقصاءات الأسر المعيشية إن لم يشمل الثكنات العسكرية وسكان القوارب والبدو الرحل وغيرهم من المجموعات السكانية الفرعية التي لا تعيش في ترتيبات أسر معيشية تقليدية. الأخطاء الواردة في أطر التعداد، الجديدة منها والقديمة، تشكل مشاكل متنوعة من بينها مشاكل تتعلق بتكرار قائمات الأسر المعيشية، وإغفال تسجيل الأسر، أو الأسر المعدودة أو المشفرة في منطقة التعداد الخاطئ.

٤١ - تكون الاستراتيجيات الملائمة لمعالجة أطر التعداد القديمة أو غير الدقيقة أو غير المكتملة، جزئياً، رهينة: (أ) أهداف الاستقصاء و(ب) عمر الإطار. ففيما يتعلق بأهداف القياس، إذا صُمم الاستقصاء عن قصد ليشمل الأسر المعيشية غير المتنقلة فقط فإن إطار التعداد الذي يستثني أسر البدو الرحل يكفي. ومن جهة أخرى، ينبغي تطوير إجراء لإنشاء إطار لأسر البدو الرحل إذا قصد بالاستقصاء أن يشملهم (في البلدان التي يوجدون فيها). وفي هذا الصدد، تتوقف مسألة ما إذا كان إطار التعداد كاملاً أم لا على تعريف المجموعة السكانية المستهدفة أو المجموعات السكانية الفرعية المستهدفة التي يُراد شمولها بالاستقصاء.

٤٢ - تختلف إصلاحات المشاكل الناجمة عن البلى والأخطاء، وذلك يتوقف على مدى تقادم التعداد. فمع أنه ربما لا يكون من الحكمة أن نضع قاعدة دقيقة، ينبغي بسبب تنوع الأحوال الوطنية أن نكتفي بقاعدة تقريبية تسترشد بها استراتيجية ملائمة لتصحيح الإطار أو تحديثه، وتحدد ما إذا كان التعداد قد مضى عليه أكثر من سنتين. وفيما يتعلق بالأخطاء التي هي من النوع المذكور في الفقرة ٤٠، تنطبق عليها الإصلاحات الواردة في الفرعين ٤ - ١ - ٢ و ٤ - ١ - ٨.

## ٤ - ١ - ٨ - ١ أطر التعداد التي مضى عليها أكثر من سنتين

٤٣ - ينطبق الوضع الأول على البلدان التي أجرت تعدادات قديمة — تعدادات مضى عليها سنتان أو أكثر. هذه الأطر القديمة هي التي تشكل أكبر التحديات لتصميم العينات لاستقصاء الأسر المعيشية، لا سيما في المدن السريعة النمو. وإن تحديث إطار التعداد القديم في البلد بأسره هو الحل الأمثل لأنه، إن نجح، سيضمن أن تكون بيانات الاستقصاء الناشئة أكثر دقة بكثير، من حيث تغطية الاستقصاء وأكثر ما تكون موثوقة. وهي من سوء الحظ أعلى ما تكون تكلفة واستنزافاً للوقت وهي لذلك غير عملية. ومع ذلك، ربما لا يوجد بديل في هذه البلدان التي تقادم فيها التعداد إلى درجة خطيرة.

٤٤ - يوجد، بدلاً من التحديث الكامل، حلٌ وسط يُحدثُ بموجبه الإطار في المناطق المستهدفة فقط، ويعرّفُ هذه المناطق خبراء قُطريون لهم معرفة بأنماط النمو والتحوّلات الديمغرافية. وإن تحديث إطار التعداد مسألة

بسيطة إلى حد ما: فكل ما يلزم هو مقياس جديد للحجم. لأغراض تحديث الإطار، يُعرّف مقياس الحجم بأنه عدد وحدات السكن، لا عدد الأسر المعيشية أو الأشخاص.

٤٥ - من الأهمية بمكان الاعتراف بأن مقياس الحجم لا يحتاج إلى أن يكون دقيقاً لكي تكون منهجية العيّنة صحيحة. فإذا ساد الظنُّ، مثلاً، بأن منطقة تعداد تضم ١٢٢ أسرة معيشية بناءً على أساس التعداد السابق، لا داعي للقلق إذا كان فيها الآن ١١٥ أو ١٣٢ أسرة. لهذا السبب ليس من المفيد محاولة تحديث الإطار في الأحياء القديمة الراسخة التي تغيرت قليلاً على مر العقود، حتى وإن كان السكان الأفراد يذهبون ويجيئون. إنما الذي يهم، بدلاً من ذلك، هو الفرق الجذري في الوضع الراهن مقارنةً بما كان عليه في وقت إجراء التعداد — بمعنى أن يوجد ٢٥٠ أسرة معيشية بينما كان يُتوقع أن تكون هناك ١٠٠ أسرة. هذه الأوضاع يرجح أن تحدث في أحياء تمر بنمو كبير أو تُهدم فيها منازل، مثل أحياء المستوطنين على أطراف المدينة، وإقامة المباني الشاهقة في مواقع وهدم بيوت في مواقع أخرى. هذه المناطق هي المناطق المستهدفة للتحديث. وسوف يُعتمد على المتعاونين والخبراء القطريين للمساعدة في تعيين المناطق المستهدفة ويشمل هذا طبعاً المناطق التي حدثت فيها تغيرات بعد التعداد فقط.

٤٦ - يستتبع التحديث بوجه عام عدة خطوات من بينها (أ) تعيين مناطق التعداد التي تشكل المناطق المستهدفة، و(ب) طواف سريع لعدّ مناطق التعداد المتأثرة لمعرفة مقياس الحجم الحالي، و(ج) مراجعة ملف التعداد لرؤية مقياس الحجم المحدث. وحقيقة أن مقياساً تقريبياً للحجم يكفي، كما ذكر، هي السبب في الدعوة إلى الطواف لإجراء عدّ سريع لتعيين وحدات السكن بدلاً من الأسر المعيشية. وليس من الضروري في الطواف للعدّ السريع أن تُطرق الأبواب لعدّ الوحدات السكنية إلا، ربما، في المساكن المتعددة الوحدات، التي لا يكون فيها عدد الوحدات ظاهراً إلا بدخول المبنى.

٤٧ - تحديث أطر التعداد القديم بالطريقة التي وُصفت أعلاه ضروري لتثبيت احتمالات الاختيار لوحدات المرحلة قبل الأخيرة، من ثمّ موثوقية تقديرات الاستقصاء. يساعد التحديث، على وجه الخصوص، على التحكم لا في الحجم الإجمالي للعيّنة فقط، وإنما أيضاً في أعباء عمل الموظفين الميدانيين في وضع القوائم وإجراء المقابلات. علاوة على ذلك، يقلل احتمالات مواجهة مجموعات كبيرة في الميدان يتبين أنها أكبر كثيراً ممّا كان متوقعاً والاضطرار إلى تقسيمها إلى عيّنات فرعية أو اتخاذ إجراء آخر بشأنها. ومما يتصل بالنقطة الأخيرة حقيقة أن التقسيم إلى عيّنات فرعية يتطلب مُعاملات ترجيح للتعديلات — وفي هذا تعقيد لتجهيز البيانات. سيُخفّض هذا الاحتمال إلى حد لا تواجه فيه مجموعات كبيرة غير متوقعة بعد اختيار العيّنات في المرحلة قبل الأخيرة.

٤٨ - يرجّح أن يستتبع تصميم العيّنة المعيارية لاستقصاء الأسر المعيشية تجميع القوائم الراهنة للأسر المعيشية في المجموعات المختارة للعيّنة. في هذه الحالة، يحدث أيضاً تحديث الاختيار في المرحلة قبل الأخيرة (انظر الإشارة إلى الاحتمال المناسب مع الحجم المقدّر في اختيار العيّنات، الذي ينطبق — هنا أيضاً — في الفرع التالي). لهذا السبب ربما تشبه القوائم الراهنة للمجموعات المختارة للعيّنة، التي لم يتم تحديثها، قوائم التعداد شبيهاً قوياً (وإن كان هذا غير مضمون). غير أنه يمكن أن نتوقع أن تسفر المجموعات المختارة للعيّنة من الفرع المحدث من إطار التعداد عن قوائم حديثة تختلف اختلافاً كبيراً عن قوائم التعداد — من حيث مجموع عدد الأسر المعيشية وتعيينها على وجه التحديد أيضاً.

٤٩ - ينبغي هنا إبداء نقطة أخيرة واحدة حول استخدام تعداد قديم، وهذا يتعلق بصحة العيّنة لا بتباين العيّنات. تُختار المجموعات عادةً، كما ذكر سابقاً، باحتمال متناسب مع الحجم. ومن شأن عدم تحديث مقياس الحجم في حالة المجموعات التي شهدت نمواً كبيراً، قبل اختيار العيّنة، أن يسفر عن تمثيل المناطق التي كانت فيها

أعداد صغيرة من الأسر المعيشية في وقت التعداد، لكنها شهدت نمواً كبيراً بعد ذلك، تمثيلاً قليلاً إلى حد خطير. وستكون نتائج الاستقصاء متحيزة ومضللة طبعاً، لأنه يرجح أن تكون خصائص الأشخاص الذين يعيشون في هذه المناطق التي شهدت نمواً كبيراً مختلفة جداً عن خصائص الأشخاص الذين يعيشون في الأحياء الأكثر استقراراً.

#### ٤ - ١ - ٨ - ٢ أطر التعداد التي مضى عليها سنتان أو أقل

٥٠ - ينطبق هذا الفرع على البلدان التي أجريت فيها تعدادات — حديثة نسبياً إذ أجريت خلال السنة الماضية أو السنتين الماضيتين — ولذلك لا تحتاج إلى تحديث عام للإطار. في هذه الحالات، تُختار المجموعات باستخدام مقياس الحجم للتعداد الأصلي، لأن من المتوقع أن يكون دقيقاً. فالتحديث في حد ذاته لا يُجرى إلا في المرحلة قبل الأخيرة من مراحل الاختيار عندما يقوم موظفو الميدان بإعداد قوائم جديدة للأسر المعيشية في مجموعات العينة. وستُختار الأسر المعيشية للعينة من القوائم الراهنة وتعدّل مُعاملات الترجيح في العينات، حسب الاقتضاء، وفقاً للإجراءات التي نوقشت في الفرع ٣ - ٧ - ٢ فيما يتعلق بكون احتمال الاختيار للعينة متناسباً مع الحجم المقدّر.

٥١ - بينما يمكن أن تكون بضع مجموعات في عالم الإطار قد نمت نمواً كبيراً منذ إجراء التعداد، لا يُتوقع أن يكون عدد هذه الحالات كبيراً جداً بحيث يؤثر تأثيراً كبيراً إما في العمليات الميدانية وإما في دقة الاستقصاء. وإذا حدث أن وقعت مجموعات من هذا القبيل في العينة يمكن تقسيمها إلى قطاعات فرعية إذا لزم ذلك. ومما يُذكر أن التقسيم إلى قطاعات فرعية، أو "التقطيع"، كما يُقال عنه، إجراء ميداني يُقصد به تخفيف عبء العمل في وضع القوائم. وينطوي هذا الإجراء على (أ) تقسيم المجموعة الأصلية إلى أجزاء، تكون في العادة أرباعاً، و(ب) اختيار واحد منها اختياراً عشوائياً لإدراجه في القائمة، و(ج) اختيار الأسر المعيشية المراد مقابلتها من ذلك القطاع. هذا التقسيم إلى قطاعات لا يُحسّن موثوقية المعاينة لأن كل قطعة من العينة ستحمل مُعاملَ ترجيح مساوياً لعدد القطع في المجموعة — عامل مقداره ٤ إن كان المجموعة قد قسمت إلى أرباع. غير أن التقسيم الفرعي يساعد في احتواء التكاليف الميدانية. وربما تحدث الحاجة إلى التقطيع حتى وإن كان التعداد حديث العهد، وذلك في مناطق التعداد التي شهدت نمواً كبيراً منذ إجراء التعداد. وعندما يكون التعداد حديثاً جداً يُتوقع أن يكون عدد هذه المناطق قليلاً جداً.

٥٢ - ينبغي ملاحظة أن أنواع الأخطاء التي ذُكرت من قبل ( تكرار التسجيل، أو إغفال تسجيل الأسر المعيشية، أو الخطأ في تعيين منطقة التعداد) تصحّح جزئياً عندما يُجرى تحديث يستتبع وضع قوائم جديدة للأسر المعيشية في المرحلة قبل الأخيرة. هذا طبعاً سبب قوي آخر للحصول على قوائم جديدة للأسر المعيشية في الاستقصاءات.

#### ٤ - ١ - ٨ - ٣ عندما يُستخدم الإطار لغرض آخر

٥٣ - يتساءل مديرو الاستقصاءات أحياناً إن كان إطار الأسر المعيشية الذي وُضع خصيصاً لنوع من الاستقصاءات يمكن أن يصلح لاستقصاء آخر. هل يمكن لإطار معاينة يقصد به استقصاء للقوى العاملة، مثلاً، أن يُستخدم في تصميم عينة لقياس الأحوال الصحية أو الإعاقة أو الفقر أو الحيازات الزراعية؟ غير أنه ليس الإطار نفسه هو الذي تكون فيه المشكلة عادةً، وإنما الطريقة التي يُقسّم بها إلى طبقات. يمكن أن تُستخدم الأطر، في معظم الحالات، لاستقصاءات مختلفة ما لم تكن غير كاملة وغير دقيقة وقد فات وأنها للاستقصاء المقصود

استعمالها فيه. انظر أدناه مثلاً للاطلاع على إطار غير ملائم لاستعمالات متعددة. فإذا كان استقصاء يركز على تكلفة المعيشة، مثلاً، يستند إلى المجتمعات الحضرية فقط (هذا ما يكون عليه الحال في الواقع غالباً)، سيستثني إطار المعاينة المناطق الريفية. ومن الواضح أن هذا الإطار لن يكون مناسباً لتقدير الفقر في البلدان التي يكون الفقر فيها أساساً ظاهرة ريفية.

٥٤ - غير أن معظم استقصاءات الأسر المعيشية تكون استقصاءات ذات غرض عام، لا من حيث مضمونها فقط، وإنما أيضاً من حيث تصميم عيناتها. واستقصاء القوى العاملة، مثلاً، يشمل في العادة معلومات مساعدة في الديمغرافيا، والتحصيل العلمي، ومواضيع أخرى. وفي هذه الحالات، يكون تصميم العيّنات المناسب ذا غرض عام أيضاً، مما يوحي باستخدام إطار معاينة تقليدي — إطار يشمل جميع الأسر المعيشية في الدولة. ويمكن تقسيم الإطار إلى طبقات، حول متغير يُعنى خصيصاً بقياس القوى العاملة. وربما تصنف مناطق التعداد، مثلاً، وفقاً للمتغير، الذي هو النسبة المئوية للعاطلين عن العمل في التعداد الأخير. ربما تُنشأ آنذاك ثلاث طبقات من مناطق التعداد — واحدة تكون فيها نسبة البطالة منخفضة، والثانية متوسطة والثالثة عالية. هذا، كما ذكر آنفاً، قرار تقسيم إلى طبقات. الإطار نفسه لا يتأثر به. وسيكون الحل "إلغاء تقسيم الإطار إلى طبقات" إذا أُريد استعماله في استقصاء آخر كاستقصاء صحي، مثلاً.

٥٥ - من المهام الأساسية لخبير المعاينة الإحصائي تقدير إطار المعاينة في عين المكان، عندما يراد استخدامه لنوع آخر من الاستقصاءات. وسوف يستتبع التقدير ضمان كون الإطار بالشكل الذي وُضع فيه يستطيع تلبية أهداف القياس في الاستقصاء المقترح، لا سيما من حيث الكمال والدقة والحدثة (التي تناقش في كل فروع هذا الفصل).

## ٤ - ٢ أطر المعاينة الرئيسية

٥٦ - يمكن أن تكون العيّنات الرئيسية فعّالة التكلفة وذات كفاءة من حيث التكلفة عندما يُجري البلد عدداً كافياً من الاستقصاءات أو حلقات دورية من الاستقصاء نفسه لاستدامة استخدامها. وربما يكون غنياً عن البيان أنه يجب تصميمها تصميماً صحيحاً لكن من الأهمية بمكان أيضاً صيانتها على الوجه الصحيح على مر الزمن. ومما يُذكر أن الأمم المتحدة (١٩٨٦) تصف معاملة أشمل بكثير لأطر المعاينة الرئيسية واستخداماتها.

### ٤ - ٢ - ١ تعريف العيّنة الرئيسية واستخداماتها

٥٧ - إطار (أو أطر) المعاينة في المرحلة الأولى من مراحل الاختيار في استقصاء للأسر المعيشية يجب أن يشمل جميع السكان المستهدفين. وعندما يُستخدم هذا الإطار لاستقصاءات متعددة أو حلقات متعددة من الاستقصاء نفسه، يعرف باسم إطار المعاينة الرئيسي، أو ببساطة العيّنة الرئيسية.

٥٨ - استخدام إطار العيّنة الرئيسية هو الاستراتيجية المفضلة لأي بلد يوجد لديه برنامج استقصاءات للأسر المعيشية مستمر وعلى نطاق واسع بين عمليات التعداد. وعلى العكس من ذلك، عندما لا يوجد برنامج استقصاءات مستمر، لا يوصى بالعيّنات الرئيسية بوجه العموم. وهناك وفورات ناتجة عن الحجم الكبير في استخدام وحدات الإطار نفسها على مر الزمن لأن جانباً كبيراً من تكلفة المعاينة تحدث في عمليات تطوير الإطار الرئيسي، لا في كل مرة يُجرى فيها استقصاء. أما البلدان التي تجري استقصاءً وطنياً في بعض الأحيان فقط بين عمليتي تعداد للسكان، من جهة أخرى، فلن تستفيد كثيراً من استغلال تصميم إطار العيّنة الرئيسية.

٥٩ - تشمل خصائص العينة الرئيسية عدد الوحدات في المرحلة الأولى من مراحل الاختيار وأحجامها وأنواعها. تتألف العينة الرئيسية، بوجه عام، من الاختيار الأولي لوحدات المعاينة الأولية التي تظل ثابتة لكل عينة فرعية. لاحظ أن المراحل اللاحقة تكون في العادة متغيرة. والمرحلة الأخيرة من مراحل الاختيار، مثلاً الأسر المعيشية الخاصة التي تُختار للمقابلة، تكون في العادة مختلفة في الاستقصاءات المستقلة، بينما يمكن أن تكون هي نفسها أو تتداخل جزئياً مع أسر أخرى في الاستقصاءات المتكررة.

#### ٤ - ٢ - ٢ الخصائص المثالية لوحدات المعاينة الأولية المستخدمة في إطار العينة الرئيسية

٦٠ - تظهر المبادئ التي تحكم إنشاء إطار العينة الرئيسية بضعة اختلافات عن مبادئ أطر المعاينة بوجه عام. فيجب أن يكون الإطار الرئيسي أكمل وأدق وأحدث ما يمكن عملياً. أما إطار العينة الرئيسية لاستقصاء الأسر المعيشية فهو مجرد إطار عينة عادية يطور في الأحوال النمطية من أحدث تعداد للسكان. ولأن الإطار الرئيسي يمكن أن يستخدم في كامل الفترة الفاصلة بين تعدادين، فهو يتطلب عادةً تحديثاً دورياً ومنتظماً، مرة كل سنتين أو ثلاث سنوات، مثلاً. وهذا يخالف الإطار العادي الذي يرجح أن يحدث على أساس مخصص، ولا يفعل ذلك إلا عندما يخطط لإجراء استقصاء معين.

٦١ - الخصائص التي تشجع على وضع إطار رئيسي هي — وهذا شيء متنبأ به — شبيهة بخصائص أطر المعاينة بوجه عام. ويكون تعريف الوحدات التي يُراد استخدامها باعتبارها وحدات المعاينة الأولية، مثلاً، مضغوطاً بمطلب أن تكون من قبل وحدات منطقة مرسومة خرائطها. غير أن هذه ليست مضابطة شديدة لأن وحدات الإطار ستعرف دون استثناء بأنها وحدات إدارية أنشئت من قبل لأغراض تعداد السكان. غير أن ثمة مطلباً هاماً يمكن أن يختلف عن مطالب أطر المعاينة العادية، وهو أن حجم وحدة المعاينة الأولية يجب أن يكون كبيراً إلى حد يكفي لاستيعاب استقصاءات متعددة دون ضرورة مقابلة نفس المستجيبين مقابلات متكررة؛ لكن حتى هذه الظاهرة يمكن التخلي عنها في تطبيقات معينة.

#### مثال

هناك نوع خاص من أنواع الإطار الرئيسي استخدم في بعض الأوضاع يقوم على أساس تصميم ذي مرحلتين. تنطوي المرحلة الأولى على عينة كبيرة من مناطق التعداد (أو وحدات مناطق صغيرة ورُسمت خرائطها على نحو مشابه). تُختار عينة فرعية من عينة وحدة التعداد الرئيسية لكل استقصاء مستقل يستخدم الإطار. وتوضع قائمة لكل عينة فرعية أو تقسم بطريقة أخرى إلى قطاعات فرعية لتطبيق الاستقصاء في الوقت الذي يُخطط فيه فعلياً للاستقصاء. ونقول، زيادة في الإيضاح: إن العينة الرئيسية يمكن أن تشمل ١٠ ٠٠٠ منطقة تعداد، منها ١ ٠٠٠ وحدة عبارة عن عينات فرعية لاستقصاء العمالة. وتوضع قائمة بالأسر المعيشية في مناطق التعداد الألف التي ستختار منها عينة للمرحلة الثانية قوامها ١٥ أسرة معيشية تُختار للاستقصاء من كل منطقة تعداد. في السنة التالية، تُختار من العينة الرئيسية عينة فرعية أخرى قوامها ٨٠٠ منطقة تعداد لاستخدامها في استقصاء صحي، وهلم جراً. بهذه الطريقة، لا تُستخدم أي منطقة تعداد أكثر من مرة، ومن هنا سيكون حجم وحدة المعاينة الأولية غير ذي شأن.

٦٢ - غير أن حجم وحدة المعاينة الأولية هام لأن جميع العينات الفرعية التي يولدها الإطار الرئيسي يجب أن تأتي من نفس مجموعة وحدات المعاينة الأولية؛ وفي المثال الوارد أعلاه، مناطق التعداد هي وحدات المعاينة الأولية

وتُستخدم مجموعة فرعية مختلفة من مناطق التعداد في كل عيّنة فرعية. ومما يذكر أن اختيار وحدات المعاينة الأولية في عيّنة رئيسية ليس مسألة خاصة لأن الطريقة ستكون نفس الطريقة سواء أكانت لعيّنة رئيسية أم لأية عيّنة أخرى. وستكون الطريقة بوجه عام طريقة الاحتمال المتناسب مع الحجم، إلا في بعض الحالات النادرة حيث تكون وحدات المعاينة الأولية متساوية في الحجم تقريباً؛ وفي تلك الحالة يمكن استخدام عيّنة تقوم على تساوي الاحتمالات مؤلفة من وحدات معاينة أولية.

#### ٤ - ٢ - ٣ استخدام العيّنات الرئيسية لدعم الاستقصاءات

٦٣ - ناقش الفرع ٣ - ٣ - ٧ كيف تلزم عيّنة كبيرة للعينات الرئيسية بغية توفير عدد كافٍ من الأسر المعيشية لدعم استقصاءات متعددة على مدى بضع سنوات دون حاجة إلى مقابلة نفس المستجيبين مراراً وتكراراً. وتشكّل أحجام العيّنات المتوقعة لجميع الاستقصاءات المتوقعة والمحتملة، التي يمكن أن تستخدم إطار العيّنة الرئيسية، المَعْلَمَات الرئيسية في تصميم إطارها. ويُتَوَقَّع، مثلاً، أنه تُقَابِل ٥٠ ٠٠٠ أسرة في مختلف الاستقصاءات التي ستُخَدَّمُها العيّنة الرئيسية، وستكون لدى فريق المعاينة المعلومات الأساسية التي يحتاج إليها لتقرير عدد وحجم وحدات المعاينة الأولية. علاوةً على ذلك، يمكن تطوير خطة لتنفيذ الاستقصاء من حيث استخدام العيّنة الرئيسية، كما يظهر في العيّنة التالية (انظر الشكل الوارد في الفرع ٣ - ٣ - ٧ للمقارنة).

#### مثال

كما جاء في المثال الوارد في الفرع ٣ - ٣ - ٧، تضمّ العيّنة الرئيسية في البلد ألف ٥٠ ٠٠٠ أسرة معيشية. وقد حُطِّط لثلاثة استقصاءات ستكون أحجام العيّنة والمجموعات فيها كما يلي: ١٦ ٠٠٠ أسرة معيشية، ٦ أسر من كل مجموعة لاستقصاء الإيرادات والنفقات؛ و ١٢ ٠٠٠ أسرة، و ١٢ أسرة من كل مجموعة لاستقصاء القوى العاملة؛ و ١٠ ٠٠٠ أسرة معيشية، و ٢٠ أسرة من كل مجموعة للاستقصاء الصحي. وقد اختيرت أحجام المجموعات المختلفة لكي تتصدّى للآثار المختلفة (بحسب نوع الاستقصاء) لآثار التصميم. بالإضافة إلى ذلك، توجد ١٢ ٠٠٠ أسرة معيشية إضافية تبقى احتياطاً لاستخدامها في استقصاءات أخرى إذا لزم. تتطلب الاستقصاءات الثلاثة المخطط لها ما مجموعه ١٦٧ ٤ وحدة معاينة أولية (١٦٠٠٠/٦ + ١٢٠٠٠/١٢ + ١٠٠٠٠/٢٠). ولما كان مضمون الاستقصاءات التي يمكن أن تستخدم العيّنة الفرعية الاحتياطية غير معلوم، تقرر التخطيط لحجم مجموعة مقداره ١٢ وحدة، مما يضيف ١ ٠٠٠ وحدة معاينة أولية، حتى يبلغ المجموع الكلي ١٦٧ ٥ وحدة. لذلك يقرر فريق تصميم العيّنة الرئيسية أن ينشئ عيّنة رئيسية مقدارها ٢٠٠ ٥ وحدة معاينة أولية. ويجب أن يأخذ تعريف وحدة المعاينة الأولية في الحساب عدد الأسر المعيشية المراد مقابلتها. في هذا الشكل، يجب أن تكون كل وحدة معاينة أولية كبيرة إلى حد أن ينتج عنها ٥٠ أسرة معيشية لمقابلتها. عندئذٍ يستطيع فريق المعاينة، بهذه المعلومات، أن يقرّر أيّ وحدة جغرافية تخدم على أفضل وجه لتعريف وحدات المعاينة الرئيسية. فإذا كان في البلد ألف، مثلاً، مناطق تعداد تضمّ الواحدة منها في المتوسط ١٠٠ أسرة معيشية، والتباين بين أحجامها والمتوسط لا يكاد يذكر، يكون من المعقول استخدام مناطق التعداد هذه وحدات معاينة رئيسية.

#### ٤ - ٢ - ٣ - ١ ميزات استخدام إطار العيّنة الرئيسية عدة مرات

٦٤ - توجد ميزات بارزة لاستخدام عيّنة رئيسية. الأولى والأهم هي أن خطة العيّنة الرئيسية تعمل بمثابة أداة تنسيق للوزارات الفنية وغيرها، التي لها مصلحة في أيّ برنامج إحصاء وطني. وهذا ينطبق على عدة جوانب لعملية

إجراء الاستقصاءات تتجاوز اعتبارات المعاينة في حد ذاتها، يتعلق معظمها بالتحكم بالتكاليف ووضع إجراءات موحدة في مختلف القطاعات فيما يتعلق بالتعاريف الإحصائية، وصياغة أسئلة الاستقصاءات وإجراءات تفسير البيانات.

٦٥ - توجد ميزة رئيسية لبرنامج العينات الرئيسية، وهي أنه يستخدم وحدات المعاينة الأولية نفسها. ويمكن تنظيم مجموعة موظفين ميدانيين وإبقاؤها طيلة دورة حياة العينة الرئيسية. فيمكن، مثلاً، استئجار المُقابلين وتدريبهم وإبقاؤهم متاحين في بداية برنامج العينات الرئيسية عندما يُعرف مكان وجود وحدات المعاينة الأولية لجميع الاستقصاءات التي سُنستعمل الإطار لمدة عشر سنين، مثلاً. ويمكن استئجار المُقابلين محلياً، بالعدد اللازم، من بين المقيمين في وحدات المعاينة الأولية للعينة الرئيسية أو بالقرب منها. ويمكن توليد مواد الاستقصاء، كخرائط وحدات المعاينة الأولية وصحائف قوائم الأسر المعيشية، عند بداية برنامج العينات الرئيسية، مما يوفر الوقت وسداد نسبة كبيرة من تكاليف عملية الاستقصاء لجميع الاستقصاءات المتوقعة. وتتيح الاستخدامات المتعددة للعينة، بالإضافة إلى ذلك، فرصة للتحكم بالأخطاء من غير أخطاء المعاينة وحتى عدم الاستجابة. ذلك لأن الزيارات المتكررة لنفس المستجيبين قد تمكن من تسجيل مواقف هؤلاء المستجيبين وملاحظة المشاكل التي تظهر في مجالات مختلفة، بغية وضع تدابير تصحيحية في الاستقصاءات اللاحقة. غير أن من المهم هنا أيضاً توكيد أن هذه الفوائد لا تحدث إلا عندما تُستخدم العينة الرئيسية استخداماً مكثفاً.

٦٦ - تشمل الميزات الأخرى للعينات الرئيسية، بوجه عام، سواءً أُستعملت وحدات المعاينة الأولية نفسها في تصميم ذي ثلاث مراحل أم وحدات معاينة أولية مختلفة في تصميم ذي مرحلتين، إمكانية (أ) دمج البيانات تحليلياً من تطبيقين اثنين أو أكثر للعينة الرئيسية باستخدام مضامين مختلفة، و(ب) الاستجابة بسرعة للاحتياجات غير المرئية إلى جمع البيانات.

#### ٤ - ٢ - ٣ - ٢ أوجه القصور في الاستخدامات المتعددة لإطار العينات الرئيسية

٦٧ - توجد في العينة الرئيسية بعض أوجه القصور الممكنة، كتلك التي تستنفد وحدات المعاينة الأولية، أي حين لا تبقى أسر معيشية إذا أفرط في استخدام الإطار الرئيسي. لكن يمكن تجنب ذلك بواسطة التخطيط الكافي مسبقاً. ومع أنه ليس من الممكن التنبؤ بكل استخدامات العينة الرئيسية الممكنة طيلة دورة حياتها، لكنّه يمكن تخصيص العينات الفرعية الاحتياطية لاستخدامات ممكنة ما دامت العينة الرئيسية كبيرة بما فيه الكفاية.

٦٨ - يوجد وجه قصور آخر، وهو مقدار التحيز المتزايد باستمرار، الذي يحدث حين لا يُنفذ التحديث الملائم كلما تقادمت العينة الرئيسية. وأخيراً، العينات الرئيسية غير مناسبة لتوفير البيانات في استقصاءات "الاحتياجات الخاصة"، كالتي تُجرى في أقاليم خاصة أو لمجموعات سكانية نادرة تكون موضع اهتمام.

#### ٤ - ٢ - ٤ المخصّصات للمجالات المختلفة (المناطق الإدارية، إلخ).

٦٩ - كانت المكاتب الإحصائية الوطنية تتعرض لضغط متزايد لجدولة وتحليل بياناتها المتعلقة باستقصاء الأسر المعيشية لمناطق إدارية هامة دون الوطنية كالمناطق والأقاليم الهامة والمدن الكبيرة. ويتوقع من بعض البلدان، مثل فييت نام، أن تقدم بيانات روتينية على مستوى المقاطعة. هذه المتطلبات والتوقعات مدفوعة باحتياجات سياسية مشروعة، لأسباب تقوم بوجه عام على أن البرامج الاجتماعية - الاقتصادية تركز على مناطق محلية وقد وضعت لهذه المناطق خلافاً للدولة بأسرها.

٧٠ - هذه، بلغة العيّنات الإحصائية، تقديرات للمجالات، كما سبقت مناقشتها من قبل. ونظراً إلى أن أحجام العيّنات الضرورية لتحقيق نتائج موثوقة أحجام هائلة، فهي تأتي بتكاليف باهظة غالباً ما تتجاوز ميزانيات الاستقصاء التي تستطيع الحكومة بوجه عام أن تدبرها. وتؤثر الحاجة إلى بيانات المجالات أيضاً في تطوير إطار العيّنة الرئيسية.

٧١ - نُظِرَ في عدد ونوع المجالات المراد إنشاؤها في الفرع ٣ - ٣ - ٤، المتعلق بأحجام العيّنات، ولن تتكرر المناقشة هنا. وحالما تُتَّخَذُ تلك القرارات يمكن إنشاء إطار العيّنة الرئيسية. فقد يقرّر بلد ما، مثلاً، أن الاستقصاءات التي ستجرى في إطار برنامج للعيّنات الرئيسية ستقدم بيانات لمجالين اثنين فقط — المجال الحضري والمجال الريفي. وقد يقرّر بلد آخر فيه ١٢ إقليمياً، ويود أن يقدم تقديرات لكل واحد منها، أن موارد استقصاءاته تستطيع دعم أحجام العيّنات الإضافية التي تلزم إذا عوملت الأقاليم معاملة المجالات. وثمة بلد ثالث فيه ٥٠ إقليمياً قد يقرّر أن إنتاج بيانات في كل واحد من هذه الأقاليم عملية باهظة التكلفة. وقد يقرّر، بدلاً من ذلك، تعريف ٨ مناطق جغرافية رئيسية بأنها مجالات وتقسّم عليها الـ ٥٠ إقليمياً. وقد يقرّر بلد رابع ألا ينشئ مجالات في حد ذاتها، وإنما يُجدول — بدلاً من ذلك — عيّنته الوطنية المخصّصة تناسبياً، بحسب المنطقة والإقليم والمنطقة الحضرية والمنطقة الريفية، ومدن كبرى مختارة، بقصد نشر بيانات على الجمهور متعلقة بتلك المناطق الفرعية، التي يعتبر حجم العيّنة فيها كبيراً إلى حد يكفي لإعطاء نتائج موثوقة إلى درجة معقولة.

٧٢ - لا يعتبر تخصيص الموارد للمجالات في الأمثلة المذكورة أعلاه ذا صلة، لأن العيّنة موزعة بالتناسب بين المناطق الوطنية الفرعية التي هي موضع اهتمام. ففي حالة الأمثلة الثلاثة الأولى، يجب اتخاذ خطوات خاصة لتخصيص وحدات المعاينة الأولية التابعة للعيّنة الرئيسية. وبما أن تقدير المجالات يفترض التساوي في الموثوقية بين المجموعات أو المناطق السكانية الفرعية، يجب اختيار نفس العدد من وحدات المعاينة الأولية من كل مجال — وهذا مطلب ثابت بغض النظر عما إذا كان تصميم العيّنة مستنداً إلى عيّنة رئيسية أم لا.

#### ٤ - ٢ - ٥ صيانة العيّنات الرئيسية وتحديثها

٧٣ - من حيث التأثير في تغطية السكان، تعتبر الصيانة الصحيحة لإطار العيّنة الرئيسية عنصراً أساسياً في تطويرها وفي التخطيط لاستخدامها. وتستخدم العيّنة الرئيسية للبلد — أي بلد — في الأحوال النمطية، للعقد الذي يفصل بين كل تعدادين، ويرجح أن يحدث بينهما تحوُّلات بعيدة المدى في تحركات السكان. من الضروري تحديث الإطار بصورة دورية ليعكس التغيرات السكانية لكي يظل "ممثلاً" للسكان.

٧٤ - يوجد نوعان هامان من أنواع التحديث. النوع الأول، وهو الأبسط، يستتبع إعداد قوائم جديدة للأسر المعيشية في مجموعات العيّنات المختارة في المرحلة قبل الأخيرة، وهذا إجراء يوصى به بوجه عام في كل أجزاء هذا الدليل، سواء أكان ذلك لعينات رئيسية أم لتصاميم عينات تُستخدم استخداماً منفرداً. وبذلك تُحدَّث مجموعات العيّنات تحديتاً تلقائياً لكي تعكس الهجرة والمواليد والوفيات. ومن شأن هذا النوع من التحديث، إذا انحصر في مجموعات العيّنات، أن يساعد على تقليل أخطاء التغطية (غير أخطاء المعاينة) إلى الحد الأدنى؛ لكن التباين في المعاينة يزداد على مر الزمن ما لم يتم تحديث الإطار بأسره.

٧٥ - توجد أيضاً حاجة إلى تحديث الإطار بأسره، بصورة دورية، لكي يؤخذ في الحسبان على الوجه الصحيح النمو الذي يحدث على نطاق كبير بعد التعداد. هذا النمو، كما نوقش من قبل، من النوع الذي يحدث في بناء المباني السكنية الشاهقة وتوسُّع مناطق المستوطنين في المدن. وإن مناطق التعداد، التي يحدث فيها هذا النمو المرتفع،

كانت أصغر كثيراً، بلا استثناء، في الوقت الذي أنشئ فيه إطار العينة الرئيسية. ونتيجة لذلك، تصبح مقاسات أحجامها أقل في تقديرها من الحقيقة إلى حد خطير بسبب تطور النمو. لذلك تقل فرص اختيارها باحتمال متناسب مع الحجم إلى الحد الأدنى. ويمكن أن يكون الأثر على تباين المعاينة جذرياً عندما يحدث أن تُختار مناطق التعداد هذه، لأن مقياس الحجم الحالي ربما يكون أضعاف المقياس الأصلي.

٧٦ - يمكن تخفيض مشاكل مناطق النمو المرتفع وآثارها الضارة بالعينات الرئيسية تخفيضاً كبيراً بتنقيح الإطار بانتظام، كل سنتين أو ثلاث سنوات، أو نحو ذلك.

#### ٤ - ٢ - ٦ تناوب وحدات المعاينة الأولية في العينات الرئيسية

٧٧ - يُحال القارئ إلى الفرع ٣ - ٩ - ٢ من الفصل الثالث، المعنون "المعاينة لتقدير التغير أو الاتجاه"، للاطلاع على مناقشة مفصلة فيما يتصل بالاستقصاءات المتكررة أو المستمرة، التي تهدف إلى قياس التغير أو الاتجاه. وتوحي العينات المتراكبة بوجود مشروع عينات يستخدم أسراً معيشية بديلة عندما تكرر الاستقصاءات. ومن الأهمية بمكان أن نؤكد من جديد أن استخدام العينات المتراكبة هو التقنية المفضلة لتقدير التغير من سنة إلى السنة التالية، مثلاً. ومن أساليب تبديل العينات التناوب الذي يمكن من حدوث تراكم جزئي من استقصاء إلى استقصاء أو من مناسبة إلى مناسبة.

٧٨ - أشير في الفرع السابق إلى أن موثوقية المعاينة (وهذا شيء مرغوب)، والأخطاء غير أخطاء العينات (وهذا شيء غير مرغوب) أكبر ما يكونان عندما تُستخدم نفس الأسر المعيشية في كل حلقة من حلقات الاستقصاء. ونتيجة لذلك، يسعى القائمون بالاستقصاء عادةً إلى حل وسط باستخدام تراكم جزئي في العينة من حلقة إلى الحلقة التي تليها، لا سيما عندما يُكرر الاستقصاء ثلاث مرات أو أكثر (انظر الفرع ٣ - ٩ - ٢ للاطلاع على الأساس المنطقي للتراكم الجزئي).

٧٩ - من أساليب إدخال التراكم الجزئي تبديل أو تناوب وحدات المعاينة الأولية في العينة (خلافًا لتبديل الأسر المعيشية في وحدات المعاينة الأولية في العينة). وعندما توجد عينة رئيسية مؤلفة من وحدات معاينة أولية، تُستخدم ليس فقط لحلقات نفس الاستقصاء، وإنما لاستقصاءات متعددة، لا يقل عن ذلك أهمية أن يُنظر بكل عناية في إجراء تناوب بينها.

٨٠ - لتنفيذ خطة التناوب تنفيذاً عملياً ولإعطاء نتائج ذات معنى ينبغي أن تكون درجة التراكم بين الفترات الزمنية مشابهة وثابتة طول الوقت. فإذا كان التراكم بين السنتين الأولى والثانية، مثلاً،  $K$  في المائة، ينبغي أيضاً أن تكون  $K$  في المائة بين السنتين الثانية والثالثة، وبين الثالثة والرابعة، وهلمّ جراً. وبناءً على ذلك، عندما يجري التناوب بين وحدات المعاينة الأولية كاملةً ينبغي إدراج هذه السمة في تصميم التناوب.

#### ٤ - ٢ - ٧ أمثلة قُطرية للعينات الرئيسية

٨١ - يعطي هذا الفرع توصيفات للعينات الرئيسية في أربعة بلدان نامية، هي: كمبوديا، والإمارات العربية المتحدة، وفيت نام، وموزامبيق. ويوضح كل واحد منها سمات ومبادئ المعاينة الرئيسية، كالمعاينة ذات المرحلة الواحدة أو المرحلتين من وحدات المعاينة الرئيسية والتطبيق المرن لاستقصاءات محددة تناقش في هذا الفصل. بالإضافة إلى ذلك، توضح سمات أخرى لتصميم العينة أبرزت في هذا الدليل. وتشمل هذه السمات، فيما تشمله، التقسيم الضمني إلى طبقات، والاختيار الأمثل لحجم المجموعة، بغية تخفيض آثار تصميم العينة وتخصيص العينات للمجالات.

## ٤ - ٦ - ١ - كمبوديا، ١٩٩٨ - ١٩٩٩

٨٢ - توضّح العيّنة الرئيسية لكمبوديا استخدام التصميم ذي المرحلتين. فثمة عيّنة كبيرة من وحدات المعاينة الأولى تقدّم قائمة رئيسية بقطاعات منطقة من المرحلة الثانية أُدرجت في عيّنات لاستخدامها في استقصاءات محددة.

٨٣ - وضع المعهد الوطني للإحصاءات في كمبوديا عيّنة رئيسية في سنة ١٩٩٩ لاستخدامها في البرنامج الحكومي لاستقصاءات الأسر المعيشية بين كل تعدادين للسكان، وهي تتألف من استقصاء اجتماعي - اقتصادي دوري ويحتمل كذلك أن تجرى استقصاءات للصحة، والقوى العاملة، والإيرادات والنفقات، والديمغرافيا، واستقصاءات مخصّصة. كان تعداد السكان لسنة ١٩٩٧ بمثابة إطار لتصميم العيّنة الرئيسية، التي نُفّذت في مرحلتين. كانت المرحلة الأولى اختيار وحدات المعاينة الأولى باحتمال متناسب مع الحجم، وعُرّفت هذه الوحدات بأنها قُررى كان مقياس حجمها عدد الأسر المعيشية كما جاء في التعداد. وأُجري اختيار وحدات المعاينة الأولى كعملية حاسوبية. وتناولت المرحلة الثانية إنشاء قطاعات منطقة داخل وحدات المعاينة الأولى، التي أُجريت كعملية يدوية.

٨٤ - تقرّر استخدام القرى كوحدات معاينة أولية لأنها كبيرة بما فيه الكفاية، في المتوسط لتضم عدداً كافياً من الأسر المعيشية (٢٤٥ أسرة في المناطق الحضرية و١٥٥ أسرة في المناطق الريفية)، لاستيعاب عدة استقصاءات فيما بين التعدادين. وهذا يجنّب الناس عبء المقابلات المتكررة مع نفس المستجيبين. وقد نُظر في البديل المتمثل في استخدام مناطق التعداد، لكنه رُفض لأن حجمها يعادل في المتوسط نصف حجم القرية فقط. ولم تُشمل المجموعات السكانية الخاصة، كالمجموعات العابرة أو المقيمة في مؤسسات، ولا سكان الثكنات العسكرية، في العيّنة الرئيسية.

٨٥ - اختير ما مجموعه ٦٠٠ وحدة معاينة أولية في العيّنة الرئيسية لأنه ساد شعور بأن هذا العدد يكفي لنشر الاستقصاء في جميع أنحاء البلد على نحو يمثل جميع الأقاليم تمثيلاً كافياً. واستُخدم في اختيار العيّنة التقسيم الضمني إلى طبقات، ونُفّذ باستعراض ملف القرية بتتابع جغرافي - المناطق الحضرية والريفية وبحسب الإقليم، والمقاطعة والكميونة. وهكذا وُزعت العيّنة الرئيسية توزيعاً تناسبياً بصورة تلقائية، بحسب المناطق الحضرية والريفية والأقاليم.

٨٦ - كانت في العيّنة الرئيسية سمة طريفة في كمبوديا، وهي عملية المعاينة في المرحلة الثانية. فقد استدعى ذلك، كما ذُكر أعلاه، إنشاء قطاعات منطقة داخل وحدات المعاينة الأولى المختارة. وينبغي الإشارة إلى أنه ينبغي عدم الخلط بين مفهوم المرحلة الثانية من بناء العيّنة الرئيسية ومفهوم المرحلة الثانية من اختيار العيّنة المتعلقة باختيار الأسر المعيشية لاستقصاء معيّن. فقد شكّلت في كل وحدة معاينة أولية في العيّنة الرئيسية المختارة قطاعات منطقة يضم واحداً ١٠ أسر (في المتوسط) وكان تشكيلها مهمة كتابية في عملية المكتب. وفي هذا السياق، لم يَسْتَدِرْ تشكيلها أيّ عمل ميداني إلا في الحالات غير الاعتيادية، لأن دفاتر قائمات التعداد لسنة ١٩٩٧ ومسودات الخرائط الموجودة استُعملت لهذه الغاية. وقد حُسب عدد القطاعات التي أنشئت في كل وحدة معاينة أولية في العيّنة الرئيسية بتقسيم عدد الأسر الوارد في تعداد السكان على ١٠ مع تقريب حاصل القسمة إلى أقرب عدد صحيح. فالقرية التي تضم ١٨٧ أسرة، مثلاً، بحسب تعداد سنة ١٩٩٧ قُسمت إلى ١٩ قطاعاً.

٨٧ - شكّلت القطاعات المنشأة على هذا النحو لبنات البناء لكي تُبنى بها العيّنات أو العيّنات الفرعية فيما يتعلق بتطبيقها على استقصاءات محددة. ويتم اختيار واحد من القطاعات أو أكثر من جميع وحدات المعاينة الأولى أو مجموعة فرعية منها، لكل استقصاء أو حلقة استقصاء تستخدم العيّنة الرئيسية. وتوجد ميزة هامة في ذلك،

وهي أن إنشاء العينة الرئيسية بالطريقة التي وُصفت تتيح فرصة لكل واحد من الاستقصاءات المحددة التي تستخدمها لكي يضع مُعاملات ترجيحه ذاتياً، بحسب تفاصيل تصميم عيّنته.

٨٨ - ثمة ميزة رئيسية في تصميم العينة الرئيسية، وهي أنها تسمح بكثير من المرونة في الطريقة التي تُقسم بها إلى عيّنات فرعية لاستخدامها في استقصاءات محدّدة. ومن شأن اختيار المجموعات (أي القطاعات، بعبارة أخرى) لكل استقصاء أن ينتج مجموعةً مختلفة إن أُريد ذلك. وتضمّ وحدة المعاينة الأولية في الأحوال النمطية ١٨ - ٣٠ قطاعاً، فتوفر عدداً من القطاعات في كل وحدة معاينة أولية يكفي للمشاركة في جميع الاستقصاءات. علاوةً على ذلك، يمكن تكرار الاستقصاء الاجتماعي - الاقتصادي على أساس سنوي مع مجموعة من القطاعات المختلفة كل سنة. وفي خيار آخر، يمكن أيضاً الاحتفاظ بعدد من القطاعات (٢٥ في المائة، مثلاً) من سنة إلى أخرى بنمط من أنماط التناوب، وتُبدل ٧٥ في المائة من القطاعات كل سنة.

٨٩ - استخدام المجموعات المترابطة (جميع الأسر المعيشية في قطاع العينة متجاورة) يفرض وجهاً من أوجه القصور على تصميم العينة الرئيسية. وهذا يزيد أثر التصميم عنه في القطاعات غير المترابطة، أي بمعنى عينة منهجية من الأسر المعيشية في مجموعة أكبر لأنه ساد الاعتقاد بأن أثر التصميم سيكون أقل إلى حد ما بتقليل حجم المجموعة، فاتفق في النهاية على قطاعات حجمها ١٠ أسر معيشية فقط.

٩٠ - كان من المتوقع تحديث العينة مرة كل سنتين أو ثلاث سنوات. مع أنه اعترف بأن من الأفضل تحديث العينة الرئيسية بأسرها، تقرّر تحديث وحدات المعاينة الأولية فقط للاستقصاء بالعينة، الذي كان مخطئاً له بالذات في ذلك الوقت. وقد تألف التحديث من زيارات ميدانية اضطلع بها لإعداد قوائم جديدة للأسر المعيشية في القطاعات المعنية. وقد أُعيد إدراج نفس مساحة الأرض في القطاعات التي تضمّ المجموعة الأولى من الأسر، وهذا من الأسباب التي تجعل حدود القطاعات على درجة كبيرة من الأهمية.

٩١ - من الطريف أن يلاحظ المرء أنه طُلب، فيما يتعلق بتصميم العينة الرئيسية الكمبودية، تعاونُ مختبر القرى في عمليات التحديث. فمن المعروف أنهم يحتفظون بسجلات دقيقة لكل الأسر في قراهم؛ علاوةً على ذلك، تبقى السجلات بصورة روتينية محتوية على أحدث القيود. ويُعتقد أن قائماتهم في معظم الحالات غاية في الدقة. سيضاف إلى ذلك أن مختبر القرى يشكّلون أيضاً موارد قيّمة من حيث تعيين قطع الأرض الملائمة لأيّ قطاع بعينه والتعريف بمواقعها.

#### ٤ - ٢ - ٦ - ٢ الإمارات العربية المتحدة، ١٩٩٩

٩٢ - تُبرز العينة الرئيسية للإمارات العربية المتحدة سمتين هامتين في التصميم. الأولى هي أن تصميم العينة الرئيسية يستخدم تقسيماً خاصاً إلى طبقات لأخذ المجموعتين السكانيّتين المختلفتين في الإمارات العربية المتحدة — المواطنين وغير المواطنين — في الاعتبار. والثانية هي أن التصميم يوضح كيف يمكن استغلال تصميم العينة المعيارية (انظر الفرع ٣ - ٨ - ٢) لمعالجة تباين أحجام عيّنات التعداد وعمليات التعداد القديمة.

٩٣ - وصفت وزارة التخطيط العينة الرئيسية للإمارات العربية المتحدة بأنها عينة فائقة تضم ٥٠٠ وحدة معاينة أولية، تستند إلى تعداد السكان لعام ١٩٩٥ باعتباره إطار المعاينة لها. ويقصد به أن يُستخدم لاستقصاءات محددة للأسر المعيشية حتى إجراء التعداد التالي للسكان. وعُرِّفت وحدات المعاينة الأولية بأنها مناطق تعداد السكان، أو أجزاء من هذه المناطق، لكي تشمل وحدة المعاينة الأولية، في المتوسط، ٦٠ أسرة معيشية — سواءً في ذلك المواطنون وغير المواطنين.

٩٤ - أنشئت طبقتان قبل اختيار وحدات المعاينة الأولية. تتكون الطبقة الأولى من مناطق تعداد كان تُثُكُّ الأسر التي فيها أو أكثر، في وقت التعداد، مواطنين. وتتألف الطبقة الثانية من جميع وحدات التعداد الأخرى. صُنِّفَ ما مجموعه ١٦٨٦ منطقة تعداد في الطبقة الأولى، و٢٩٨٦ في الطبقة الثانية. وباستخدام الاختيار المنهجي باحتمال متناسب مع الحجم، اختيرت عينة قوامها ٢٥٠ وحدة معاينة أولية في كل واحدة من الطبقتين، فبلغ المجموع ٥٠٠ وحدة معاينة أولية في الدولة بأسرها. وكان من المتوقع أن تنتج هذه العينة الرئيسية مجموعتين من الأسر الوطنية والأسر غير الوطنية متساويتين في العدد تقريباً. أولاً، قُسمت وحدات المعاينة الأولية الكبيرة (التي تضم ٩٠ أسرة معيشية أو أكثر) إلى قطاعات، واختير قطاع واحد من كل وحدة معاينة كبيرة بهذا الحجم اختياراً عشوائياً. وأنشئت قائمة جديدة محدّثة من الأسر المعيشية في وحدات المعاينة الأولية الـ ٥٠٠ لتحديث الإطار. وأسفرت عملية إعداد القوائم عن نحو ٣٠٠٠٠ أسرة معيشية في وحدات المعاينة الأولية التي تشكّل العينة، وسوف تُستخدم هذه الوحدات في توليفات مختلفة لإجراء استقصاءات معيّنة. ولتيسير التطبيق المرن، قُسمت الأسر في كل وحدة معاينة أولية من العينة إلى ١٢ مجموعة فرعية، أو فريق، يضم كل منها ٥ أسر معيشية.

٩٥ - توجد ظاهرة جديرة بالملاحظة في العينة الرئيسية، وهي أنها ليست ذاتية الترحيح، لأن الطبقتين غير متساويتين في الحجم. كان أول استقصاء يستخدم العينة الرئيسية هو الاستقصاء الوطني لمرض السكري، الذي أُجْرِيَ في سنة ١٩٩٩، برعاية وزارة الصحة. وكانت ثمة استقصاءات أخرى ستُجرى، وهي استقصاء القوى العاملة، واستقصاء الإيرادات والنفقات (أو استقصاء ميزانيات الأسر).

٩٦ - ترد أدناه بضع تفاصيل أخرى فيما يتعلق بظروف خاصة معيّنة في الإمارات العربية المتحدة تملّي كيفية بناء تصميم العينة الرئيسية. وهناك اعتباران غالبان يؤخذان بعناية في الحساب، وهما المجموعات السكانية المستهدفة وإطار العينة.

٩٧ - يوجد، كما لاحظنا، مجموعتان سكانيتان هامتان مستهدفتان في البلد — المواطنون وغير المواطنين. بينما تشكّل المجموعة الأولى ٤٣ في المائة من السكان، وفقاً لتعداد السكان لسنة ١٩٩٥، فهي تشكّل رُبع الأسر فقط في الدولة. ويعود السبب في ذلك إلى أن الأسر المعيشية لغير المواطنين أصغر حجماً بكثير من حيث عدد أفراد الأسرة المعيشية. وأثر ذلك على تصميم العينة هو أنه إذا اختيرت عينة تناسبية من أسر الدولة فإن نحو ثلاثة أرباع الأسر المستجيبة تكون من غير المواطنين. وثمة أثر آخر، وهو أن موثوقية تقديرات الاستقصاء الناتجة لأسر غير المواطنين ستكون أكبر من تقديرات أسر المواطنين بنحو ثلاثة أضعاف. وحيث إن النتائج ستستخدم لوضع برامج سياسية وتخطيطية، فإن هذا التفاوت في موثوقية التقديرات لا يُعتقد بأنه مستحسن أو مفيد. والحل من حيث تصميم العينة هو معاملة المجموعتين السكائيتين المختلفتين وغير المتساويتين ككيانين مستقلين، بواسطة تطبيق التقسيم الصحيح إلى طبقتين، كما ذُكر أعلاه.

٩٨ - استُخدم مستوى آخر للتقسيم إلى طبقات أيضاً، وهو التقسيم الجغرافي إلى طبقات، لضمان التوزيع الملائم للعينة بحسب الإمارة وبحسب المناطق الحضرية والريفية. فُرِزَ ملف مناطق التعداد بالتتابع التالي قبل اختيار العينة: أولاً، طبقة المواطنين بحسب رموز المناطق الحضرية وفي كل منطقة حضرية، وفي هذه الطبقة بحسب رموز الإمارة، وداخل الإمارة بحسب رموز منطقة التعداد، مرتبةً ترتيباً تصاعدياً بالنسبة المئوية للمواطنين، تتبعها رموز الطبقة الريفية، فرموز الإمارة، فرموز منطقة التعداد؛ وفُرِزَت بعدئذٍ طبقة غير المواطنين بنفس التتابع.

٩٩ - اعترِفَ بأنه يجب أن تكون ثمة سمة رئيسية في إطار العينة الرئيسية، وهي وجود مجموعة واضحة من الخرائط تعين حدود وحدات المناطق المراد تعيينها كمناطق مدرجة في العينة، بمعنى وحدات المعاينة الأولية. ويجب أن تكون وحدة المنطقة صغيرة بما فيه الكفاية لكي توضع في القائمة بصورة ملائمة لكنها في الوقت

نفسه كبيرة إلى حدٍ كافٍ لكي يتم تعريفها بوضوح فيما يتعلق بالحدود الطبيعية (لتيسير تعيين موقعها). وقد اعتُقد أن وحدات تعداد السكان هي وحدة المنطقة الممكنة عملياً التي تفي بهذه المعايير المزدوجة. من المؤسف أن الخرائط لم تُستخدم في تعداد السكان؛ ولذلك لم تكن مناطق التعداد محددة بوضوح من حيث الحدود المعروفة. وكان من الضروري، نتيجةً لذلك، ضمان وضع معلومات جيدة عن الحدود في وحدات المعاينة الأولية المدرجة في العينة الرئيسية (مناطق التعداد).

١٠٠ - استخدمت الأعمال التحضيرية للعينة الرئيسية لوحدات المعاينة الأولية "تصميم القطاع المعياري" الوارد وصفه في الفصل الثالث. هذه منهجية استُخدمت بنجاح في كثير من البلدان بواسطة برنامج الاستقصاءات الديمغرافية والصحية، وفي استقصاء الدول العربية لصحة الأم والطفل.

١٠١ - تقرّر استخدام تصميم القطاع المعياري لأن مناطق تعداد السكان في الإمارات العربية المتحدة متفاوتة جداً من حيث الحجم. أنشئت قطاعات معيارية يضم الواحد منها نحو ٦٠ أسرة معيشية. حُسب عدد القطاعات المعيارية في كل وحدة معاينة أولية مدرجة في العينة بأنه حاصل قسمة مجموع الأسر المعيشية (أي المواطنين وغير المواطنين) على ٦٠ مع تقريب حاصل القسمة إلى أقرب عدد صحيح.

١٠٢ - في الحالات التي كان فيها عدد القطاعات المعيارية - بمعنى مقياس الحجم - اثنين أو أكثر، قُسمت وحدة المعاينة الأولية إلى قطاعات منطقة. وقد اقتضى ذلك إجراءً ميدانياً يستدعي زيارة لمنطقة التعداد (وحدة المعاينة الأولية) وإعداد مسودة خارطة باستخدام عدّ سريع للمساكن (لا الأسر المعيشية) وتحديد مواقعها على الخارطة. بعد التقسيم إلى قطاعات اختير قطاع واحد من كل وحدة معاينة أولية اختياراً عشوائياً. وأصبح هذا القطاع هو المنطقة الجغرافية الفعلية للمعاينة في العينة الرئيسية. ونُظمت زيارة أخرى للميدان للحصول على القائمة الراهنة للأسر المعيشية في كل قطاع من قطاعات العينة. واعتُبر الإجراء الأخير، اللازم لتحديث العينة الرئيسية البالغ عمرها ثلاث سنوات، بالتالي عنصراً حيوياً في عملية المعاينة.

١٠٣ - تكونت العملية النهائية للعينة الرئيسية من تقسيم الأسر المعيشية التي وضعت بها قائمة جديدة في كل قطاع من قطاعات العينة إلى ١٢ مجموعة فرعية أو فريقاً. ويستخدم فريق واحد أو أكثر منها في استقصاءات معينة. ولما كان متوسط حجم القطاع نحو ٦٠ أسرة معيشية، فإن كل فريق يحتوي على خمس أسر في المتوسط.

١٠٤ - السبب الرئيسي لتشكيل ١٢ فريقاً هو المرونة التي يعطيها هذا العدد في تشكيل توليفات لاستخدامها في الاستقصاءات. ويتوقف الخيار الفعلي لكل استقصاء بعينه على عوامل مختلفة، من بينها أهداف الاستقصاء، وحجم المجموعة المستنوب، والحجم الكلي للعينة اللازم للاستقصاء. فقد استُخدم خمساً وحدات المعاينة الأولية، مثلاً، في الاستقصاء الوطني لمرض السكري. لذلك استخدمت أربعة أفرقة من الأفرقة الاثني عشر من الأسر المعيشية الموجودة في وحدات المعاينة الأولية هذه في الاستقصاء. وأسفرت هذه التوليفة عن خطة عينات إجمالية مقدارها ٢٠٠ وحدة معاينة أولية مع مجموعات حجمها ٢٠ (أي ٤ مضروبة بخمس أسر) وحجم عينة مجموعته ٤٠٠٠ أسرة معيشية تقريباً.

٤ - ٢ - ٦ - ٣ فيبيت نام، ٢٠٠١

١٠٥ - توجد سمتان مميزتان في العينة الرئيسية لفيبيت نام. فهي تُبرز استخدام المرحلتين في اختيار العينة الرئيسية ومرحلةً ثالثةً عندما تطبق على استقصاءات معينة. ثانياً، تُبرز كيف يمكن تخصيص عينة رئيسية لمجالات جغرافية.

١٠٦ - تتضمن العيّنة الرئيسية، التي تستند إلى تعداد سنة ١٩٩٩ باعتباره إطار المعاينة، تصميمًا ذا مرحلتين. وقد عُرِّفت وحدات المعاينة الأولى بأنها كميونات في المناطق الريفية وأحياء في المناطق الحضرية. وكان تعريفها بهذا الشكل لأنه تقرر أنه يلزم ما لا يقل عن ٣٠٠ أسرة معيشية في كل وحدة معاينة أولية لتكون عيّنة رئيسية. واعتُبرت مناطق التعداد، بدلاً من ذلك، وحدات معاينة أولية، لكنها كانت صغيرة جداً وكان يلزم جمعها مع مناطق تعداد مجاورة لكي تصبح مؤهلة بشكل مُرضٍ لأن تكون وحدات معاينة أولية. واعتُقد بأن المهمة الأخيرة ستكون مضيئة جداً ومستنزفة للوقت. ومن جهة أخرى كان عدد الكميونات/الأحياء التي لزم جمعها مع وحدات مجاورة لصغر حجمها مجرد ٥٢٩ كميونة/حياً من بين ما يزيد عن ١٠٠٠٠ كميونة/حي.

١٠٧ - اختير ما مجموعه ٣٠٠٠ وحدة معاينة أولية باحتمال متناسب مع الحجم للعيّنة الرئيسية. وتحتوي كل وحدة معاينة أولية أدرجت في العيّنة، في المتوسط، ٢٥ منطقة تعداد في المناطق الحضرية و١٤ منطقة تعداد في المناطق الريفية. وفي المرحلة الثانية من مراحل الاختيار، اختيرت ثلاث مناطق تعداد في كل وحدة معاينة أولية تشكل جزءاً من العيّنة مع احتمال متناسب مع الحجم. وكانت وحدات المرحلة الثانية، مناطق التعداد، تضم كل واحدة منها ١٠٠ أسرة معيشية، وفقاً لما جاء في تعداد سنة ١٩٩٩: ١٠٥ أسر في المناطق الحضرية و٩٩ أسرة في المناطق الريفية.

١٠٨ - كان من أهداف العيّنة الرئيسية القدرة على توفير بيانات موثوقة إلى حد ما لكل واحدة من مناطق فييت نام الجغرافية الثماني. واضطُبع باختيار العيّنات بصورة مستقلة داخل كل إقليم على حدة. فكانت هذه الأقاليم بمثابة طبقات للعيّنة الرئيسية. وسادت الرغبة في الإفراط في المعاينة في أقاليم صغيرة معيّنة، تضم مجموعات سكانية صغيرة. بناءً على ذلك، كان توزيع العيّنة بين الأقاليم بأسلوب الاحتمال المتناسب مع الجذر التربيعي لحجم الإقليم. واستخدم التوزيع النسبي بين المناطق الحضرية والمناطق الريفية.

١٠٩ - بالإضافة إلى التقسيم الطبقي على مستوى الإقليم، الذي ذُكر أعلاه، استُخدم تقسيم جغرافي ضمني إلى طبقات داخل الإقليم. وفي تطبيق العيّنة الرئيسية على استقصاءات محدّدة لزم استخدام مجموعات فرعية من مناطق التعداد — فاستخدم ثلثها، مثلاً، لاستقصاء الأسر المعيشية المتعدد الأغراض. ولتطبيق الاستقصاء، استُخدمت مرحلة ثالثة اختير فيها عدد محدد من الأسر المعيشية من كل منطقة تعداد أدرجت في العيّنة. وقد يختلف هذا العدد من استقصاء إلى آخر ومن منطقة حضرية إلى منطقة ريفية. فقد تُختار، مثلاً، ٢٠ أسرة من كل منطقة تعداد في مناطق التعداد الريفية و ١٠ أسرة من كل منطقة تعداد في مناطق التعداد الحضرية.

#### ٤ - ٦ - ٢ - ٤ موزامبيق، ١٩٩٨ - ١٩٩٩

١١٠ - تبين العيّنة الرئيسية لموزامبيق حالة يُستخدَم فيها اختيار وحدات معاينة أولية في مرحلة واحدة لجميع الاستقصاءات الوطنية التي تدعمها الحكومة في البرنامج الوطني لاستقصاءات الأسر المعيشية بين كل تعدادين للسكان. وتبين أيضاً كيف يمكن اعتماد عيّنة رئيسية مرنة لتحقيق أهداف القياس من استقصاء محدد.

١١١ - عُرِّفت وحدات المعاينة الأولى التي تشكّل جزءاً من العيّنة الرئيسية في موزامبيق وأدخلت في العيّنة بطريق مباشر، كما وُصِفَ في أجزاء مختلفة من هذا الدليل. وقد أنشئت وحدات المعاينة الأولى من تعداد السكان لسنة ١٩٩٧ باعتباره إطار المعاينة. وهي مشكّلة من تجمّعات جغرافية تضم الواحدة منها بوجه عام ما يتراوح من ٣ إلى ٧ مناطق تعداد للسكان، تضم الواحدة منها نحو ١٠٠ أسرة في المتوسط. واختيرت وحدات المعاينة الأولى المدرجة في العيّنة الرئيسية باحتمال متناسب مع الحجم.

١١٢ - اختير ما مجموعه ١٥١١ وحدة معاينة أولية لتشكّل إطاراً للمعاينة يُطبّق في نظام موزامبيق المتكامل لاستقصاءات الأسر المعيشية. وقُسمت وحدات المعاينة الأولية التي تشكّل أجزاءً في العينة الرئيسية على أفرقة، يشكّل كل فريق منها مجموعة فرعية منهجية، وهي لذلك عينة احتمالية في حد ذاتها. كانت هناك ١٠ أفرقة، يضم كل فريق منها نحو ١٥١ وحدة معاينة أولية. وفي خطة السنوات الخمس (٢٠٠٠ - ٢٠٠٤)، كان استبيان مؤشرات الرفاهية المركزية، الذي وضعه البنك الدولي، أول استقصاء يستخدم العينة الرئيسية.

١١٣ - صُممت خطة عينة استبيان مؤشرات الرفاهية المركزية سعياً إلى تحقيق هدفين اثنين للقياس. الأول هو الحصول على المؤشرات ذات الصلة التي تلزم لتوصيف رفاهية الأشخاص والأسر في موزامبيق. والثاني تقديم تقديرات موثوقة لهذه المؤشرات على الصعيد الوطني، للمناطق الحضرية والمناطق الريفية، كل على حدة، ولكل واحد من أقاليم البلد الأحد عشر. استخدمت منهجية المعاينة لاستبيان، مؤشرات الرفاهية المركزية، العينة الرئيسية لموزامبيق لاختيار نحو ١٤٥٠٠ أسرة في مجموعات مقسّمة طبقاً. وبناءً على ذلك، كانت المرحلة الأولى من مراحل الاختيار، طبعاً، وحدات المعاينة الأولية التي أُدرجت في العينة الرئيسية.

١١٤ - كانت المرحلة الثانية من مراحل الاختيار لاستبيان مؤشرات الرفاهية المركزية اختيار عينات فرعية من وحدات المعاينة الأولية التي تشكّل جزءاً من العينة المركزية. قُسم ما مجموعه ٦٧٥ وحدة معاينة أولية من الوحدات الـ ١٥١١، المدرجة في العينة المركزية، إلى عينات فرعية. واختيرت بطريقة منهجية بناءً على تساوي الاحتمالات، وبالإضافة إلى ذلك، اختيرت بالتساوي من أقاليم موزامبيق الأحد عشر. وفي المرحلة الثالثة، اختيرت عينة مؤلفة من منطقة تعداد واحدة من كل وحدة معاينة أولية لاستبيان مؤشرات الرفاهية المركزية. وهكذا كان في عينة الاستبيان ٦٧٥ مجموعة — منها ٤٧٥ مجموعة ريفية و ٢٠٠ مجموعة حضرية. اختيرت مناطق التعداد بناءً على تساوي الاحتمالات لأن أحجامها كانت متماثلة تقريباً — كما دُكر أعلاه، نحو ١٠٠ أسرة في المتوسط، مع أنه كان هناك شيء من التفاوت بينها. حدثت المرحلة الثالثة من مراحل الاختيار بعد العمل الميداني الذي قام فيه المعهد الوطني للإحصاء بزيارة المجموعات لتجميع قائمة جديدة بالأسر المعيشية، لتحديث إطار العينة لسنة ١٩٩٧. ومن القوائم التي جُمعت على هذا النحو، اختيرت عينة منهجية مؤلفة من ٢٠ أسرة معيشية في المناطق الريفية و ٢٥ أسرة في المناطق الحضرية لإجراء مقابلات الاستقصاء بناءً على أساس الاستبيان. لذلك كانت عملية اختيار العينات للاستبيان من أربع مراحل، مع أن العينة الرئيسية التي استندت إليها اختيرت في مرحلة واحدة.

١١٥ - هناك سمتان اثنتان في تصميم استبيان مؤشرات الرفاهية المركزية تبيّنان المرونة التي يمكن أن تُعتمد بها العينة الرئيسية لتتناسب مع المتطلبات المحددة لتطبيق الاستقصاء.

١١٦ - أولاً، باستخدام العينة الرئيسية لاستبيان مؤشرات الرفاهية المركزية، كان ثمة اهتمام لدى المعهد الوطني للإحصاء في استخدام الأفرقة التي كانت قد عُيّنت من قبل، كما دُكر أعلاه. ونظراً إلى الرغبة في استخدام نحو ٦٠٠ وحدة معاينة أولية للاستبيان، كان من المأمول أنه يمكن استخدام أربعة من الأفرقة. غير أن هذه الفكرة رُفضت عندما أدرك أن عدد الأفرقة التي يلزم استخدامها للاستبيان سيكون مختلفاً في كل إقليم عنه في الأقاليم الأخرى، لأن هدف القياس يتطلب تساوي أحجام العينات تقريباً في كل إقليم. واختيرت بدلاً من ذلك وحدات المعاينة الأولية اللازمة للاستبيان والبالغ عددها ٦٧٥ وحدة، بصورة منهجية من ملف العينة الرئيسية برمته دون اعتبار الأفرقة. وهذا العدول عن الهدف الأصلي للعينة الرئيسية المتمثل في توفير عينة تناسب كل إقليم حصل لأن موثوقية للاستبيان متكافئة مرغوب فيها على مستوى الأقاليم مقابل التصميم الأصلي للعينة الرئيسية. كان من المتوقع في خطة العينة الرئيسية، التي وُضعت أصلاً، أن تأخذ التقديرات الأولية على المستوى الوطني.

١١٧ - هناك مسألة هامة أخرى بشأن تصميم عيّنة الاستبيان، وهي مسألة حجم المجموعة. اتفق على أن أحجام المجموعات يجب أن تكون مختلفة من حيث عدد الأسر المعيشية الحضرية والريفية بحجة أن أثر تصميم العيّنة أعلى في المناطق الريفية حيث أسباب المعيشة الاقتصادية تُستمد من الزراعة الكافية. بعبارة أخرى، يرجح أن تكون الأسر المعيشية في المناطق الريفية ذات خصائص متشابهة جداً. وتتيح العيّنة الرئيسية إمكانية اختيار عدد محدد مختلف للأسر المعيشية (٢٥ في المناطق الحضرية و ٢٠ في المناطق الريفية) في المرحلة النهائية.

### ٤ - ٣ موجز المبادئ التوجيهية

١١٨ - يوجز هذا الفرع المبادئ التوجيهية الرئيسية المراد استخلاصها من الفصل. تقدم هذه المبادئ في شكل قائمة، وهي، كما جاء في الفصل الثالث، قواعد تقريبية أكثر منها توصيات محددة، وهي تقترح ما يلي:

- استخدام أطُر عينات تكون كاملة ودقيقة وحديثة قدر الإمكان.
- ضمان أن يشمل إطار العيّنة المجموعة السكانية المقصود استهدافها.
- استخدام أحدث تعداد سكان كإطار لاستقصاءات الأسر المعيشية إن أمكن ذلك.
- تعريف وحدات المعاينة الأولية في الإطار من حيث وحدات المناطق، مثل مناطق تعداد السكان بحدود مرسومة جيداً، وخرائطها مرسومة، وتكون أعداد سكانها معروفة.
- عدم استخدام قوائم الأسر المعيشية الموجودة في تعداد السكان إطاراً للمرحلة الأخيرة إلا إذا كان التعداد حديثاً جداً — لم يكن قد مضى على إجرائه أكثر من سنة واحدة في العادة.
- استخدام الأطُر الثنائية أو المتعددة بحذر وذلك بضمن كون الإجراءات موجودة للتعامل مع حالات التكرار.
- تحديث إطار التعداد إن كان قد مضى على إجرائه أكثر من سنتين — سواءً على الصعيد الوطني أم في مناطق مستهدفة على وجه التحديد، معروفة بأنها شهدت نمواً مرتفعاً، لكي يشمل:
- استخدام زيارات عدّ سريع لتحديث الإطار القديم.
- في المجموعات المدرجة في العيّنة، التحديث بوضع قوائم جديدة بالأسر المعيشية.
- التحديث في مجموعات العيّنات فقط إذا لم يكن قد مضى على التعداد أكثر من سنتين ليشمل:
- التحديث بوضع قوائم جديدة بالأسر المعيشية.
- عدم استخدام العيّنة الرئيسية أو إطار العيّنة الرئيسية إلا عندما يكون من المخطط تنفيذ برنامج استقصاءات مستمر أو كان هذا البرنامج جاري التنفيذ.
- تعريف وحدات معاينة أولية للعيّنة الرئيسية تكون كبيرة بما يكفي أو متعددة إلى حد يكفي لدعم استقصاءات كثيرة، أو تكرار حلقات استقصاء، في الفترة الفاصلة بين تعدادين.
- تحديث أطُر العيّنات الرئيسية باستخدام المبادئ التوجيهية كما اقترح أعلاه، لأطُر استقصاء منفرد.
- استخدام نظام لتناوب العيّنات — إما الأسر المعيشية وإما وحدات المعاينة الأولية — في الاستقصاءات المتكررة التي تستخدم عيّنات رئيسية.

## المراجع ومواد أخرى للقراءة

- Cochran, W.G. (1977), *Sampling Techniques*, 3rd ed., New York: Wiley.
- Hansen, M.H., W.N. Hurwitz, and W.G. Madow (1953). *Sample Survey Methods and Theory*, New York: Wiley.
- International Statistical Institute (1975). *Manual on Sample Design*. World Fertility Survey Basic Documentation, Voorburg, Netherlands.
- Kalton, G. (1983). *Introduction to Survey Sampling*, Beverly Hills, California: Sage. Publications Kish, L. (1965), *Survey Sampling*. New York: Wiley.
- League of Arab States (1990). *Sampling Manual, Arab Maternal and Child Health Survey*, Basic Documentation 5. Cairo: Pan Arab Project for Child Development (PAPCHILD).
- Macro International, Inc. (1996). *Sampling Manual*. DHS-III Basic Documentation, No. 6. Calverton, Maryland.
- Pettersson, Hans. (2001), Mission report: recommendations regarding design of master sample for household surveys of Viet Nam. Unpublished. General Statistical Office, Hanoi. 25 November
- \_\_\_\_\_ (2005). *Design of master sampling frames and master samples for household surveys in developing countries*. In Household Sample Surveys in Developing and Transition Countries. Studies in Methods, No. 96 Sales No. E.05.XVII.6.
- Turner, A. (1998), Mission report to the Kingdom of Cambodia, National Institute of Statistics, 11-24 November. National Institute of Statistics, Phnom Penh: Unpublished.
- \_\_\_\_\_ (1999), Mission report to United Arab Emirates, Ministry of Health and Central Department of Statistics 23 January–3 February. Abu Dhabi: Unpublished. Central Department of Statistics.
- \_\_\_\_\_ (2000). Mission report to Mozambique, Instituto Nacional de Estatística, 13-26 August. Instituto Nacional de Estatística, Maputo: Unpublished.
- United Nations Statistics Division (1984). *Handbook of Household Surveys* (Revised Edition). Studies in Methods, No. 31. Sales No. E.83.XVII.13.
- \_\_\_\_\_ (1986), National household Survey Capability Programme: *Sampling frames and sample designs for integrated household survey programmes*. Preliminary Version. DP/UN/INT-84-014/5E. New York: United Nations Department of Technical Co-operation for Development and Statistical Office.
- United Nations Children's Fund (2000). *End-Decade Multiple Indicator Survey Manual*. New York: UNICEF chap. 4.
- United States Bureau of the Census (1978). *Current Population Survey Design and Methodology*. Technical Paper 40. Washington, D.C.: Bureau of the Census.
- Verma, Vijay (1991). *Sampling Methods*. Training Handbook. Tokyo: Statistical Institute for Asia and the Pacific.
- World Bank (1999). *Core Welfare Indicators Questionnaire (CWIQ) Handbook*. Washington D.C.: World Bank. chap. 4.



## الفصل الخامس

# توثيق تصاميم العيّنات وتقييمها

### ٥ - ١ مقدمة

١ - لهذا الفصل، وإن كان قصيراً نسبياً، دور مركزي في هذا الدليل. فتوثيق تصاميم العيّنات وتقييمها بوجه خاص، ومنهجية الاستقصاء بوجه عام، كثيراً ما تُهمل بسبب الإسراع إلى إصدار نتائج الاستقصاء. ويصح هذا بوجه خاص في البلدان التي لا توجد لديها خبرة مسبقة تُذكر في إجراء استقصاءات للأسر المعيشية، حيث يكون توثيق البيانات الوصفية في الغالب توثيقاً رديئاً في صحائف عمل الاستقصاءات وتقاريرها. وفي بعض البلدان لا تُعطى قيم تُذكر لتوثيق إجراءات الاستقصاء، بما في ذلك تنفيذ العيّنات. لذلك، لا يوجد تقليد لتوفير الوثائق، ناهيك عن طلبها. ومع أن الباحثين يتعهدون بتحليل البيانات التي تولدها الاستقصاءات، ينبغي لهم أن يهتموا أيضاً بمعرفة المنهجيات المتصلة بتصميم العيّنات. لذلك، يجب إعداد عدد من الوثائق لوضع تفاصيل الإجراءات التي استخدمت في الاستقصاء.

٢ - غالباً ما يتجاهل الناس تقييم نتائج الاستقصاء تجاهلاً تاماً. ومن عواقب ذلك التجاهل تسرّب الأخطاء إلى تحليل الاستقصاء. ويعود هذا بوجه عام إلى حقيقة أن القيود المفروضة على الميزانية غالباً ما تحول دون تطوير أيّ دراسات نظامية أو أساليب لتقييم الأخطاء الكثيرة من غير أخطاء المعاينة، التي يمكن أن تظهر في استقصاءات الأسر المعيشية. ومع ذلك، توجد مقاييس أخرى لجودة البيانات، ينبغي أن تكون متوفرة بيسر (كنسبة عدم الاستجابة، مثلاً) وهذه أيضاً لا يرد لها ذكر في تقارير الاستقصاء في أغلب الأحيان.

٣ - يشدد هذا الفصل أيضاً على أهمية تقديم المعلومات ذات الصلة إلى المستخدمين عن أوجه القصور المعروفة في البيانات، حتى عندما لا تُجرى دراسات نظامية للتقييم؛ لكن من الأهمية بمكان، في هذا الصدد، ملاحظة أن مناقشة تقنيات إجراء تقييمات رسمية لمنهجية الاستقصاء، التي يوجد منها الكثير، تقع خارج نطاقه<sup>١</sup>. يركز الفصل، بدلاً من ذلك، على ماهية المعلومات التي ينبغي إعطاؤها إلى المستخدمين لمساعدتهم على تقييم جودة الاستقصاء، مع التركيز على جوانب المعاينة.

### ٥ - ٢ الحاجة إلى توثيق العيّنات وتقييمها وأنواع هذا التوثيق والتقييم

٤ - يوجد نوعان من الوثائق اللازمة في استقصاءات الأسر المعيشية. النوع الأول يتألف من سجلات معدة بعناية للاستقصاء وإجراءات المعاينة بينما هي تُنفذ عملياً في عملية الاستقصاء. ومن دون هذا التوثيق تتسلسل

<sup>١</sup> تشمل الدراسات الخاصة التي تصمّم لتقييم أنواع محددة من الأخطاء غير أخطاء المعاينة، بما في ذلك استقصاءات إعادة المقابلة (لتقييم تباين الردود)، واستقصاءات ما بعد التعداد (لتقييم التغطية والمضمون)، والتداخل في العيّنات (لتقييم تفاوت المقابلين، وإجراء فحوص عكسية للسجلات (لتقييم أخطاء المستجيبين بعد التذكيرة) وما أشبه ذلك.

الأخطاء إلى تحليل الاستقصاء. فاحتمالات الاختيار، مثلاً، ربما لا تكون معروفة جيداً في وقت التحليل دون حفظ السجلات بعناية.

٥ - لذلك، يجب أن يتخذ تقنيّ المعاينة الخطوات الضرورية لا لتوثيق خطة العيّنات للاستقصاء المحدد الجاري بعناية فقط، وإنما توثيق تنفيذه أيضاً. وتستدعي تصاميم العيّنات في الغالب تكييفاً في مراحل مختلفة من العمل الميداني بسبب الأوضاع غير المنظورة التي تبرز في إجراء الاستقصاء. ومن الأهمية بمكان أن تسجل — خطوة خطوة — جميع الإجراءات المستخدمة في تنفيذ خطة العيّنات لكي تضمن أن يتبع التنفيذ التصميم بأمانة. وعندما لا يكون الأمر كذلك، فإنّ ما هو أكبر أهمية تسجيل جميع حالات الخروج عن التصميم، حتى الصغيرة منها. هذه المعلومات ستكون ضرورية في المستقبل، في مرحلة التحليل، إذا ما دعت الحاجة إلى إجراء أيّ تعديلات.

٦ - النوع الثاني من التوثيق يتكون من التقارير. وينبغي إعداد نوعين من التقارير الفنية لكل استقصاء للأسر المعيشية: وصف موجز إلى حد ما وميسور للمستخدم لمنهجية الاستقصاء، بما في ذلك خطة العيّنات وتنفيذها؛ ووصف أكثر تفصيلاً لمنهجية الاستقصاء. فالوصف الأول يغطي في العادة الأجزاء "التقنية" (أو التذييلات) من التقارير الفنية التي تُنشر لمناقشة النتائج الفنية للاستقصاء<sup>٢</sup> وتفسيرها، بما في ذلك جزء فرعي عما يعرف عن أوجه قصور البيانات (انظر أدناه).

٧ - النوع الثاني من التقرير الفني يُقصدُ به أن يكون أكثر توجّهاً إلى الباحثين المهنيين وعلماء الاجتماع والإحصائيين منه إلى راسمي السياسة والجمهور العام، وينبغي أن يحتوي على وصف أكثر تفصيلاً لمنهجية الاستقصاء ويجب أن يكون مستقلاً بذاته، لا أن يكون جزءاً في سلسلة من التقارير الفنية. لقد أنتج مكتب الولايات المتحدة للإحصاء (١٩٧٨) نصاً ممتازاً من هذا التقرير. من الأفضل طبعاً أن تنتج التقرير الفني المفصل وتقارير الاستقصاء العادية في الوقت نفسه، وإن كان التقرير الأول في الأحوال النمطية يُنتج في وقت لاحق متأخر كثيراً، إن أنتج على الإطلاق. ومن المفيد أيضاً إصدار التقرير الفني، أو نص مختصر منه، في مجلة إحصائية لضمان بقائه طويلاً.

٨ - التقارير من كلا النوعين هامة جداً، لذلك يوصى بأن تعيّن مكاتب الإحصاء الوطنية مكتباً خاصاً أو موظفاً خاصاً لإعدادها بشكل روتيني لاستقصاءات الأسر المعيشية.

## ٥ - ٣ وضع بطاقات عناوين لمتغيرات التصميم

٩ - يناقش هذا الفرع والفروع من ٥ - ٢ حتى ٥ - ٧ توثيق النوع الأول، الذي يُعنى بحفظ سجلات عمليات الاستقصاء المتعلقة بالمعاينة.

١٠ - ينبغي وضع بطاقات تعريف واضحة وفريدة على وحدات الاختيار التي تُعيّن في كل مرحلة. هذا يعني، في التصميم متعدد المراحل، وضع رموز لوحدة المعاينة الأولية والثانية والثالثة والأخيرة (هذا يتوقف على عدد المراحل التي يشملها التصميم). ويكفي في العادة رمز من أربعة أرقام للمرحلة الأولى من مراحل الانتخاب، ورمز من ثلاثة أرقام للمراحل الباقية. وينبغي أيضاً وضع بطاقات تعريف صحيحة للمجالات الجغرافية. بالإضافة إلى

<sup>٢</sup> يمكن العثور على المبادئ التوجيهية المتعلقة بما ينبغي إدراجه في التقرير عن نتائج الاستقصاء بالعيّنة في تقرير للجنة الفرعية للأمم المتحدة المعنية بالمعاينة الإحصائية (الأمم المتحدة، ١٩٦٤).

ذلك، ينبغي أن تكون الرموز الإدارية التي تعرّف الهيكل الجغرافي والإداري للمناطق التي تنتمي إليها وحدات المعاينة جزءاً من عملية وضع بطاقات التعريف. وينبغي أيضاً تعريف وحدات التحليل بوضوح.

### مثال

افترض أن عيّنة تضم ١٢٠٠ وحدة معاينة عُرِّفت في مناطق تعداد السكان قد اختيرت لتصميم يضم مرحلتين، وأن ٦٠٠ وحدة في كل واحد من مجالين عُرِّفاً بأنهما مجال حضري ومجال ريفي. الطريقة الملائمة لترميز وحدات المعاينة الأولية هي من ٠٠٠١ إلى ١٢٠٠. علاوةً على ذلك، من المفيد أن نعيّن هذه الرموز بنفس الترتيب التتابعي المستخدم في اختيار وحدات المعاينة الأولية. وقد تلزم هذه السمة للتطبيق في حساب أوجه التباين في العيّنات. فإذا اختيرت وحدات المعاينة الأولية الريفية أولاً ستكون رموزها ٠٠٠١ إلى ٠٦٠٠، بينما تكون رموز الوحدات الحضرية من ٠٦٠١ إلى ١٢٠٠. نظام الترميز هذا له ميزتان، هما: أولاً، كل وحدة معاينة أولية لها رمز فريد يختلف عن رموز الوحدات الأخرى وتعريف فريد؛ ثانياً، يستطيع المحللون أن يعرفوا بمجرد نظرة إن كانت وحدة المعاينة الأولية حضرية أو ريفية من رمز التعريف. في المرحلة التالية من المعاينة يرد اسم كل وحدة معاينة أولية في القائمة، وتُختار منها ٢٠ أسرة معيشية للمقابلة. في هذه المرحلة، تُعطى جميع وحدات المعاينة الأولية الواردة في القائمة رموزاً من ثلاثة أرقام (أو أربعة أرقام إذا كانت بعض مناطق التعداد تضم أكثر من ٩٩٩ أسرة معيشية)، وهذه أيضاً تكون بنفس التسلسل الذي سُجلت به في القائمة. ونظراً إلى أن الأسر المعيشية المشمولة بالعيّنة ستحتفظ بالرموز التي أُعطيتُها بهذه الطريقة، لن تُعطى الأسر المختارة رموزاً جديدة، كـ ٠١ - ٢٠، مثلاً. وأخيراً، تُعطى الرموز الإدارية حسب الاقتضاء. وهكذا فإن الرمز ٠١٢٨ ٠٠٨ ٠٠٣ ٠٠٩ - ٠٨٠ عندما يُعطى لأسرة من أسر العيّنة فهو يعرفها بأنها الأسرة المعيشية الثمانين (٠٨٠) المدرجة في القائمة (والمختارة للمقابلة) في وحدة المعاينة الأولية رقم ١٢٨ (٠١٢٨) التي تنتمي إلى القسم المدني رقم ٨ (٠٠٨) في المقاطعة رقم ٣ (٠٠٣) من الإقليم رقم ٩ (٠٩). يُضاف إلى ذلك أن رقم وحدة المعاينة الأولية يُعرّف الأسرة على الفور بأنها في المجال الريفي. وإذا حصل الاستقصاء على معلومات عن أفراد الأسرة، يُعطى كل فرد منهم رمزاً فريداً مؤلفاً من رقمين، من ٠١ إلى ٩٩.

١١ - وضع بطاقات التعريف الصحيحة أساسي أولاً لمراقبة الجودة: حيث تُعطى المهام إلى المُقابلين وتُعاد الاستبيانات من الميدان، يمكن مقابلتها مع قائمة رئيسية لضمان أن جميع أسر العيّنة قد عُرِف مصيرها. ثانياً، نظم الترميز الفريد نُظِمَ قيّمةً لأنها تمكّن من وضع الجدولة بحسب الموقع الجغرافي.

١٢ - من المرغوب فيه جداً، في حالة البلدان التي يوجد لديها برامج استقصاءات متعددة، أن توضع بطاقات تعريف لمتغيرات التصميم بطريقة منسقة ومعيارية في جميع الاستقصاءات. لذلك يكون للقضاء على التشوش بين منتجي البيانات ومستخدميها ميزات واضحة في إعداد البيانات وتقديم النتائج.

١٣ - فيما يتعلق بالنقطة الأخيرة، سيستفيد برنامج الاستقصاءات المتعددة من إعطاء رموز لوحدة المعاينة الأولية لعالم وحدات المعاينة الأولية بأسره، كما ورد وصفه في المثال السابق، ليس فقط للوحدات المدرجة في العيّنة. هذا لازم لأن وحدات المعاينة الأولية المختلفة تُرمز في الغالب من قبل استقصاءات مختلفة، وهذا يحدث في كثير من الأحيان عندما تُستخدم العيّنات الرئيسية.

١٤ - تحتاج العيّنات الرئيسية بوجه عام إلى بطاقات تعريف للتصميم أكثر حتى مما تحتاج إليها عيّنات الاستقصاء الذي يجري مرة واحدة فقط. ومن الاستعمالات الرئيسية للعيّنات الرئيسية، كما نوقشت من قبل، الحلقات المتكررة من الاستقصاء نفسه. ووضع بطاقات التعريف الصحيحة على متغيرات التصميم التي تعرّف

مراحل الاختيار ضروري جداً، بغية متابعة الحالات التي تشكّل الأجزاء المترابكة من العيّنة من استقصاء إلى الذي يليه. وفي كثير من الأحيان تُعَيّن أفرقة التناوب ( المجموعات الفرعية المنهجية من العيّنة الكاملة) بقصد تسهيل تعريف الوحدات ( الأسر المعيشية، أو المجموعات أو وحدات المعاينة الأولية) المراد تبديلها في الحلقات التالية من الاستقصاء. وهي تحتاج طبعاً إلى رموز التعريف الخاصة بها. يضاف إلى ذلك أن الأسر المعيشية التي تضاف إلى العيّنة أثناء عمليات التحديث الدورية يجب ترميزها على الوجه الصحيح. وينبغي أيضاً تعيين نظام الترميز على هذا النحو للتمييز بين الأسر المعيشية الجديدة والقديمة.

## ٥ - ٤ احتمالات الاختيار

١٥ - ثمة بند واحد من بنود المعلومات يُغفل عنه في كثير من الأحيان في توثيق العيّنات، وهو تسجيل احتمالات الاختيار في المراحل المختلفة. فحيثما وجدت المعلومات غالباً ما تنحصر في مُعاملٍ ترجيح العيّنة الإجمالية (الذي يمكن منه حساب الاحتمال الإجمالي بسهولة) لكل حالة عيّنة.

١٦ - يوجد شيء هام من التوثيق الصحيح يجب أن يؤخذ في الاعتبار عند إجراء تقسيم العيّنة إلى عيّنات فرعية في الميدان أثناء جمع البيانات، كما يحدث عندما يكون قطاع/مجموعة العيّنة كبيراً جداً. فكما نوقش أعلاه، قد يلزم أن تقسم المجموعة التي يتبين أنها أكبر ممّا كان متوقّعاً إلى أربعة أجزاء، مثلاً، متساوية تقريباً في الحجم. وعندئذٍ يُختار جزء منها بصورة عشوائية لوضعه في القائمة ومقابلته. في تلك الحالة يكون الاحتمال الإجمالي للجزء المدرج في العيّنة (والأسرة المعيشية/الأفراد المختارين منه) يساوي رُبع احتمال المجموعة الأصلية؛ ومُعاملٍ ترجيحه عكس الرُبع، أو أربعة. ينبغي أن ينعكس مُعامل الترحيح هذا في الحسابات عند تحليل البيانات.

١٧ - يمكن أن يحدث تقسيم العيّنة إلى عيّنات فرعية أيضاً عندما يوجد أكثر من أسرة معيشية واحدة في المسكن (حين يكون المسكن هو الوحدة المدرجة في القائمة). أحد الخيارات، وهو غير متحيز، أن يقابل مدير الاستقصاء جميع الأسر المعيشية التي يجدها في المسكن؛ وغالباً ما يكون هذا هو النهج المتبع عندما توجد أُسرتان فقط في المسكن. غير أنه يمكن أن توجد خمس أسر، مثلاً، بينما كان من المتوقّع وجود أسرة واحدة. ربما تملّي اعتبارات التكلفة اختيار أسرة واحدة منها فقط — بطريقة عشوائية طبعاً — لمقابلتها. وهنا أيضاً يجب تسجيل نسبة تقسيم العيّنة (١/٥ في هذا المثال) لكي يتسنى حساب احتمال اختيار الأسرة المتأثرة بذلك حساباً دقيقاً من قبل موظفي المعاينة وتعديل مُعامل الترحيح تعديلاً صحيحاً (بمُعامل مقداره ٥).

١٨ - من المفيد أيضاً تسجيل احتمالات الاختيار في كل مرحلة من مراحل الاختيار، كما ذُكر في الفقرة الأولى من هذا الفرع. فاحتمال اختيار كل وحدة معاينة أولية، مثلاً، يكون مختلفاً كلما استخدمت المعاينة بعيّنة احتمالها متناسب مع الحجم. هذا صحيح حتى عندما يكون تصميم العيّنة الإجمالية ذاتي الترحيح. فإذا حدث تجاهل احتمالات اختيار وحدات المعاينة الأولية أو سُجّلت تسجيلاً خاطئاً ربما لا يتسنى تحديد مُعاملات ترجيح المعاينة الإجمالية. ومن المفيد، مثلاً، أن نعرف ما هي احتمالات الاختيار الأصلية لكي نحدد إجراءات تقسيم العيّنة إلى عيّنات فرعية تحديداً دقيقاً.

## ٥ - ٥ نسب الاستجابة ونسب التغطية في المراحل المختلفة من مراحل اختيار العيّنات

١٩ - من الأساسي، كجزء من عملية تقييم تنفيذ الاستقصاء بالعيّنة، تقديم معلومات إلى المستخدمين عن نسب الاستجابة ونسب التغطية. ومن المفيد إعطاء أكثر ما هو متاح من التفاصيل. ومن المهم عدم إعطاء نسبة

الاستجابة فقط (أو عنصر منها، هو نسبة عدم الاستجابة)، وإنما إعطاء جدولة أسباب عدم الاستجابة أيضاً. من المرجح أن تشمل أسباب عدم الاستجابة ما يلي:

- لا يوجد أحد في البيت
- وحدة السكن خالية
- وحدة السكن مهدومة أو غير صالحة للسكن
- رفض
- خارج البيت مؤقتاً (عطلة، إلخ).

٢٠ - ربما يختلف تعريف نسبة الاستجابة، من حيث الفئات المشمولة، من بلد إلى بلد. غير أن الاستجابة الكاملة، في الأحوال النمطية، تشمل الفئات الأولى والرابعة والخامسة من الفئات المذكورة أعلاه. وتشكّل هذه الفئات حالات يجب الحصول فيها على إجابة إن أمكن ذلك. وحدة السكن الخالية أو المهدومة يتم تجاهلها عادةً (في حساب نسبة الاستجابة) بحجة أنه لا يمكن بحكم التعريف الحصول على إجابة من هذه الوحدات. وبذلك يمكن، مثلاً، أن يختار البلد ٥٠٠٠ أسرة معيشية ويحصل على النتائج التالية: ٧٧٢ ٤ أتمت المقابلات، ٧٥ حالة "لا يوجد أحد في البيت"، ٣١ وحدة خالية، ١٧ وحدة مهدومة، ١٢ حالة رفضوا الإجابة، ٩٣ حالة "خارج البيت مؤقتاً". تُحسب نسبة الاستجابة، عادةً باستثناء الوحدات الخالية والمهدومة، بأنها  $٧٧٢ / (٥٠٠٠ - ٣١ - ١٧) = ٩٦,٤$  في المائة.

٢١ - عندما تشمل المجموعة السكانية المستهدفة للاستقصاء أسرةً معيشيةً — لمتغيرات مثل الإيرادات أو الحصول على الخدمات — وأفراداً (لمتغيرات كالحالة الصحية للنساء البالغات، مثلاً)، جرت العادة بحساب نسب الاستجابة على مستوى الأسر المعيشية ومستوى الأفراد كليهما. فيمكن، مثلاً، أن تستجيب ٩٨ في المائة من الأسر، بينما توجد داخل الأسر المستجيبة نسبة صغيرة من الأفراد الذين لا يستجيبون.

٢٢ - في أحيان كثيرة لا تُقَابَلُ مجموعات كاملة لأسباب مختلفة، من بينها مسائل أمنية، كأن يكون هناك حرب أهلية أو اضطرابات، أو عدم إمكانية الوصول إليها بسبب وعورة التضاريس أو الطقس. ويحدث مراراً، عندما تظهر مشاكل من هذا النوع، أن يُستعاضَ عن هذه المجموعات بمجموعات أخرى، وهذا إجراء ينطوي على درجة خطيرة من التحيز، لأن من المرجح أن يكون سكان المجموعات البديلة في كل الأحيان تقريباً مختلفين اختلافات كبيرة جداً عن سكان المجموعات التي حُلّت محلها. مع ذلك، يجب إذا حدث مثل هذا التبديل على فريق الاستقصاء أن يسجّل عدد هذه المجموعات ومواقعها. ويمكن أن يتم ذلك بتقدير عدد الأشخاص في المجموعة (أو المجموعات) السكانية المستهدفة، التي يُعتقد أنها مقيمة في مناطق المجموعات المثلة المستعاض عنها.

٢٣ - من المفيد ملاحظة أن المشاكل التي هي من النوع المذكور أعلاه مباشرة يمكن تقليلها إلى حد ما بالقيام قبل اختيار العيّنة، بتعيين مناطق البلد التي تكون "خارج النطاق" لأغراض المقابلة للاستقصاء، بسبب مخاوف أمنية أو مخاوف من عدم إمكانية الوصول إليها. ويجب توثيق هذه المناطق واستبعادها من عالم الاستقصاء قبل اختيار العيّنات. ويجب توثيق المناطق المستثناة في التقرير وإدراج بيان يقول إن نتائج الاستقصاء لا تنطبق على المنطقة المستثناة.

## ٥ - ٦ الترجيح: مُعامِلات الترجيح الأساسية، وعدم الاستجابة، وتعديلات أخرى

٢٤ - يرد حساب مُعامِلات الترجيح في الاستقصاء في الفصل السادس. ويشدد هذا الفصل على أهمية توثيق هذه الحسابات.

٢٥ - ينطوي الترجيح في استقصاءات الأسر المعيشية بوجه عام على ما يصل إلى ثلاث عمليات — حساب مُعامِلات الترجيح الأساسية أو التصميمية، والتعديلات المدخلة بسبب عدم الاستجابة، والتعديلات المدخلة فيما بعد التقسيم الطبقي. في كثير من التطبيقات، لا يُستخدَم إلا مُعامِلات الترجيح في التصميم، بينما في تطبيقات أخرى يمكن تعديل مُعامِلات الترجيح عند التصميم بعامل إضافي يعكس عدم الاستجابة. وفي تطبيقات قليلة نسبياً يمكن أن يعكس الترجيح عاملاً آخر، إما بالتعديل الذي يعكس عدم الاستجابة وإما بدونه، يُقصد به تعديل توزيع السكان الذي حُصلَ عليه من العيّنة، لجعله يتفق مع التوزيع الذي حُصلَ عليه من مصدر مستقل للبيانات، كتعداد حديث للسكان، مثلاً. هذا ما يشار إليه غالباً بعبارة الترجيح الحاصل بعد التقسيم الطبقي. وفي بعض التطبيقات لا يحصل ترجيح على الإطلاق؛ وهذا لا يحدث إلا في حالة توفّر شرطين اثنين فقط، وهما: أن تكون العيّنة ذاتية الترجيح تماماً وأن تكون البيانات المولدة محصورةً بالتوزيع بالنسب المئوية، وبالحصص والنسب، خلافاً لتقديرات المجاميع أو الأرقام المطلقة.

٢٦ - عندما يُستخدَم الترجيح، من الضروري طبعاً أن تُسجّل الحسابات بعناية. وكما ذُكر آنفاً، ينبغي أن تُحسب مُعامِلات الترجيح (أو الاحتمالات) في كل مرحلة وتُسجّل. وينبغي أيضاً تسجيل مُعامِلات الترجيح المستقلة في كل مرحلة من مراحل عمليات البيانات، وهي بعبارة أخرى: (أ) مُعامِلات الترجيح عند التصميم، و(ب) مُعامِلات الترجيح عند التصميم بعد ضربها بعامل (أو عوامل) التعديل مقابل عدم الاستجابة، و(ج) مُعامِلات الترجيح عند التصميم في (ب) بعد تطبيق عوامل التعديل لما بعد التقسيم الطبقي.

٢٧ - من المهم ان نلاحظ أن مُعامِلات ترجيح التصميم تختلف لكل مجال عندما يشتمل تصميم العيّنة على تقدير للمجال. بعبارة أخرى، حتى عندما تكون العيّنة ذاتية الترجيح داخل المجال، سيكون لكل مجال مُعامِل ترجيح خاص به. علاوة على ذلك يوجد لكل مجال مجموعة مختلفة من مُعامِلات الترجيح إذا لم يكن التصميم ذاتي الترجيح داخل المجالات. يضاف إلى ذلك أنه يجب ملاحظة أن تعديلات عدم الاستجابة تطبق غالباً منفصلة بحسب مناطق فرعية جغرافية هامة، كالمناطق الرئيسية، مثلاً، بغضّ النظر عما إذا كان تقدير المجال موجوداً في التصميم. وأخيراً، يمكن ضرب مُعامِل الترجيح نفسه بعامل إضافي لمجموعات أو أسر معيشية معيّنة. ويُفعل هذا عندما يُستخدَم تقسيم العيّنة إلى عيّنات فرعية (انظر الفرع ٥ - ٣).

## ٥ - ٧ معلومات عن تكاليف المعاينة وتنفيذ الاستقصاء

٢٨ - بينما توضع ميزانيات استقصاءات الأسر المعيشية بكل عناية، مما لا يقل عن ذلك أهمية الاحتفاظ بسجلاتّ النفقات الفعلية على مختلف عملياتها. فتسجيل تكاليف الاستقصاء مفيد، مثلاً، لتخطيط عيّنات رئيسية، وحفظ سجلّاتّ مختلف أنشطة الاستقصاء هاماً لتخطيط الاستقصاءات التي تُجرى في المستقبل.

٢٩ - عندما تُستخدَم العيّنات الرئيسية توجد تكلفة ابتدائية كبيرة لتفعيل تطويرها. وهي تشمل بوجه عام جوانب (أ) استخدام الحاسوب في التعامل مع ملفات تعداد السكان لإنشاء إطار العيّنة، و(ب) رسم الخرائط

والرسوم لإنشاء وحدات معاينة أولية، و(ج) اختيار وحدات المعاينة الأولية للعيّنة بالحاسوب. غالباً ما تشارك في دفع تكاليف العمليات الابتدائية، كما ذُكر في الفصل السابق، الوزارات التي ستستخدم العيّنة الرئيسية أثناء دورة حياتها. وينبغي أيضاً توزيع هذه التكاليف على جميع الاستقصاءات التي يُقصد استخدام العيّنة الرئيسية لها، بقدر ما تكون هذه التكاليف معروفة مسبقاً. لذلك، من الأساسي إنشاء السجلات بعناية فائقة فيما يتعلق بتطوير العيّنة الرئيسية وكذلك تخطيط جوانب المعاينة للاستقصاءات التي ستُجرى في المستقبل، والاحتفاظ بها.

٣٠ - عندما يُنتهى من وضع العيّنة الرئيسية، ينبغي وضع سجلات للتكاليف المتعلقة بالاحتفاظ بها. وكما لوحظ من قبل، يجري تحديث العيّنات الرئيسية بصورة دورية، وتحتاج تكاليفها طبعاً إلى المراقبة بكل عناية.

٣١ - تشمل عمليات المعاينة، التي ينبغي الحصول على تكاليفها بانتظام، التكاليف الواردة في القائمة التالية، التي تنطبق على الاستقصاءات بالعيّنة التي تجرى مرة واحدة وعلى العيّنات الرئيسية أيضاً:

(أ) مرتبات مصممي العيّنة بما في ذلك الأتعاب التي تُدفع للخبراء الاستشاريين الخارجيين؛

(ب) التكاليف الميدانية لتحديث إطار العيّنات، بما في ذلك تكاليف الموظفين وإعداد مواد مساعدة كالخرائط، مثلاً؛

(ج) التكاليف الحاسوبية لإعداد إطار العيّنات لاختيار العيّنة من وحدات المعاينة الأولية؛

(د) تكاليف الموظفين لاختيار العيّنة من وحدات المعاينة الأولية (إذا لم تُنفذ بالحاسوب)؛

(هـ) التكاليف الميدانية لإجراء عملية وضع قوائم وحدات المعاينة في المرحلة قبل الأخيرة، بما في ذلك تكاليف الموظفين وإعداد مواد، كملفات المجموعات؛

(و) تكاليف الموظفين لجمع البيانات من عيّنات الأسر المعيشية داخل مجموعات العيّنات.

٣٢ - بالإضافة إلى تكاليف العيّنات للاستقصاء، ينبغي حفظ السجلات المتصلة بتكاليف التنفيذ. ويمكن أن تشمل هذه التكاليف: أجور العدّادين، والمشرفين؛ وعلاوة الإعاشة اليومية والخدمات الميدانية للموظفين العاديين لمنظمة الاستقصاء؛ ونفقات السفر؛ واللوازم المكتبية؛ ونفقات التدريب؛ وتكاليف الوقود؛ وخدمات الاتصالات، وإعداد البيانات.

## ٥ - ٨ التقييم: أوجه القصور في بيانات الاستقصاء

٣٣ - كثير من الوثائق المتعلقة بحفظ السجلات على الوجه الصحيح، التي نوقشت أعلاه، بالإضافة إلى كونها هامة لتجهيز نتائج الاستقصاء، ووثائق مفيدة لتقييم جوانب تصميم العيّنات وتنفيذ الاستقصاء. فالمعلومات عن نسب الاستجابة، مثلاً، تساعد على تقييم ما إذا كان التحيز الآتي من عدم الاستجابة خطيراً أم لا. ويمكن استخدام المعلومات عن تكاليف المعاينة لتقييم الفعالية "الاقتصادية" لتصميم العيّنات واستخدامه للاستقصاءات في المستقبل.

٣٤ - كما ذُكر سابقاً، يشمل التقييم الرسمي للاستقصاءات بالعيّنة أوجهاً متعددة لأخطاء غير أخطاء المعاينة التي تتجاوز نطاق هذا الدليل إلى حد كبير (انظر الأمم المتحدة (١٩٨٤) — لمعالجة الموضوع معالجة شاملة). غير أنه ينبغي أن نذكر أن تقييم الخطأ غير أخطاء المعاينة ينبغي أن يشمل، مثلاً، أنشطة في مجالي التنفيذ والتجهيز. ومن جهة أخرى، يمكن تقدير خطأ المعاينة وتجري مناقشته بمزيد من التفصيل أدناه.

٣٥ - على الرغم من أن دراسات التقييم النظامية لا تُجرى في الغالب لاستقصاءات الأسر المعيشية، يظل مع ذلك من الأهمية بمكان أن يشمل توثيق الاستقصاء معلومات عن أوجه القصور في البيانات. وينبغي أن يُكرّس لهذا الموضوع جزء موجز من التقارير الفنية عن النتائج، غالباً ما يكون بعنوان بسيط هو "أوجه القصور في بيانات الاستقصاء". وينبغي إبلاغ القارئ في هذا الجزء عن جوانب المعاينة وغير المعاينة في أخطاء الاستقصاء.

٣٦ - يصف منشور قيّم صادر عن مكتب الولايات المتحدة للإحصاء (١٩٧٤) كيفية عرض المعلومات عن أخطاء الاستقصاء. وتقرّح الفقرات التالية المقتبسة من ذلك المنشور (التذييل الأول، الصفحة أولاً - ١) نوع المعلومات التي ينبغي عرضها للمستخدمين، عندما تُنشر نتائج الاستقصاء:

الإحصاءات الواردة في هذا التقرير هي تقديرات مستمدة من استقصاء بالعيّنة. يوجد نوعان من الأخطاء الممكنة في التقدير بناءً على الاستقصاء بالعيّنة — أخطاء معاينة وأخطاء غير أخطاء المعاينة. تحدث أخطاء المعاينة بسبب ملاحظات تُبدى على العيّنة فقط، لا على السكان ككل. والأخطاء التي هي من غير أخطاء المعاينة (التي تناقش في الفصل الثامن) يمكن أن تُعزى إلى مصادر كثيرة: العجز عن الحصول على معلومات عن جميع الحالات المشمولة بالعيّنة، وصعوبات متعلقة بالتعريف، واختلافات في تفسير الأسئلة، وعدم القدرة أو عدم الاستعداد من جانب المستجيبين لتقديم معلومات صحيحة، أخطاء في تسجيل أو ترميز البيانات التي حُصلَ عليها، وأخطاء أخرى متعلقة بجمع البيانات والاستجابة والتجهيز والتغطية وتقدير البيانات غير الموجودة. وتحدث أخطاء غير أخطاء المعاينة أيضاً في عمليات التعداد الكاملة. فدقّة نتيجة الاستقصاء تُحدّد بالآثار المشتركة لأخطاء المعاينة وأخطاء غير أخطاء المعاينة.

العيّنة الخاصة التي استخدمت في هذا الاستقصاء واحدة من عدد كبير من جميع العيّنات الممكنة ذات الحجم الواحد التي كان يمكن اختيارها باستخدام تصميم العيّنة نفسه. وتختلف التقديرات المستمدة من عيّنات مختلفة بعضها عن بعض. وخروج تقدير العيّنة عن متوسط جميع العيّنات الممكنة يُسمّى خطأ معاينة. أما الخطأ المعياري في تقدير الاستقصاء فهو مقدار التفاوت بين التقديرات الآتية من العيّنات الممكنة، وهو لذلك مقياس للدقة التي يكون بها التقدير الآتي من عيّنة محددة قريباً من متوسط نتائج جميع العيّنات الممكنة. ويُعرّف الخطأ المعياري النسبي بأنه الخطأ المعياري مقسوماً على القيمة الجارية تقديرها.

يقيس الخطأ المعياري، كما حُسب في هذا التقرير، قياساً جزئياً أيضاً أثر الأخطاء غير أخطاء المعاينة، لكنه لا يقيس أيّ تحيّزات منهجية في البيانات. والتحيز هو الفرق، كمتوسط لجميع العيّنات الممكنة، بين التقدير والقيمة المستتوية. ومن البديهي أن دقة نتيجة الاستقصاء تتوقف على أخطاء المعاينة والأخطاء غير أخطاء المعاينة كليهما، مقيسة بالخطأ المعياري، والتحيز والأنواع الأخرى من الأخطاء غير أخطاء المعاينة، غير المقيسة بالخطأ المعياري.

٣٧ - ثمة عنصر هام في تقييم العيّنة، كما ورد أعلاه ضمناً، وهو تقدير أخطاء المعاينة، الذي ينبغي إجراؤه بتقديرات الاستقصاء الرئيسية. وإن واحدة من الخصائص المميزة للعيّنة الاحتمالية، كما نوقش من قبل، هي أن العيّنة نفسها يمكن أن تُستخدم لتقدير الأخطاء المعاينة. فأساليب تقدير التفاوت والأخطاء المعاينة تُبحث بالتفصيل في الفصل السادس. بالإضافة إلى ذلك توجد برامج حاسوبية ذات كفاءة وموثوقية متاحة لتقدير الأخطاء المعاينة، ينبغي أن يُستفاد منها ما أمكن ذلك.

٣٨ - تُعدُّ تقديرات الأخطاء المعيارية للخصائص الرئيسية التي هي موضع اهتمام في الاستقصاء، بوجه عام، لأنه ليس عملياً ولا ضرورياً أن تُحسب لجميع البنود. فالأخطاء المعيارية تعطي الوسائل طبعاً إلى المستخدمين لتقييم موثوقية تقديرات الاستقصاء وإنشاء فترات ثقة حول تقديرات النقاط.

٣٩ - يمكن أيضاً استخدام الأخطاء المعيارية لتقييم تصميم العيّنة نفسه. وتوجد إحصائية مفيدة بوجه خاص للقيام بذلك، وهي أثر تصميم العيّنة ( $deff$ )، أو أكثر دقة ( $deft$ )، وهو الجذر التربيعي لأثر تصميم العيّنة. من السهل حساب الجذر التربيعي لأثر تصميم العيّنة، حساباً مباشراً إلى حد ما، لكل بند من بنود البيانات يُقدر الخطأ المعياري المتعلق به. وهذا يستتبع مجرد تقسيم الخطأ المعياري المقدر لبند معين على الخطأ المعياري الآتي من عيّنة بسيطة مختارة عشوائياً لها نفس حجم العيّنة، وذلك  $pq/n$ ، حيث  $p$  هي النسبة المقدرة؛ و  $q$  هي  $1-p$  و  $n$  هو حجم العيّنة. وتعمل العملية على توكيد أو دحض آثار التصميم التي افترضت افتراضاً عند تصميم العيّنة، لأن الآثار الفعلية لتصميم العيّنة والجذر التربيعي لآثار تصميم العيّنة لا يمكن معرفتهما حتى ما بعد إجراء الاستقصاء، وإعداد البيانات وتقدير الأخطاء المعيارية.

٤٠ - يستطيع خبير المعاينة الإحصائي أن يستخدم آثار التصميم المحسوبة لتقرير ما إذا كانت أحجام المجموعات أحجاماً معقولة لبنود البيانات الرئيسية، ويمكنه أن يتخذ تدابير تصحيحية عند الاقتضاء. فإذا كان الجذر التربيعي لأثر تصميم العيّنة، مثلاً، أكبر كثيراً من الحجم المتوقع لبنود رئيسية معيّنة، يمكن تصميم العيّنة لاستقصاء يُجرى في المستقبل على نحو تُستخدم فيه أحجام أصغر للمجموعات.

## ٥ - ٩ موجز المبادئ التوجيهية

٤١ - يوجز هذا الفرع في شكل قائمة، المبادئ التوجيهية الرئيسية الواردة في هذا الفصل، التي تُعرض، كما جاء في الفصول السابقة، في شكل قواعد تقريبية أكثر منها توصيات ثابتة و:

- توثيق جوانب المعاينة في الاستقصاءات بطريقتين: حفظ السجلات على الوجه الصحيح وتقديم المعلومات التقنية إلى المستخدمين
- حفظ سجلات مفصلة لعمليات المعاينة، بما في ذلك التكاليف
- وضع رموز لمتغيرات تصميم العيّنات: المناطق الإدارية، وحدات المعاينة الأولية، المجموعات، الأسر المعيشية، الأشخاص، إلخ.
- العمل بجهد على وضع رموز موحدة لمتغيرات التصميم تكون متسقة في كل الاستقصاءات
- تسجيل جميع حالات الخروج عن خطة العيّنات الأصلية الأخرى التي تحدث في تنفيذ الاستقصاء
- حساب وتسجيل احتمالات الاختيار في كل مرحلة من مراحل المعاينة
- تسجيل معلومات، على وجه الخصوص، عن تقسيم العيّنات إلى عيّنات فرعية، الذي يحدث في العمل الميداني
- تسجيل المعلومات عن عدد حالات عدم الاستجابة وأنواعها
- تسجيل مُعاملات ترجيح التصميم، والتعديلات بسبب عدم الاستجابة، والتعديلات المدخلة بعد التقسيم الطبقي

- الاحتفاظ بسجلات تكاليف مفصلة لكل عملية في تصميم العيّنات وتنفيذها
- في حالة العيّنات الرئيسية، الاحتفاظ بسجلات تكاليف تطويرها (التكاليف الابتدائية) وصيانتها
- إعداد تقارير تقنية للمستخدمين عن المعاينة ومنهجية الاستقصاء
- إعداد تقرير موجز عن أوجه القصور في جميع المنشورات الفنية التي تقدّم تقارير عن نتائج الاستقصاء
- إعداد تقرير فني أكثر كثافة عن جميع جوانب منهجية المعاينة
- حساب أخطاء المعاينة في المتغيرات الرئيسية وعرض هذه الأخطاء في تقارير تقنية
- حساب آثار التصميم (*deff or deft*) للمتغيرات الرئيسية
- تعيين موظف يتولّى مسؤولية التوثيق

### المراجع ومواد أخرى للقراءة

- Casley, D.J., and D.A. Lury, (1981). *Data Collection in Developing Countries*. Oxford, United Kingdom: Clarendon Press.
- International Statistical Institute (1975). *Manual on Sample Design*. World Fertility Survey Basic Documentation, Voorburg, Netherlands.
- League of Arab States (1990). *Sampling Manual, Arab Maternal and Child Health Survey*, Basic Documentation 5. Cairo: Pan Arab Project for Child Development (PAPCHILD).
- Macro International, Inc. (1996), *Sampling Manual*. DHS-III Basic Documentation No. 6. Calverton, Maryland.
- United Nations (1964). Recommendations for the Preparation of Sample Survey Reports (Provisional Issue). Statistical Papers, Series C, No. 1, Rev.2.
- \_\_\_\_\_ (1984), *Handbook of Household Surveys*, (Revised Edition). Studies in Methods, No. 31. Sales No. E.83.XVII.13.
- United States Bureau of the Census (1974). Standards for Discussion and Presentation of Errors in Data. Technical Paper 32. Washington, D.C.: United States Bureau of the Census.
- \_\_\_\_\_ (1978). *Current Population Survey Design and Methodology*. Technical Paper 40. Washington, D.C.: Bureau of the Census.
- World Bank (1999). *Core Welfare Indicators Questionnaire (CWIQ) Handbook*. Washington, D.C.: World Bank.

## الفصل السادس

# بناء مُعامِلات ترجيح العيّنات واستخدامها

### ٦ - ١ مقدمة

١ - يناقش هذا الفصل مختلف المراحل في وضع مُعامِلات ترجيح العيّنات واستخدامها في حساب تقديرات الخصائص التي هي موضع اهتمام من بيانات استقصاء الأسر المعيشية. ويوصّف، على وجه الخصوص، تعديل مُعامِلات الترجيح في العيّنات للتعويض عن مختلف النواقص الحاصلة في العيّنة المختارة. يقتصر الانتباه على تقديرات وصفية تُنتج على نطاق واسع في معظم تقارير الاستقصاءات. تُوضّح الأفكار الهامة المعروضة باستخدام أمثلة حقيقية من استقصاءات راهنة أُجريت في بلدان نامية، أو نماذج محاكاة لأوضاع استقصاءات حقيقية.

### ٦ - ٢ الحاجة إلى مُعامِلات ترجيح العيّنات

٢ - تستند استقصاءات الأسر المعيشية بوجه عام إلى تصاميم عيّنات مركّبة، بالدرجة الأولى لضبط التكاليف. ويرجح أن يكون في العيّنات الناتجة نواقص ربما تؤدي إلى تحيُّز ومفارقات أخرى بين العيّنة والمجموعة السكانية المشار إليها. ومن هذه النواقص اختيار عيّنات غير متساوية الاحتمالات، وعدم تغطية المجموعات السكانية، وعدم الاستجابة. مُعامِلات ترجيح العيّنات لازمة لتصحيح هذه النواقص، وبذلك تستمد تقديرات ملائمة للخصائص التي هي موضع اهتمام. مقاصد الترجيح، باختصار، هي:

(أ) التعويض عن احتمالات الاختيار غير المتساوية؛

(ب) التعويض عن عدم استجابة (الوحدات)؛

(ج) تعديل توزيع العيّنات المرّجّح للمتغيرات الرئيسية التي هي موضع اهتمام (السن، مثلاً، والعرق، ونوع الجنس)، لجعلها تتفق مع توزيع سكاني معلوم.

٣ - ستناقش الإجراءات المستخدمة لكل واحد من هذه السيناريوهات بالتفصيل في الفروع التالية. حالما يُعوّض عن النواقص في العيّنة، يمكن استخدام مُعامِلات الترجيح في تقدير خصائص السكان التي هي موضع اهتمام وكذلك في تقدير أخطاء المعاينة في تقديرات الاستقصاء المولّدة.

٤ - عندما لا تُستخدم مُعاملات الترجيح للتعويض عن نسب الاختيار المتباينة داخل الطبقات (حينما تكون العيّنة قد صُمّمت على هذا النحو) ولنواقص العيّنة المذكورة أعلاه، ستكون التقديرات الناتجة لِمُعَلَمَات السكان متحيزة. انظر الفروع ٦ - ٣ و ٦ - ٤ و ٦ - ٥، للاطلاع على أمثلة لإجراءات الترجيح المستخدمة في إطار كل سيناريو، بما في ذلك تعويض عن التقديرات المرّجحة وغير المرّجحة في كل حالة.

## ٦ - ٢ - ١ نظرة عامة

٥ - يتناول الفرع ٦ - ٣ تطوير مُعاملات ترجيح العيّنات، بما في ذلك تعديل مُعاملات ترجيح العيّنة لتفسير عمليات التكرار في العيّنة والوحدات التي لا تُعرف أهليتها للاستخدام في الاستقصاء في وقت اختيار العيّنة. ويناقش الفرع ٦ - ٤ ترجيح احتمالات الاختيار غير المتساوية؛ ويقدم بضعة أمثلة للتعداد، بما في ذلك دراسة حالة إفرادية لوضع مُعاملات الترجيح لاستقصاء وطني للأسر المعيشية؛ وفي الختام يناقش عيّنات ذاتية الترجيح. تُناقش مسألتنا عدم الاستجابة وعدم التغطية في استقصاءات الأسر المعيشية في الفرعين ٦ - ٥ و ٦ - ٦، على التوالي. وتُناقش مصادر وعواقب عدم الاستجابة وعدم التغطية. وتُعرض أيضاً طرق تعويض عن عدم الاستجابة وعدم التغطية، بما في ذلك أمثلة رقمية توضح تعديل مُعاملات ترجيح العيّنة مقابل عدم الاستجابة وعدم التغطية. يناقش الفرع ٦ - ٧ مسألة التضخم المالي في تباين تقديرات الاستقصاء، نتيجةً لاستخدام مُعاملات ترجيح العيّنة في تحليل بيانات استقصاء الأسر المعيشية. ويقدم أيضاً مثال رقمي لكي يوضّح حساب الزيادة في التباين الحاصلة نتيجةً للترجيح. ويناقش الفرع ٦ - ٨ مسألة تخفيف مُعاملات الترجيح ويعرض مثلاً لإجراء التخفيف الذي يعاد بموجبه قياس مُعاملات الترجيح المخففة بطريقة يجعلها تضيف إلى مجموع مُعاملات الترجيح الأصلية. وتُقدم بعض الملاحظات الختامية في الفرع ٦ - ٩.

## ٦ - ٣ تطوير مُعاملات ترجيح العيّنات

٦ - حالما تتقرر احتمالات اختيار الوحدات المدرجة في العيّنة يبدأ بناء مُعاملات ترجيح العيّنات. ويتوقف احتمال اختيار الوحدة المدرجة في العيّنة على تصميم العيّنة المستخدم لاختيار الوحدة. لقد عرض الفصل الثالث أوصافاً تفصيلية لتصاميم المعاينة الأكثر استعمالاً، واحتمالات الاختيار المتصلة بهذه التصاميم. ويُفترض في جميع الحالات أن تكون احتمالات الاختيار قد تقرر.

٧ - يُعتبر وضع مُعاملات ترجيح العيّنات أحياناً الخطوة الأولى في تحليل بيانات الاستطلاع. وهو يبدأ عادةً ببناء مُعامل الترجيح الأساسي أو مُعامل ترجيح التصميم لكل وحدة مدرجة في العيّنة لكي تعكس احتمالات اختيارها غير المتساوية. فمُعامل الترجيح الأساسي للوحدة المدرجة في العيّنة هو معكوس احتمال اختيارها لإدراجها في العيّنة. وفي الترميز الرياضي، إذا أُدرجت وحدة في العيّنة وكان احتمالها  $p_i$  فإن مُعامل ترجيحها الأساسي، المشار إليه بالرمز  $w_i$  يُعطى بـ

$$(٦ - ١)$$

$$w_i = 1/p_i$$

٨ - فإذا اختيرت وحدة مدرجة في العيّنة، مثلاً، باحتمال نسبته  $١/٥٠$  تمثل  $٥٠$  وحدة في السكان الذين اختيرت منهم العيّنة. وبذلك تعمل مُعاملات الترجيح في العيّنة بمثابة عوامل تضخم مصممة لتمثيل عدد الوحدات في المجموعة السكانية المشمولة بالاستقصاء، التي تمثلها الوحدة المدرجة في العيّنة التي أُعطيت مُعامل الترجيح. ويعطي مجموع مُعاملات الترجيح في العيّنة تقديراً غير متحيز لمجموع عدد الوحدات في المجموعة السكانية المستهدفة.

٩ - في حالة التصاميم المتعددة المراحل، يجب أن تعكس مُعاملات الترجيح الأساسية احتمالات الاختيار في كل مرحلة. ففي حالة تصميم ذي مرحلتين، مثلاً، تُختار فيه وحدة المعاينة الأولى رقم  $i^{th}$  مع احتمال نسبته  $p_i$  في المرحلة الأولى، تُختار الأسرة المعيشية رقم  $j^{th}$  من وحدة معاينة أولية مدرجة في العيّنة مع احتمال نسبته  $p_{j(i)}$  في المرحلة الثانية، فعندئذٍ يُعطى الاحتمال الإجمالي  $(p_{ij})$  لاختيار كل أسرة في العيّنة بحاصل جمع الاحتمالين، أو

(٢ - ٦)

$$p_{ij} = p_i * p_{j(i)}$$

ويُحصّل على مُعامل الترجيح الأساسي للأسرة المعيشية كما حصل من قبل، بأخذ معكوس الاحتمال الإجمالي لاختيارها. وبالتناسب مع ذلك، إذا كان مُعامل الترجيح الأساسي للأسرة المعيشية رقم  $b^{th}$  يساوي  $w_{ij,b}$ ، يكون مُعامل الترجيح المعزوّ إلى التعويض عن عدم الاستجابة  $w_{ij,nr}$ ، ومُعامل الترجيح المعزوّ إلى التعويض عن عدم التغطية  $w_{ij,nc}$ ، فعندئذٍ يُعطى مُعامل الترجيح الإجمالي للأسرة المعيشية بـ

(٣ - ٦)

$$w_{ij} = w_{ij,b} * w_{ij,nr} * w_{ij,nc}$$

### ٦ - ٣ - ١ تعديلات مُعاملات ترجيح العيّنة للأهلية غير المعلومة

١٠ - توجد، أثناء جمع البيانات في استقصاءات الأسر المعيشية، في بعض الأحيان حالات تكون فيها أهلية الأسرة المعيشية موضع شك. فربما لا يجد المقابل أحداً في البيت، مثلاً، في منزل أُدرج في العيّنة أثناء جمع البيانات أو بعد زيارات متكررة. في هذه الحالة لا يكون معروفاً إن كانت الوحدة السكنية مشغولة أم لا. فإن كانت مشغولة يجب تصنيفها بأنها وحدة سكنية غير مستجيبة (تحت فئة "لا أحد في البيت"). وإلا فهي خارج نطاق الاستقصاء، وهي لذلك غير مؤهلة لعدّها كوحدة مدرجة في العيّنة. وفي بعض الأحيان يفترض المقابلون أنه إذا لم يوجد أحد في البيت أثناء زيارات متكررة أن تلك الوحدة السكنية غير مشغولة ولذلك فهي غير مؤهلة. هذا افتراض غير صحيح بوجه العموم، ويؤدي في كثير من الأحيان إلى رفع نسبة الاستجابة خطأً.

١١ - عندما تكون أهلية بعض الوحدات السكنية غير معروفة، يجب تعديل مُعاملات ترجيحها لكي تعكس هذه الحقيقة. والفكرة هي أن تُفترض بعض الافتراضات التي تسمح بتقدير نسبة الوحدات السكنية التي تكون أهليتها غير معروفة وهي في الواقع مؤهلة. وأبسط نهج لتحقيق ذلك هو أن تؤخذ نسبة الوحدات المدرجة في العيّنة والمعروف أنها مؤهلة أو غير مؤهلة، وتطبق تلك النسبة على الوحدات غير المعروفة أهليتها. افترض، مثلاً، أن عيّنة من ٣٠٠ وحدة سكنية كانت ردودها كما هو مبين في الجدول ٦ - ١.

الجدول ٦ - ١

#### فئات الاستجابة في الاستقصاء

عدد الوحدات السكنية	فئة الاستجابة
٢١٥	مقابلات كاملة
٢٥	وحدات مؤهلة غير مستجيبة
١٠	وحدات غير مؤهلة
٥٠	وحدات غير مؤهلة غير معروفة

١٢ - لاحظ أن نسبة الوحدات السكنية المعروف أنها مؤهلة التي هي فعلاً مؤهلة هي  $(215+25)/ (10+25+215) = 0,96$ . لذلك يمكننا أن نفترض أن نفس النسبة  $(0,96)$  من الوحدات السكنية غير المعروفة أهليتها يمكن اعتبارها مؤهلة. بعبارة أخرى، ٩٦ في المائة من الـ ٥٠ وحدة سكنية غير المعروفة أهليتها (أو ٤٨ وحدة سكنية) في الواقع مؤهلة. عندئذٍ نعدّل مُعاملات ترجيح الوحدات السكنية المؤهلة (المقابلات الكاملة والوحدات المؤهلة غير المستجيبة) باستخدام عامل التعديل المعرف كما يلي:

(٤ - ٦)

$$F_{ue} = \frac{\sum_c w_{ij,b} + \sum_{nr} w_{ij,b} + \varepsilon \times \sum_{ue} w_{ij,b}}{\sum_c w_{ij,b} + \sum_{nr} w_{ij,b}}$$

حيث يشير  $\varepsilon$  إلى نسبة الحالات غير المعروفة أهليتها التي يقدّر أنها مؤهلة ( $\varepsilon = 0,96$  في هذا المثال). وتشير المجاميع التي هي فوق  $c$  و  $nr$  و  $ue$  في المعادلة الواردة أعلاه، على التوالي، إلى مجموع مُعاملات الترجيح الأساسية للوحدات السكنية التي أُجريت معها مقابلات كاملة، والوحدات المؤهلة غير المستجيبة، والوحدات غير المعروفة أهليتها. ثم تُحسب مُعاملات الترجيح الأساسية المعدلة للوحدات السكنية التي أُجريت معها مقابلات كاملة والوحدات المؤهلة غير المستجيبة بضرب مُعاملات الترجيح الأساسية الأولى  $w_{ij,b}$  بالعامل  $F_{ue}$ .

### ٦ - ٣ - ٢ تعديلات مُعاملات ترجيح العيّنة لمراعاة النسخ المكررة

١٣ - إذا عُرف من قبل أن هناك نسخاً مكررة من بعض الوحدات في الإطار، يمكن تعويض الاحتمال الزائد لاختيار هذه الوحدات بإعطائها مُعاملات ترجيح هي معكوسات عدد النسخ المكررة في إطار المعاينة، إذا أُدرجت هذه الوحدات في العيّنة. غير أن النسخ المكررة لا تُكتشف في أغلب الأحيان إلا بعد اختيار العيّنة، وينبغي تعديل احتمالات اختيار هذه الوحدات المدرجة في العيّنة لمراعاة النسخ المكررة. يُنفذ هذا التعديل على النحو التالي. افترض أن الوحدة رقم  $i^{th}$  المدرجة في العيّنة لها احتمال اختيار مشار إليه بـ  $p_{i1}$  وافترض أنه توجد سجلات إضافية عددها  $k-1$  في إطار المعاينة، عرّفها الوحدة المدرجة في العيّنة بأنها نسخ مكررة، ولكل نسخة منها احتمالات اختيار نِسبها كما يلي:  $p_{i2}, \dots, p_{ik}$ . عندئذٍ يُعطى احتمال الاختيار المعدل للوحدة المعنية المدرجة في العيّنة بـ

(٥ - ٦)

$$p_i = 1 - (1 - p_{i1})(1 - p_{i2}) \dots (1 - p_{ik})$$

عندئذٍ تُرجح الوحدة المدرجة في العيّنة بناءً على ذلك، بمعنى أنها تُرجح بـ  $1/p_i$ .

١٤ - ترد أدناه توضيحات بأمثلة محددة لإجراءات بناء مُعاملات ترجيح العيّنة بموجب السيناريوهات الموجزة أعلاه.

### ٦ - ٤ الترجيح لاحتمالات الاختيار غير المتساوية

١٥ - لتيسير الشرح دعونا ننظر في تصميم ذي مرحلتين واستخدام مناطق التعداد كوحدات معاينة أولية، والأسر المعيشية كوحدات للمرحلة الثانية. لنفترض أننا اخترنا عيّنة متساوية الاحتمالات فيها العدد  $m$  من وحدات المعاينة الأولية، اخترت ممّا مجموعه  $N$  في المرحلة الأولى، ثم اخترت  $m$  أسر معيشية من كل وحدة معاينة أولية أُدرجت في العيّنة. من البديهي أن احتمال انتخاب الأسرة المعيشية يتوقف على عدد الأسر المعيشية في وحدة المعاينة

الأولى التي توجد الأسرة فيها. دع  $M_i$  يشير إلى عدد الأسر المعيشية في وحدة المعاينة الأولى  $i$ . عندئذ يكون احتمال اختيار وحدة المعاينة الأولى  $n/N$  والاحتمال المشروط لاختيار أسرة من وحدة المعاينة الأولى رقم  $i^b$  المدرجة في العيّنة  $m/M_i$ . لذلك، يُعطى الاحتمال الإجمالي لاختيار أسرة بـ

$$(٦-٦) \quad p_{ij} = p_i \times p_{j(i)} = \frac{n}{N} \times \frac{m}{M_i} = \frac{nm}{N} \times \frac{1}{M_i}$$

وكذلك يُعطى مُعامِل ترجيح الأسرة المدرجة في العيّنة بموجب هذا التصميم بـ

$$(٧-٦) \quad w_i = \frac{1}{p_{ij}} = \frac{N}{nm} \times M_i$$

### مثال

اختبرت عيّنة متساوية الاحتمالات تضم ٥ أسر معيشية من ٢٥٠ أسرة. واختير شخص بالغ اختياريًا عشوائيًا من كل أسرة مدرجة في العيّنة. وقد سُجّل الدخل الشهري ( $y_{ij}$ ) ومستوى التعليم ( $z_{ij} = 1$  إن كان قد أنهى الدراسة الثانوية أو أعلى منها؛ وصفر إن لم يُنه الثانوية) للشخص البالغ رقم  $i^b$  الذي أُدرج في العيّنة في الأسرة. دع  $M_i$  تشير إلى عدد الأشخاص البالغين في الأسرة رقم  $i^b$ . عندئذ يُعطى الاحتمال الإجمالي لاختيار الشخص البالغ بـ

$$p_{ij} = p_i \times p_{j(i)} = \frac{5}{250} \times \frac{1}{M_i} = \frac{1}{50} \times \frac{1}{M_i}$$

لذلك يُعطى مُعامِل ترجيح الشخص البالغ الذي أُدرج في العيّنة بـ

$$.w_i = \frac{1}{p_{ij}} = 50 \times M_i$$

١٦ - نوضّح الآن حساب التقديرات الأساسية بموجب هذا التصميم. دعونا نفترض أن البيانات التي حُصل عليها من الشخص البالغ الأعزب المدرج في العيّنة لكل أسرة معيشية في عيّنة المرحلة الأولى التي تضم ٥ أسر، كما هي واردة في الجدول ٦ - ٢ أدناه. لاحظ أن عدد الأشخاص البالغين في كل أسرة معيشية، ومُعامِل الترجيح الإجمالي المتناسق مع هذا للشخص البالغ المدرج في العيّنة هو المعطى في العمودين الثاني والثالث على التوالي.

الجدول ٦ - ٢

مُعامِلات الترجيح بموجب احتمالات الاختيار غير المتساوية

الأسرة المعيشية المدرجة في العيّنة	$M_i$	$w_i$	$y_{ij}$	$z_{ij}$	$w \cdot y_{ij}$	$w \cdot z_{ij}$	$w \cdot z_{ij} \cdot y_{ij}$
١	٣	١٥٠	٧٠	١	١٠٥٠٠	١٥٠	١٠٥٠٠
٢	١	٥٠	٣٠	صفر	صفر	صفر	صفر
٣	٣	١٥٠	٩٠	١	١٣٥٠٠	١٥٠	١٣٥٠٠
٤	٥	٢٥٠	٥٠	١	١٢٥٠٠	٢٥٠	١٢٥٠٠
٥	٤	٢٠٠	٦٠	صفر	صفر	صفر	صفر
المجموع	١٦	٨٠٠	٣٠٠	٣	٣٦٥٠٠	٥٥٠	٥٠٠٠٠

١٧ - يمكن عندئذٍ الحصول على تقديرات لمختلف الخصائص من الجدول ٦ - ٢ كما يلي:

تقدير متوسط الدخل الشهري هو

$$\bar{y}_w = \frac{\sum w_i y_{ij}}{\sum w_i} = \frac{50\,000}{800} = 62.5$$

إذا لم تُستخدم مُعامِلات الترجيح يكون هذا التقدير ٦٠ (أو ٥/٣٠٠).

تقدير نسبة الأشخاص الذين أتموا المرحلة الثانوية أو التعليم العالي هو

$$\bar{y}_w = \frac{\sum w_i z_{ij}}{\sum w_i} = \frac{550}{800} = 0.68.75\% \text{ أو } 68.75\%$$

إذا لم تُستخدم مُعامِلات الترجيح يكون هذا التقدير ٥/٣ أو ٠,٦٠ في المائة.

مجموع الأشخاص الذين أتموا المرحلة الثانوية أو التعليم العالي هو

$$\hat{t} = \sum w_i z_{ij} = 550$$

تقدير متوسط الدخل الشهري للبالغين الذي يحملون شهادة الدراسة الثانوية أو أعلى منها هو

$$\bar{y}_w = \frac{\sum w_i z_{ij} y_{ij}}{\sum w_i z_{ij}} = \frac{36\,500}{550} = 66.36$$

١٨ - "تربّع" مُعامِلات ترجيح العيّنات أحياناً، بمعنى أن مُعامِلات الترجيح تُضرب بالنسبة

(٦ - ٨)

عدد المستجيبين

مجموع مُعامِلات ترجيح جميع المستجيبين

١٩ - لذلك، كان مجموع مُعامِلات الترجيح العادية هو حجم العيّنة المحقق للتحليل (عدد المستجيبين). لاحظ أن معايير الترجيح المربعة لا يمكن استخدامها لتقدير المجاميع، مثل مجموع عدد الأشخاص البالغين الحاصلين على شهادة الدراسة الثانوية أو أعلى منها. في هذه الحالة ينبغي ترجيح الوحدة المدرجة في العيّنة بمعكوس احتمالات اختيارها، بعبارة أخرى يجب استخدام مُعامِلات الترجيح العادية. غير أنه لتقدير المتوسطات والنسب تحتاج مُعامِلات الترجيح إلى أن تكون متناسبة فقط مع معكوسات احتمالات الاختيار. بعبارة أخرى، لا يهم إن استُعملت مُعامِلات الترجيح العادية أو المربعة (التي تكون متناسبة مع مُعامِلات الترجيح العادية) للحصول على تقديرات متوسطات مَعَلَمات السكان، كمتوسط عدد النساء أو نسبة عدد النساء اللاتي هن في سن الحمل والولادة الحاصلات على العناية الصحية الأولية. فكلتا النوعين من مُعامِلات الترجيح يأتي بنفس النتيجة.

٢٠ - في المثال السابق، مثلاً، تكون مُعامِلات الترجيح متناسبة مع  $w_i$ . فإذا استخدمنا  $M_i$  ( $w_i = 50 * M_i$ ) كمُعامِلات ترجيح كان تقدير نسبة الأشخاص الحاصلين على شهادة الدراسة الثانوية أو أعلى منها هو

$$\hat{p} = \frac{\sum M_i z_{ij}}{\sum M_i} = \frac{3 \times 1 + 1 \times 0 + 3 \times 1 + 5 \times 1 + 4 \times 0}{3 + 1 + 3 + 5 + 4} = \frac{11}{16} = 0.6875$$

أو ٦٨,٧٥ في المائة، وهو نفس العدد السابق بالضبط. غير أنه لتقدير مجموع عدد البالغين الحاصلين على شهادة الدراسة الثانوية أو أعلى منها، يجب استخدام مُعامِلات الترجيح العادية للحصول على النتيجة الصحيحة، بمعنى

$$\hat{t}_s = \sum (50 \times M_i) z_{ij} = 50 \sum M_i z_{ij} = 50 \times 11 = 550$$

تُختار عيّنة على مرحلتين من الأسرة المعيشية في المناطق الريفية من البلد. في المرحلة الأولى تُدخل ٥٠ قرية في العيّنة باحتمال متناسب مع عدد الأسر القاطنة فيها في وقت آخر تعداد السكان. ومما يذكر أن مجموع عدد الأسر المعيشية في المناطق الريفية في وقت آخر تعداد السكان كان ٣٠٠٠٠٠ أسرة. وأتبع اختيار العيّنة في المرحلة الأولى بعملية إعداد قوائم تهدف إلى تجميع قوائم الوحدات السكنية لكل واحدة من القرى المختارة. وقد وُجد أحياناً أن الوحدة السكنية المنفردة تضم أكثر من أسرة معيشية واحدة.

٢١ - ننظر الآن في عدة خيارات لتصميم تقسيم العيّنة إلى عيّنات فرعية (لاختيار أسر معيشية من وحدات سكنية مختارة) ونحدد معادلة الاختيار للاحتمال الإجمالي لاختيار أسرة معيشية لإدراجها في العيّنة. دع  $D_i$  تشير إلى عدد الوحدات السكنية في القرية  $i$  ودع  $H_{ij}$  تشير إلى عدد الأسر في الوحدة السكنية  $j$  من القرية  $i$ . فإن مجموع عدد الأسر المعيشية في القرية، مشاراً إليه بـ  $H_i$ ، يُعطى بـ

$$\sum_i H_i = \sum_i \sum_j H_{ij} = 300\,000 \quad \text{لاحظ أن } H_i = \sum_j H_{ij}$$

تستند احتمالات الاختيار المحسوبة هنا إلى معادلات قُدّمت في الفصل الثالث.

### خيار التصميم ١

٢٢ - اختيرت ١٥ وحدة سكنية بالمعينة البسيطة المختارة عشوائياً دون تبديلها من قائمة كل قرية مختارة. وأدرجت في العيّنة جميع الأسر القاطنة في الوحدات السكنية المختارة، لذلك توجد مرحلتان فقط لاختيار العيّنة: مرحلة اختيار القرى ومرحلة اختيار الوحدات السكنية. بموجب هذا التصميم، فإن معادلة الاختيار للاحتمال الإجمالي لاختيار الأسر المعيشية لإدراجها في العيّنة يُعطى بـ

$$P_{ij} = pr \text{ (اختيرت الوحدة السكنية } j \text{ بحكم أن } i \text{ اختيرت)} * pr \text{ (اختيرت القرية } i \text{)}$$

ثمَّ:

$$p_{ij} = 50 \times \frac{H_i}{\sum_i H_i} \times \frac{15}{D_i} = \frac{750}{\sum_i M_i} \times \frac{H_i}{D_i}$$

ويعطى مُعامِل الترجيح الأساسي بـ

$$w_{ij} = \frac{\sum_i M_i}{750} \times \frac{D_i}{H_i}$$

٢٣ - لاحظ أن الاحتمال الإجمالي للاختيار سيتباين من قرية إلى أخرى، بحسب نسبة عدد الأسر المعيشية إلى عدد الوحدات السكنية  $H_i/D_i$ . لذلك نستنتج أن هذا التصميم ليس ذاتي الترجيح (للاطلاع على مزيد من مناقشة التصاميم ذاتية الترجيح انظر الفرع ٦ - ٤ - ٢ أدناه). وسيكون ذاتي الترجيح إذا كانت كل وحدة سكنية فيها أسرة معيشية واحدة، بمعنى أنه إذا كانت النسبة  $H_i/D_i$  هي نفسها لجميع القرى المدرجة في العيّنة.

## خيار التصميم ٢

٢٤ - تُدرجُ المساكن في العيّنات بطريقة منهجية داخل القرى المختارة بنسبة معاينة في القرية تكون متناسبة تناسباً عكسياً مع عدد أسرها المعيشية في وقت آخر تعداد للسكان. أُدرجت في العيّنة جميع الأسر المعيشية القاطنة في الوحدات السكنية المختارة. يمكن التعبير عن الاحتمال المشروط لاختيار مسكن في قرية مختارة رقمها  $i$  بأنه  $k/H_i$ ، حيث  $k$  هو ثابت التناسب. لذلك فإن معادلة الاختيار للاحتمال الإجمالي لاختيار أسرة معيشية لإدراجها في العيّنة بموجب هذا التصميم تُعطى بـ

$$p_{ij} = 50 \times \frac{H_i}{\sum_i H_i} \times \frac{k}{H_i} = \frac{50 \times k}{\sum_i M_i}$$

ومُعامل الترجيح الأساسي يُعطى بـ

$$w_{ij} = \frac{\sum_i M_i}{50 \times k}$$

وهو ثابت. لذلك نستنتج أن خيار التصميم ٢ تصميم ذاتي الترجيح.

## خيار التصميم ٣

٢٥ - تُدرجُ المساكن في العيّنة منهجياً داخل قرى مختارة مع نسبة معاينة للقرية تتناسب تناسباً عكسياً مع عدد الأسر التي كانت فيها وقت تعداد السكان الأخير. وتُختار أسرة معيشية واحدة اختياراً عشوائياً في كل مسكن مختار. في هذه الحالة توجد ثلاث مراحل للاختيار: القرى، والمساكن، والأسر المعيشية. لذلك، فإن معادلة الاختيار للاحتمال الإجمالي لاختيار أسرة معيشية لإدراجها في العيّنة بموجب هذا التصميم تُعطى بـ

$$p_{ij} = 50 \times \frac{H_i}{\sum_i H_i} \times \frac{k}{H_i} \times \frac{1}{H_{ij}}$$

ويُعطى مُعامل الترجيح الأساسي بـ

$$w_{ij} = \frac{\sum_i M_i}{50} \times \frac{H_{ij}}{k}$$

وهذا يتباين من وحدة سكنية إلى أخرى بحسب عدد الأسر القاطنة في الوحدة السكنية. لذلك نستنتج أن خيار التصميم ٣ ليس ذاتي الترجيح.

## ٦ - ٤ - ١ دراسة حالة إفرادية في مُعاملات الترجيح:

### استقصاء فييت نام الصحي الوطني، ٢٠٠١

٢٦ - نمضي الآن إلى توضيح بناء مُعاملات ترجيح العيّنات لأجل استقصاء فعلي، وهو الاستقصاء الصحي الوطني الذي أُجري في فييت نام في سنة ٢٠٠١. أُجري الاستقصاء على أساس تصميم عيّنة طبقية ذات ثلاث مراحل. كان هناك ما مجموعه ١٢٢ طبقة، عُرِّفت بمجالات حضرية أو ريفية في ٦١ إقليمياً. أُجري اختيار العيّنات مستقلاً عن غيره داخل كل طبقة. في المرحلة الأولى اختيرت كميونات أو أحياء باحتمال متناسب مع الحجم (عدد

الأُسُر في وقت إجراء تعداد السكان والمساكن في سنة ١٩٩٩). وفي المرحلة الثانية، اختيرت منطقتا تعداد اثنتان في كل كميونة أو حيٍّ مُدرَج في العيّنَة بمعايينة منهجية، بنسبة معايينة تتناسب تناسباً عكسياً مع عدد مناطق التعداد في الكميونة أو الحي. وفي المرحلة الثالثة والأخيرة، اختيرت ١٥ أسرة معيشية في كل منطقة تعداد مدرجة في العيّنَة، وهنا أيضاً كان اختيارها بطريقة معايينة منهجية.

٢٧ - كانت مُعاملات ترجيح العيّنَة الأساسية للأُسُر المعيشية المدرجة في العيّنَة، في إطار تصميم الاستقصاء الصحي الوطني، يمكن تطويرها كما يلي. دع  $H_i$  و  $E_i$  تشيران إلى عدد الأُسُر المعيشية وعدد مناطق التعداد، على التوالي (في وقت تعداد السكان في سنة ١٩٩٩) في الكميونة  $i$ ، ودع  $H_{ij}$  تشير إلى عدد الأُسُر المعيشية في منطقة التعداد  $j$  في الكميونة  $i$ . ثم يُعطى الاحتمال الإجمالي لاختيار الأُسُر المعيشية  $k$  في منطقة التعداد  $j$  من الكميونة  $i$ ، بـ

$$p_{ijk} = n_c \times \frac{H_i}{\sum_i H_i} \times \frac{2}{E_i} \times \frac{15}{H_{ij}}$$

حيث  $n_c$  هو عدد الكميونات المختارة في طبقة معيّنَة و

$$\sum_i H_i \text{ هو مجموع عدد الأُسُر المعيشية في تلك الطبقة.}$$

وَمُعَامِلُ ترجيح معايينة الأُسُرَة ( $w_{ijk}$ ) هو معكوس احتمال الاختيار، بمعنى أن

(٦ - ٩)

$$w_{ijk} = \frac{E_i \times H_{ij} \times \sum_i M_i}{30 \times n_c \times H_i}$$

## ٦ - ٤ - ٢ العيّنات ذاتية الترجيح

٢٨ - عندما تكون مُعاملات ترجيح جميع الوحدات المدرجة في العيّنَة متساوية، يشار إلى العيّنَة بأنها ذاتية الترجيح. حتى وإن كانت وحدات المرحلة العليا غالباً ما تُختار باحتمالات متباينة لأسباب تتعلق بكفاءة المعايينة، يمكن أن تُلغى هذه الاحتمالات المتباينة مقابل احتمالات الاختيار في مراحل لاحقة. خيار التصميم ٢ في المثال الوارد أعلاه يعطي مثلاً لهذا الوضع.

٢٩ - غير أن عيّنات استقصاء الأُسُر المعيشية، في الواقع، نادراً ما تكون ذاتية الترجيح على الصعيد الوطني، لعدة أسباب. أولاً، الوحدات المدرجة في العيّنَة غالباً ما تُختار عن قصد باحتمالات متباينة للاختيار. بل إنه مع أن وحدات المعايينة الأولية غالباً ما تُختار باحتمالات متناسبة مع الحجم، وتُختار الأُسُر المعيشية بنسبة ملائمة داخل وحدات المعايينة الأولية، لكي تنتج تصميمات ذاتية الترجيح، وقد يُلغى هذا باختيار شخص واحد للمقابلة من كل أسرة معيشية مدرجة في العيّنَة. ثانياً، غالباً ما يكون في العيّنَة المختارة أوجه قصور منها عدم الاستجابة (الفرع ٦ - ٥) وعدم التغطية (الفرع ٦ - ٦). ثالثاً، غالباً ما تستدعي الحاجة إلى تقديرات دقيقة للمجالات والمجموعات السكانية الفرعية الإفراط في معايينة هذه المجالات بغية الحصول على أحجام عيّنات كبيرة بما يكفي لتلبية احتياجات الدقة المحددة سلفاً. رابعاً، عندما يستدعي تصميم العيّنَة إعداد قوائم حديثة بالأُسُر المعيشية في المجموعات المختارة (وحدات معايينة أولية أو وحدات معايينة في المرحلة الثانية) وسيُختار عدد ثابت محدد سلفاً للأُسُر المعيشية من كل مجموعة، فإن الاحتمال الفعلي لاختيار الأُسُرَة يختلف إلى حد ما عن الاحتمال المصمّم لاختيارها، الذي يستند

إلى الأعداد الواردة في الإطار لا إلى الأعداد الراهنة للأسر؛ ومن ثم تظهر احتمالات غير متساوية للاختيار، حتى وإن كان الهدف إنتاج تصميم ذاتي الترجيح.

٣٠ - على الرغم من العوائق، ينبغي أن يكون الحصول على عيّنات ذاتية الترجيح هدف كل عملية تصميم عيّنة بسبب الميزات التي تقدمها فيما يتعلق بتنفيذ التصميم وتحليل البيانات التي يولدها التصميم. فبالعيّنات ذاتية الترجيح، يمكن أن تُستمدّ تقديرات الاستقصاء من بيانات غير مرّجحة، وبذلك يمكن أن تُضخّم النتائج، عند الضرورة، بعامل ثابت للحصول على تقديرات ملائمة لمعلومات السكان. علاوةً على ذلك، التحليلات التي تستند إلى عيّنات ذاتية الترجيح أكثر ميلاً لسلوك اتجاه مباشر، وتكون النتائج أيسر فهماً، ومقبولة لدى غير الإحصائيين والجمهور العام.

## ٦ - ٥ تعديل مُعاملات ترجيح العيّنات لعدم الاستجابة

٣١ - نادراً ما يكون الحال أن كل المعلومات المرغوبة يُحصل عليها من جميع الوحدات المدرجة في عيّنات الاستقصاءات. فيمكن، مثلاً، ألاّ تقدم بعض الأسر المعيشية أيّ بيانات على الإطلاق، بينما تقدم أسر معيشية أخرى بيانات جزئية فقط، بمعنى بيانات عن بعض الأسئلة التي في الاستقصاء لا كلها. يُسمّى النوع الأول من عدم الاستجابة عدم استجابة الوحدة أو عدم الاستجابة الكلي، بينما سُمّي النوع الأخير عدم الإجابة على بند. وإذا وجدت أيّ اختلافات منهجية بين المستجيبين وغير المستجيبين، عندها ستكون التقديرات المستندة ببراءة إلى المستجيبين فقط تقديرات متحيزة. ومن النقاط الرئيسية في الممارسات الجيدة في الاستقصاءات نقطة نؤكدُها في كل فروع هذا الدليل، وهي أن من الأهمية بمكان أن نجعل عدم الاستجابة في الاستقصاء عند أدنى حد ممكن. هذا ضروري لتخفيض إمكانية أن تكون تقديرات الاستقصاء متحيزة بطريقة ما، بعدم إدراجها جزءاً معيناً من السكان (أو إدراجها كجزء صغير غير متناسب من الإسكان). فالأشخاص الذين يعيشون في مناطق حضرية وتكون مداخيلهم مرتفعة نسبياً ربما يكونون أقل ترجيحاً للمشاركة في استقصاء متعدد الأغراض يشمل بنوداً عن الدخل. وإن عدم القدرة على الحصول على استجابات من قطاع كبير من هذه الفئة من السكان يمكن أن يؤثر في التقديرات الوطنية لمتوسط دخل الأسرة المعيشية، والتحصيل العلمي، ومعرفة القراءة والكتابة، إلخ.

## ٦ - ٥ - ١ تقليل التحيز الناتج عن عدم الاستجابة في استقصاءات الأسر المعيشية

٣٢ - حجم التحيز الناتج عن عدم الاستجابة لعيّنة متوسطة، مثلاً، هو دالّة لعاملين اثنين:

- نسبة السكان الذين لا يستجيبون
- حجم الفرق بين متوسط السكان ذوي الخاصية التي هي موضع اهتمام المجموعات المستجيبة والمجموعات غير المستجيبة

٣٣ - لذلك، يستدعي تقليل التحيز الناتج عن عدم الاستجابة إما أن تكون نسبة غير المستجيبين صغيرة، وإما أن تكون الفروق صغيرة بين المستجيبين وغير المستجيبين من الأسر المعيشية والأشخاص. يمكن، بواسطة حفظ السجلات حفظاً متقناً، لكل وحدة مدرجة في العيّنة تُختار للاستقصاء، أن تُقدّر بصورة مباشرة من بيانات الاستقصاء نسبة عدم الاستجابة للعيّنة كلها والمجالات الفرعية التي هي موضع اهتمام. علاوةً على ذلك، يمكن تنفيذ دراسات خاصة مصمّمة بعناية لتقييم الفروقات بين المستجيبين وغير المستجيبين (غروفر وكوبر، ١٩٩٨).

٣٤ - في استقصاءات الأفرقة (التي تُجمع فيها البيانات من نفس فريق الوحدات المدرجة في العيّنة مراراً وتكراراً على مرّ الزمن) يستطيع مصمّم الاستقصاء الحصول على بيانات أكثر عن آثار التحيز الناتج عن عدم الاستجابة المحتمل، وإدخال التعديلات اللازمة لمقابلة هذا التحيز، أكثر مما يحصل في الاستقصاءات التي تجرى مرة واحدة أو الاستقصاءات الشاملة لعدة قطاعات. فهنا ربما يَنْتُج عدم الاستجابة من فقدان بعض الوحدات أثناء الاستقصاء، أو رفضها المشاركة في حلقات لاحقة من الاستقصاء نتيجة للتعب من الإجابة، أو لعوامل أخرى، وما أشبه ذلك. ويمكن عندئذٍ استخدام البيانات التي جُمعت في حلقات سابقة من استقصاء الفريق، لتعلم المزيد عن الفروقات بين المستجيبين وغير المستجيبين ويمكن أن تعمل مثل التحيز لنوع التعديلات الوارد وصفها أدناه. وترد تفاصيل أكثر عن مختلف التقنيات المستخدمة للتعويض عن عدم الاستجابة في بحوث الاستقصاء، في بريك وكالتون (١٩٩٦)، ولييكوفسكي (٢٠٠٣) والمراجع المشار إليها فيهما.

### ٦ - ٥ - ٢ التعويض عن عدم الاستجابة

٣٥ - يمكن استخدام عدد من التقنيات لزيادة نسب الاستجابة، ومن ثم تخفيض التحيز المرتبط بعدم الاستجابة في استقصاءات الأسر المعيشية. من هذه التقنيات تحويل الرافض بواسطة "معاودة الاتصال"، التي يبذل فيها المُقابِلون لا محاولة واحدة، بل بضع محاولات، لإكمال مقابلة مع أسرة معيشية مدرجة في العيّنة. يمكن أيضاً تحسين نسب الاستجابة المرتفعة بتدريب المقابِلين تدريباً أفضل. غير أنه بغضّ النظر عن كبر الجهد الذي يبذل لتحسين نسب الاستجابة سيظل عدم الاستجابة دائماً ظاهرة محتومة في كل استقصاء للأسر المعيشية. من ثمّ غالباً ما يُجري مصمّمو الاستقصاءات تعديلات للتعويض عن عدم الاستجابة. والنهجان الأساسيان المُتبعان في التعديل مقابل عدم استجابة الوحدة هما:

(أ) تعديل حجم العيّنة بأخذ عيّنة أولية أكبر مما تدعو إليه الحاجة لأخذ عدم الاستجابة المتوقّع في الحساب؛

(ب) تعديل مُعاملات ترجيح العيّنة لأخذ عدم الاستجابة في الحساب.

٣٦ - يُنصَح في التعامل مع عدم استجابة الوحدة في استقصاءات الأسر المعيشية بتناولها دائماً بتعديل مُعاملات ترجيح العيّنة لكي تأخذ في الحساب الأسر المعيشية غير المستجيبة. يعطي الفرع ٦ - ٥ - ٣ موجزاً بالخطوات المتخذة لتعديل مُعاملات ترجيح العيّنة بسبب عدم الاستجابة، مشفوعة بمثال رقمي.

٣٧ - توجد بضع مشاكل مرتبطة بالتبديل، الذي يعتبر مكافئاً لعزو سجلّ الوحدة غير المستجيبة بأكمله (كالتون، ١٩٨٣). أولاً، يزيد احتمالات الاختيار للبدائل المحتملة، لأن الأسر غير المدرجة في العيّنة القريبة من الأسر غير المستجيبة المدرجة في العيّنة أكثر احتمالاً لاختيارها من الأسر القريبة من الأسر المستجيبة المدرجة في العيّنة. ثانياً، محاولات تبديل الأسر غير المستجيبة تستغرق وقتاً طويلاً، وهي عرضة للأخطاء والتحيز، ومن الصعب جداً تفحصها ومراقبتها. إذ يمكن التبديل، مثلاً، باستخدام أسرة مريحة بدلاً من الأسرة المعيّنة خصيصاً لتكون بديلة لأسرة غير مستجيبة، وبذلك إدخال مصدر آخر للتحيز. بسبب هذه المشاكل كلها، يجب عدم استخدام التبديل للتعويض عن عدم الاستجابة في استقصاءات الأسر المعيشية، ما لم يوجد مبرر جيد لتطبيق معين.

٣٨ - الأسلوب المعياري للتعويض، في حالة عدم الاستجابة الجزئي أو عدم الإجابة على بند من البنود، هو العزو، وهذا غير مشمول بهذا الدليل.

## ٦ - ٥ - ٣ تعديل مُعامِلات ترجيح العيّنة بسبب عدم الاستجابة

٣٩ - يُستعمل إجراء تعديل مُعامِلات ترجيح العيّنة في كثير من الأحيان للتعويض عن عدم الاستجابة في الاستفتاءات الكبيرة للأسر المعيشية. والتعديل، أساساً، ينقل مُعامِلات الترجيح الأساسية لجميع الوحدات المؤهلة المدرجة في العيّنة غير المستجيبة إلى الوحدات المستجيبة، ويُنفَّذ في الخطوات التالية:

- الخطوة الأولى: طَبَّق مُعامِلات ترجيح التصميم الأولى (لاحتمالات الاختيار غير المتساوية والتعديلات الأخرى التي بُجِّتت في الفروع السابقة، إن كان ذلك منطبقاً).
- الخطوة الثانية: قَسَّم العيّنة إلى عيّنات فرعية واحسب نسب الاستجابة المرجحة لكل مجموعة فرعية.
- الخطوة الثالثة: استخدم معكوس نسب الاستجابة في المجموعة الفرعية لإجراء تعديلات لتعويض عدم الاستجابة.
- الخطوة الرابعة: احسب مُعامِل الترجيح المعدل للتعويض عن عدم الاستجابة لوحدة العيّنة رقم  $i^{th}$  كما يلي:

(٦ - ١٠)

$$w_i = w_{1i} * w_{2i}$$

حيث يكون  $w_{1i}$  مُعامِل الترجيح الأوّلي، ومُعامِل الترجيح المعدل لتعويض عدم الاستجابة. لاحظ أنه يمكن تعريف نسبة عدم الاستجابة المرجح بأنه نسبة العدد المرجح للمقابلات التي أُجريت مكتملة مع الحالات المؤهلة المدرجة في العيّنة إلى العدد المرجح للحالات المؤهلة المدرجة في العيّنة.

## مثال

اختيرت عيّنة متعددة المراحل ومقسمة إلى طبقات قوامها ١٠٠٠ أسرة معيشية اختيرت من منطقتين في بلد ما (شمالية وجنوبية). أُدرجت الأسر المعيشية القاطنة في الشمال في العيّنة بنسبة ١/١٠٠ وأدرجت الأسر التي في الجنوب بنسبة ١/٢٠٠. وكانت نسب الاستجابة في المناطق الحضرية أدنى منها في المناطق الريفية. دع  $n_b$  تشير إلى عدد الأسر المعيشية المدرجة في العيّنة في الطبقة  $b$ ، ودع  $r_b$  تشير إلى عدد الأسر المعيشية المؤهلة التي استجابت للاستقصاء ودع  $t_b$  تشير إلى عدد الأسر المستجيبة التي تحصل على العناية الصحية الأولى. ثم يُعطى مُعامِل الترجيح المعدل للتعويض عن عدم الاستجابة للأسر المدرجة في الطبقة  $b$

(٦ - ١١)

$$w_b = w_{1b} * w_{2b}$$

حيث  $w_{2b} = n_b/r_b$ . افترض أن البيانات على مستوى الطبقة هي كالبيانات الواردة في الجدول ٦ - ٣.

الجدول ٦ - ٣  
تعديل مُعامِلات الترجيح تعويضاً عن عدم الاستجابة

الطبقة	$n_b$	$r_b$	$t_b$	$w_{1b}$	$w_{2b}$	$w_b$	$w_b r_b$	$w_b t_b$
الشمال - حضرية	١٠٠	٨٠	٧٠	١٠٠	١,٢٥	١٢٥	١٠٠٠٠	٨٧٥٠
الشمال - ريفية	٣٠٠	١٢٠	١٠٠	١٠٠	٢,٥٠	٢٥٠	٣٠٠٠٠	٢٥٠٠٠
الجنوب - حضرية	٢٠٠	١٧٠	١٥٠	٢٠٠	١,١٨	٢٣٦	٤٠١٢٠	٣٥٤٠٠
الجنوب - ريفية	٤٠٠	٣٦٠	١٨٠	٢٠٠	١,١١	٢٢٢	٧٩٩٢٠	٣٩٩٦٠
المجموع	١٠٠٠	٧٣٠	٥٠٠				١٦٠٠٤٠	١٠٩١١٠

لذلك تكون نسبة الأسر المعيشية التي تحصل على العناية الصحية الأولية:

$$\hat{p} = \frac{\sum w_b t_b}{\sum w_b r_b} = \frac{109\ 110}{160\ 040} = 0.682 \text{ أو } 68.2\%$$

والعدد المقدر للأسر التي تحصل على العناية الصحية الأولية هو:

$$\hat{t} = \sum w_b t_b = 109\ 110 = 68\% \text{ من } 160\ 040$$

لاحظ أن النسبة المقدرة غير المرجحة للأسر المعيشية التي تحصل على العناية الصحية الأولية، باستخدام الأسر المستجيبة فقط هي:

$$\hat{p}_{uv} = \frac{\sum t_b}{\sum r_b} = \frac{500}{730} = 0.685 \text{ أو } 68.5\%$$

وأن النسبة المقدرة باستخدام مُعامِلات الترجيح الأولية دون تعديل لتعويض عدم الاستجابة هي:

$$\hat{p}_1 = \frac{\sum w_{1b} t_b}{\sum w_{1b} r_b} = \frac{83\ 000}{126\ 000} = 0.659 \text{ أو } 65.9\%$$

٤٠ - قدّم هذا المثال أيضاً بغية توضيح كيف تعدّل مُعامِلات الترجيح الأولية لتعويض عدم الاستجابة. وتبين النتائج فرقاً كبيراً بين النسبة المقدرة باستخدام مُعامِلات الترجيح الأولية فقط والنسبة المقدرة باستخدام مُعامِلات الترجيح المعدلة لتعويض عدم الاستجابة، لكن الفرق بين النسبة غير المرجحة والنسبة المعدلة لتعويض عدم الاستجابة يبدو ضئيلاً.

٤١ - بعد تعديل مُعامِلات الترجيح لتعويض عدم الاستجابة، يمكن إجراء تعديلات أخرى لمُعامِلات الترجيح، حسب الاقتضاء. يُنظر في الفرع التالي تعديل مُعامِلات الترجيح لأخذ عدم التغطية في الحسبان.

## ٦ - ٦ تعديل مُعامِلات ترجيح العينة لتعويض عدم التغطية

٤٢ - تشير عبارة عدم التغطية إلى عدم تغطية إطار المعاينة لجميع المجموعات السكانية المستهدفة، وكانت نتيجة ذلك أن بعض الوحدات السكانية لا يوجد احتمال لاختيارها لإدراجها في العينة المختارة لاستقصاء الأسر المعيشية.

هذا مجرد وجه واحد من أوجه قصور كثيرة تقع فيها أُطر المعاينة المستخدمة لاختيار عيّنات للاستقصاءات. انظر الفصل الرابع للاطلاع على مناقشة مفصلة لأُطر المعاينة.

٤٣ - عدم التغطية من دواعي القلق الرئيسية إزاء استقصاءات الأسر المعيشية، لا سيما تلك التي تُجرى في البلدان النامية. يمكن رؤية الدليل على أثر عدم التغطية في حقيقة أن تقديرات العيّنة لعدد السكان بناءً على أساس بعض استقصاءات البلدان النامية تأتي أقل كثيراً من تقديرات السكان من مصادر أخرى. لذلك، يجب أن تكون منهجية تعريف عدم التغطية في استقصاءات الأسر المعيشية وتقييمه ومراقبته مجالاً رئيسياً من مجالات العمل والتدريب في المكاتب الوطنية للإحصاء.

٤٤ - يناقش هذا الفرع بعض مصادر عدم التغطية في استقصاءات الأسر المعيشية وواحداً من الإجراءات المستخدمة لتعويض عدم التغطية، وهو التعديل الإحصائي لمعاملات الترجيح بواسطة إجراءات ما بعد التقسيم الطبقي.

## ٦ - ٦ - ١ مصادر عدم التغطية في استقصاءات الأسر المعيشية

٤٥ - معظم استطلاعات الأسر المعيشية في البلدان النامية تستند إلى تصاميم احتمالات المناطق متعددة المراحل والمقسمة إلى طبقات. فوحدات المرحلة الأولى، أو وحدات المعاينة الأولية، هي في العادة وحدات مناطق جغرافية. وفي المرحلة الثانية تُنشأ قائمة بالأسر المعيشية تُختار منها عيّنة الأسر المعيشية. لذلك يمكن أن يحدث عدم التغطية في أي واحد من المستويات الثلاثة: مستوى وحدات المعاينة الأولية، ومستوى الأسرة المعيشية، ومستوى الشخص.

٤٦ - نظراً إلى أن وحدات المعاينة الأولية تستند بوجه عام إلى مناطق تعداد عُرِّفت واستُخدمت في عمليات تعداد سابقة للسكان وللأسر المعيشية، يُتوقَّع منها أن تغطي المدى الجغرافي للسكان المستهدفين بكامله. لذلك، يكون حجم عدم تغطية وحدات المعاينة الأولية صغيراً بوجه عام. وإن عدم تغطية وحدات المعاينة الأولية في استقصاءات الأسر المعيشية في البلدان النامية ليس بالخطورة التي ينطوي عليها عدم التغطية في المراحل اللاحقة من مراحل التصميم. غير أن عدم تغطية وحدات المعاينة الأولية يحدث في معظم الاستقصاءات. فمع أن الاستقصاء ربما يُصمَّم لتوفير تقديرات للسكان ككل في بلد معين أو منطقة من البلد، مثلاً، ربما تُستبعد بعض وحدات المعاينة الأولية عن قصد في مرحلة التصميم لأن بعض مناطق البلد يتعذر الوصول إليها بسبب حرب أهلية أو اضطرابات، أو كارثة طبيعية، أو عوامل أخرى. كذلك تُستبعد المناطق النائية التي يكون عدد الأسر أو الأشخاص فيها قليلاً جداً، لأن تغطيتها تكون باهظة التكلفة. ولأنها تمثل نسبة صغيرة من السكان يكون أثرها على أرقام السكان صغيراً جداً (انظر الفصل الرابع للاطلاع على مزيد من المناقشة، بما في ذلك أمثلة متعددة لعدم تغطية وحدات المعاينة الأولية في استقصاءات الأسر المعيشية). وينبغي الإشارة إلى استبعاد هذه المناطق إشارة واضحة في التقرير المقدم عن نتائج الاستقصاء الذي من هذا القبيل. ولا ينبغي أن يُترك انطباع بأن نتائج الاستقصاء تنطبق على البلد بأسره أو المنطقة بأسرها، بينما تُركت في الواقع نسبة من السكان دون تغطية. ويجب أن يُشار إشارة تامة في تقرير الاستقصاء إلى صفات عدم التغطية في الاستقصاء.

٤٧ - يصبح عدم التغطية مشكلة أكثر خطورة على مستوى الأسرة المعيشية. معظم الاستقصاءات تعتبر الأسرة المعيشية مجموعة من الأشخاص الذين تربطهم في العادة صلة قرى من نوع ما ويقومون في العادة في مسكن أو وحدة سكنية. توجد هنا مسائل هامة تتعلق بالتعريف ينبغي حلها، مثل من هو الشخص الذي يعتبر مقيماً عادياً، وما الذي يشكل وحدة سكنية. كيف يمكن تناول المباني المتعددة الوحدات (كالمباني المقسمة إلى شقق)

والوحدات السكنية التي فيها بضع أسر؟ ربما يكون من السهل تعريف الوحدة السكنية، لكن الهياكل الاجتماعية المعقدة تجعل من الصعب تعريف الأسر المعيشية داخل الوحدة السكنية. لذلك توجد إمكانية كبيرة لسوء التفسير أو التفسير غير المتسق في هذه المفاهيم لدى المقابِلين المختلفين، أو في البلدان أو الثقافات المختلفة. على أي حال، يلزم توفير تعليمات تشغيلية صارمة لإرشاد المقابِلين حول من ينبغي اعتباره فرداً في أسرة معيشية وحول ما الذي ينبغي اعتباره وحدة سكنية.

٤٨ - من العوامل الأخرى التي تساهم في عدم التغطية حذف وحدات سكنية عن غير قصد من القائمة التي أُعدت أثناء العمليات الميدانية، أو حذف مجموعات سكانية فرعية هي موضع اهتمام (كالأطفال الصغار أو المسنين، مثلاً)، أو الحذف الناتج عن أخطاء في القياس، أو عدم إدراج أفراد غائبين عن الأسرة، وحذف ناتج عن عدم فهم مفاهيم الاستقصاء فهماً صحيحاً. ويوجد أيضاً بُعدٌ زمني للمشكلة، بمعنى أن الوحدات السكنية ربما كانت غير مشغولة أو قيد البناء عندما وضعت القوائم، لكنها ربما تصبح مشغولة وقت جمع البيانات. في استقصاءات الأسر المعيشية في البلدان النامية، تتفاقم مشكلة عدم التغطية لأن معظم عمليات التعداد، التي هي الأساس الفريد لبناء أطر المعاينة، لا تعطي عناوين مفصلة لوحدات المعاينة على مستوى الأسرة المعيشية أو على مستوى الأشخاص. وفي كثير من الأحيان، تُستخدم قوائم إدارية للأسر تكون غير دقيقة أو فات أو أنها، وتُحذف أسماء أفراد في الأسر عن قصد أو غير قصد من قائمة المقيمين في الأسرة. ترد تفاصيل أكثر عن مصادر عدم التغطية في ليبكوفسكي (٢٠٠٣) والمراجع المشار إليها فيه.

## ٦ - ٦ - ٢ التعويض عن عدم التغطية في استقصاءات الأسر المعيشية

٤٩ - توجد بضعة نُهجٍ لمعالجة مشكلة عدم التغطية في استقصاءات الأسر المعيشية (ليبكوفسكي، ٢٠٠٣). تشمل هذه النُهج:

(أ) إجراءات ميدانية محسنة كاستعمال الأطر المتعددة، مثلاً، وإجراءات محسنة لإعداد القوائم؛

(ب) التعويض عن عدم التغطية بواسطة التعديل الإحصائي لمُعامِلات الترجيح.

٥٠ - فيما يتعلق بالنهج الثاني، إذا وجدت مجاميع رقابة موثوقة للسكان ككل، ولمجموعات فرعية محددة من السكان، يمكن محاولة تعديل مُعامِلات الترجيح لوحدات العيّنة بطريقة تجعل مجموع مُعامِلات الترجيح تتفق مع مجاميع الرقابة في المجموعات الفرعية المحددة. المجموعات الفرعية تُسمى "ما بعد الطبقات" (*post-strata*)، وإجراء التعديل الإحصائي يُسمى "ما بعد التقسيم الطبقي" (*post-stratification*). هذا الإجراء يعوض عن عدم التغطية بتعديل التوزيع المرجح في العيّنة لمنغيرات محددة، لجعلها تتفق مع التوزيع المعروف للسكان (انظر ليتونين وباخينين (١٩٩٥) للاطلاع على أمثلة عملية لكيفية تحليل بيانات الاستقصاء بعد التقسيم الطبقي). يرد أدناه توضيح بسيط.

### مثال

افترض أن عدد الأسر المعيشية المذكورة في المثال السابق معروف من مصدر مستقل كالسجل المدني الراهن، مثلاً، بأنه يساوي ٤٥ ٠٢٥ أسرة في الشمال و ١١٥ ٨٠٠ أسرة في الجنوب. وافترض كذلك أن المجاميع المرجحة في العيّنة هي ٤٠ ٠٠٠ أسرة في الشمال و ١٢٠ ٠٤٠ أسرة في الجنوب. قم بإجراء الخطوات التاليتين:

- الخطوة الأولى: احسب عوامل ما بعد التقسيم الطبقي.

في المنطقة الشمالية، يوجد لدينا:  $w_{3h} = \frac{45\,025}{40\,000} = 1.126$  و

في المنطقة الجنوبية يوجد لدينا:  $w_{3h} = \frac{115\,800}{120\,040} = 0.965$

- الخطوة الثانية: احسب مُعامل الترحيح المعدل النهائي:  $w_f = w_b \times w_{3h}$ .

يرد موجز النتائج الرقمية في الجدول ٦ - ٤.

الجدول ٦ - ٤

الترجيح فيما بعد التقسيم الطبقي لتعديل التغطية

الطبقة	$r_b$	$t_b$	$w_b$	$w_{fb}$	$w_{fb} r_b$	$w_{fb} t_b$
الشمال - حضرية	٨٠	٧٠	١٢٥	١٤٠,٧٥	١١ ٢٦٠	٩ ٨٥٢
الشمال - ريفية	١٢٠	١٠٠	٢٥٠	٢٨١,٤٠	٣٣ ٧٦٨	٢٨ ١٤٠
الجنوب - حضرية	١٧٠	١٥٠	٢٣٦	٢٢٧,٧٧	٣٨ ٧٢١	٣٤ ١٦٦
الجنوب - ريفية	٣٦٠	١٨٠	٢٢٢	٢١٤,٢٠	٧٧ ١١٢	٣٨ ٥٥٦
المجموع	٧٣٠	٥٠٠			١٦٠ ٨٦١	١١٠ ٧١٤

النسبة المقدرة للأسر المعيشية التي يمكنها الحصول على العناية الصحية الأولية هي:

$$\hat{P}_f = \frac{\sum w_{fb} t_b}{\sum w_{fb} r_b} = \frac{110\,714}{160\,861} = 0.69 \text{ أو } 69\%$$

٥١ - لاحظ أنه بتعديل مُعاملات الترحيح فيما بعد التقسيم الطبقي، تكون أعداد العيّنة المرجحة للمنطقتين الشمالية والجنوبية، على التوالي،  $٤٥٠٢٤$  ( $٣٣٧٦٨ + ١١٢٥٦$ ) و  $١١٥٨٢١$  ( $٣٨٧٠٩ + ٧٧١١٢$ )، وهي أرقام قريبة جداً من مجاميع الرقابة المستقلة المذكورة أعلاه.

## ٦ - ٧ زيادة التباين في العيّنة نتيجة للترجيح

٥٢ - حتى وإن كان استخدام مُعاملات الترحيح في تحليل بيانات الاستقصاء يميل إلى تخفيض التحيز في التقديرات، يمكن أيضاً أن يضحّم أوجه التباين في هذه التقديرات. ولتبسيط المناقشة، ننظر في تصميم ذي مرحلة واحدة مقسّم إلى طبقات بعيّنات متساوية الاحتمالات داخل الطبقات. فإذا لم تكن تباينات الطبقة (بمعنى التباينات بين الوحدات في الطبقات) متساوية في كل طبقة فإن استخدام مُعاملات ترحيح غير متساوية في الطبقة في كل الطبقات (مثل ذلك كون مُعاملات الترحيح متناسبة تناسباً عكسياً مع تباينات الطبقة) يمكن أن يؤدي إلى تقديرات للاستقصاء أكثر دقة. غير أنه إذا كانت تباينات الطبقة متساوية في كل طبقة، فإن مُعاملات الترحيح غير المتساوية ستؤدي إلى تباينات أعلى في تقديرات الاستقصاء مما ينتج عن استخدام مُعاملات ترحيح متساوية.

٥٣ - أثر استخدام مُعامِلات الترجيح لزيادة التباين في متوسط السكان المقدّر بموجب العامل

(١٢ - ٦)

$$L = n \times \frac{\sum_b n_b w_b^2}{(\sum_b n_b w_b)^2}$$

حيث  $n = \sum_b n_b$  هو مجموع حجم العيّنة المحقّق، و

$w_b$  هو مُعامِل الترجيح النهائي، و  $n_b$  هو حجم العيّنة المحقّق للطبقة  $b$ . يمكن أيضاً كتابة المعادلة الواردة أعلاه بلغة مكافئ تباين مُعامِلات الترجيح، كما يلي:

(١٣ - ٦)

$$L = n \times \frac{\sum_j w_j^2}{(\sum_j w_j)^2} = 1 + CV^2(w_j)$$

$$CV^2(w_j) = \frac{n}{(\sum_j w_j)^2} \left\{ \sum_j w_j^2 - \frac{1}{n} (\sum_j w_j)^2 \right\} = \frac{\text{تباين معاملات الترجيح}}{(\text{متوسط معاملات الترجيح})^2}$$

حيث

### مثال

الآن نحسب عامل تضخم التباين باستخدام البيانات الواردة في المثال الوارد في الفرع ٦ - ٦ - ٢، مع مُعامِلات الترجيح النهائية  $w_{jb}$  وأحجام العيّنة الطبقيّة المحددة  $r_b$  (انظر الجدول ٦ - ٤)

الجدول ٦ - ٥

### مُعامِلات الطبقة للتباين

الطبقة	$r_b$	$w_{jb}$	$w_{jb} r_b$	$w_{jb}^2 r_b$
الشمال - حضرية	٨٠	١٤٠,٧٥	١١ ٢٦٠	١ ٥٨٤ ٨٤٥
الشمال - ريفية	١٢٠	٢٨١,٤٠	٣٣ ٧٦٨	٩ ٥٠٢ ٣١٥
الجنوب - حضرية	١٧٠	٢٢٧,٧٧	٣٨ ٧٢١	٨ ٨١٩ ٤٥٩
الجنوب - ريفية	٣٦٠	٢١٤,٢٠	٧٧ ١١٢	١٦ ٥١٧ ٣٩٠
المجموع	٧٣٠		١٦٠ ٨٦١	٣٦ ٤٢٤ ٠٠٩

$$L = 730 \times \frac{36\,424\,009}{(160\,861)^2} = 1.03$$

لذلك،

بعبارة أخرى، توجد زيادة في تقديرات الاستقصاء مقدارها نحو ٣ في المائة نتيجة لاستخدام مُعامِلات الترجيح.

## ٨ - ٦ تخفيض مُعامِلات التّرجيح

٥٤ - عندما تُحسب مُعامِلات التّرجيح وتعدّل لتعويض النواقص التي نوقشت أعلاه، يُنصَحُ بتفحص توزيع مُعامِلات التّرجيح المعدلة. وإن مُعامِلات التّرجيح المفرطة في الكبر، حتى وإن كانت تؤثر في نسبة صغيرة فقط من الحالات المدرجة في العيّنة، يمكن أن تسفر عن زيادة كبيرة في تباين تقديرات الاستقصاء. لذلك جرت العادة بتخفيض مُعامِلات التّرجيح المفرطة إلى قيمة لا يجوز أن تتجاوزها بغية الحد من التباين المرتبط بمُعامِلات التّرجيح (وبذلك تخفيض تباين تقديرات الاستقصاء) وفي الوقت نفسه تحول دون تمكّن عدد صغير من الوحدات المدرجة في العيّنة من التحكم بالتقدير الإجمالي. ويُستعمل تخفيض مُعامِلات التّرجيح في أحيان كثيرة بعد تعديل مُعامِلات التّرجيح لتعويض عدم الاستجابة.

٥٥ - بينما يميل تخفيض مُعامِلات التّرجيح إلى تقليل تباين التقديرات، فهو يُدخل تحيزاً في التقديرات أيضاً. وفي بعض الحالات، ربما يتجاوز تخفيض التباين، الناتج عن تخفيض مُعامِلات التّرجيح المفرطة في الكبر، مقابلة الزيادة في التحيز الحاصل، وبذلك يخفض الخطأ البالغ مربع المتوسط في تقديرات الاستقصاء. ولا ينبغي في الواقع تخفيض مُعامِل التّرجيح إلاّ عندما يكون له ما يبرره، أي عندما يمكن التحقق من أن للتحيز المدخل نتيجةً لاستخدام مُعامِلات تّرجيح مخفضة (خلافاً لمُعامِلات التّرجيح الأصلية) أثر على مجموع الخطأ المساوي لمربع المتوسط أقل من التخفيض المقابل في التباين الذي تحقق بالتخفيض.

٥٦ - لأنيّ تصميم طبقي، ينبغي تنفيذ عملية تخفيض مُعامِل التّرجيح في الأحوال النمطية داخل كل طبقة. تبدأ العملية بتحديد الحد الأعلى لمُعامِلات التّرجيح الأصلية ثم تُعدّل مُعامِلات التّرجيح كلّها لكي يكون مجموع مُعامِلات التّرجيح المخفضة مساوياً لمُعامِلات التّرجيح الأصلية. دع  $w_{hi}$  يشير إلى مُعامِل التّرجيح النهائي للوحدة رقم  $i^{th}$  في طبقة رقمها  $h$ ، ودع  $w_{hB}$  يشير على الحد الأعلى لمُعامِلات التّرجيح المحددة للطبقة رقم  $h$ . عندئذٍ، يمكن تعريف مُعامِل التّرجيح المخفض للوحدة رقم  $i^{th}$  المدرجة في العيّنة في الطبقة  $h$ ، على النحو التالي:

(١٤ - ٦)

$$w_{hi(T)} = \begin{cases} w_{hi} & \text{إن } w_{hi} < w_{hB} \\ w_{hB} & \text{إن } w_{hi} \geq w_{hB} \end{cases}$$

٥٧ - يمكن الآن تعديل مُعامِلات التّرجيح المخفضة للعيّنة بأكملها، مرة أخرى، لكي يكون مجموعها مساوياً بالضبط لمجموع مُعامِلات التّرجيح الأصلية. لتسهيل التوضيح، سنفترض أن مُعامِلات التّرجيح ثابتة داخل الطبقات، وأسقط الرمز السفلي للدلي  $i$  لما تبقى من هذا النقاش. دع  $F_T$  يشير إلى نسبة مجموع مُعامِلات التّرجيح الأصلية إلى مجموع مُعامِلات التّرجيح المخفضة، بعبارة أخرى،

(١٥ - ٦)

$$F_T = \frac{\sum_h n_h w_h}{\sum_h n_h w_{h(T)}}$$

حيثُ تؤخذ مجاميع النسب عبر الطبقات كلها ومن ثم على جميع الوحدات المدرجة في العيّنة. وإذا عرّفنا مُعامِل التّرجيح المخفض المعدّل للطبقة رقم  $h$  كما يلي:

(١٦ - ٦)

$$w_{h(T)}^* = F_T w_{h(T)}$$

من الواضح إذن أن  $\sum_b n_b w_b^* = \sum_b n_b w_b$  ، كما هو مرغوب فيه.

المثال المعروض أدناه مصمّم لتوضيح إجراء التخفيض والمساعدة على فهمها.

٥٨ - العمودان الأولان من الجدول ٦ - ٦ أدناه يُعطيان مجموع عدد الوحدات ومُعاملِ الترجيح النهائي على التوالي، في كل طبقة من ٧ طبقات. ويُختار مُعاملِ ترجيح أقصاه ٢٥٠، ولذلك تُخفّض كل مُعاملات الترجيح الأصلية إلى ٢٥٠، كما هو مبين في العمود الثالث من الجدول.

#### الجدول ٦ - ٦ تخفيض مُعاملات الترجيح

$n_b w_b^*$	$n_b w_b(T)$	$n_b w_b$	$w_b(T)$	$w_b$	$n_b$
١١ ٨٢٣ ٠٠	١١ ٢٦٠	١١ ٢٦٠	١٤٠,٧٥	١٤٠,٧٥	٨٠
١٥ ٧٧٦ ٢٥	١٥ ٠٢٥	١٥ ٠٢٥	١٥٠,٢٥	١٥٠,٢٥	١٠٠
٢٢ ٩٦٨ ٧٥	٢١ ٨٧٥	٢١ ٨٧٥	١٧٥,٠٠	١٧٥,٠٠	١٢٥
٣١ ٥٠٠ ٠٠	٣٠ ٠٠٠	٣٠ ٠٠٠	٢٠٠,٠٠	٢٠٠,٠٠	١٥٠
٣١ ٥٠٠ ٠٠	٣٠ ٠٠٠	٣٠ ٠٠٠	٢٥٠,٠٠	٢٥٠,٠٠	١٢٠
٣١ ٥٠٠ ٠٠	٣٠ ٠٠٠	٣٣ ٠١٥	٢٥٠,٠٠	٢٧٥,١٣	١٢٠
٤٤ ٦٢٥ ٠٠	٤٢ ٥٠٠	٤٨ ٥١٨	٢٥٠,٠٠	٢٨٥,٤٠	١٧٠
١٨٩ ٦٩٣ ٠٠	١٨٠ ٦٦٠	١٨٩ ٦٩٣			٨٦٥
					المجموع

لاحظ أنه في هذه الحالة،

$$F_T = \frac{\sum_i n_b w_{bi}}{\sum_i n_b w_{bi(T)}} = \frac{189\ 693}{180\ 660} = 1.05$$

أعيد تحديد حجم مُعاملات الترجيح المخفضة لكي يساوي مجموعها المجموع الأصلي ١٢٥، بضرب كل مُعاملِ ترجيح بـ  $F_T = 1.05$ .

#### ٩ - ٦ ملاحظات ختامية

٥٩ - أصبحت مُعاملات ترجيح العيّنة الآن تعتبر جزءاً لا يتجزأ من تحليل بيانات استقصاء الأسر المعيشية في البلدان النامية، كما هي في سائر أنحاء العالم. والآن تدعو معظم برامج الاستقصاء إلى استخدام مُعاملات الترجيح حتى في الحالات النادرة التي تنطوي على عيّنات ذاتية الترجيح (في هذه الحالة تكون مُعاملات الترجيح ١). بذل مصمّمو الاستقصاءات جهوداً هائلة في الماضي للوصول إلى هدف يكاد يكون الوصول إليه مستحيلًا، وهو تحقيق عيّنات ذاتية الترجيح، ومن ثم جعل مُعاملات الترجيح غير لازمة في تحليل بيانات الاستقصاء. وكانت الحكمة العرفية أن استخدام مُعاملات الترجيح يجعل التحليلات معقدة جداً وأنه لم يكن يوجد شيء يُذكر، إن وُجدَ على الإطلاق، من البنى التحتية الحاسوبية لإجراء التحليلات المرجحة. غير أن ما تحقّق من تقدم في تكنولوجيا الحاسوب خلال العقد الماضي قد جعل هذه الحجة غير صحيحة. فأجهزة الحاسوب والبرامج الحاسوبية أصبحت

الآن في متناول اليد في كثير من البلدان النامية. يُضاف إلى ذلك أن كثيراً من البرامج الحاسوبية أصبحت الآن متاحة خصيصاً لتحليل بيانات الاستقصاء. سنستعرض هذه البيانات ونُقارن في الفصل السابع.

٦٠- إن استخدام مُعاملات الترجيح، كما سبقت مناقشته، يقلل التحيز نتيجةً لنواقص العيّنات المتصلة بعدم التغطية وبعدم الاستجابة. ومما يُذكر أن عدم الاستجابة وعدم التغطية نوعان مختلفان من أنواع الخطأ نتيجةً لعجز الاستقصاء المصمّم عن الحصول على معلومات من بعض الوحدات في المجموعة السكانية المستهدفة. في استقصاءات الأسر المعيشية في البلدان النامية، يشكّل عدم التغطية مشكلةً أخطر من عدم الاستجابة. أعطى هذا الفصل أمثلة لإجراءات لتطوير مُعاملات ترجيح أساسية وتعديلها إحصائياً للتعويض عن بعض المشاكل المتصلة باستقصاءات الأسر المعيشية واستخدام المُعاملات الترجيحية المعدلة في تقدير المُعلّات التي هي موضع اهتمام. وإن مجيء الحواسيب الفائقة السرعة وبرامج إحصائية حاسوبية في متناول اليد أو بالمجان، من شأنها أن تجعل استخدام مُعاملات الترجيح جانباً روتينياً من جوانب تحليل بيانات استقصاء الأسر المعيشية حتى في البلدان النامية. غير أن تطوير مُعاملات ترجيح للعيّنات، كما أظهر هذا الفصل، يزيد تعقيد عمليات الاستقصاء بطرق مختلفة. فمُعاملات الترجيح، مثلاً، تحتاج إلى حساب في كل مرحلة من مراحل اختيار العيّنات؛ ثم تحتاج إلى تعديل لتعويض مختلف نواقص العيّنة؛ وأخيراً، تحتاج إلى تخزينها واستعمالها بشكل ملائم في جميع التحليلات اللاحقة. ومن ثمّ، يجب الانتباه بعناية لتطوير عمليات الترجيح والحساب الفعلي لمُعاملات الترجيح التي سنستخدم في تحليل الاستقصاء.

### المراجع ومواد أخرى للقراءة

- Brick, J.M., and G. Kalton, (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol. 5, pp. 215-238.
- Cochran, W.G. (1977). *Sampling Techniques*, 3rd ed. New York: John Wiley & Sons.
- Groves, R.M., and M.P. Couper, (1998). *Non-response in Household Interview Surveys*. New York: John Wiley & Sons.
- Groves, R.M., and others (2002). *Survey Non-response*. New York: , John Wiley & Sons.
- D. Kasprzyk, (1986). The treatment of missing survey data. *Survey Methodology*, vol. 12, pp. 1-16.
- Kalton, G. (1983). *Compensating for Missing Survey Data*. Ann Arbor, Michigan: Survey Research Center, University of Michigan.
- Kish, L. (1965), *Survey Sampling*, New York: Wiley.
- I. Hess, (1950). On non-coverage of sample dwellings, *Journal of the American Statistical Association*, vol. 53, pp. 509-524.
- Lehtonen, R., and E.J. Pahkinen. (1995). *Practical Methods for Design and Analysis of Complex Surveys*, New York: Wiley.
- Lepkowski, James. (2005), *Non-observation error in household surveys in developing countries*. In Household Sample surveys in Developing and Transition Countries. Studies in Methods, No. 96. Sales No. E.05. XVII.6.
- Lessler, J. and W. Kalsbeek, (1992). *Nonsampling Error in Surveys*, , New York: John Wiley & Sons.

Levy, P. S., and S. Lemeshow. (1999), *Sampling of Populations: Methods and Applications*, 3<sup>rd</sup> ed. New York: John Wiley & Sons.

Lohr, S. (1999), *Sampling: Design and Analysis*, Duxbury Press, Pacific Grove.

Yansaneh, I. S. (2004), *Overview of Sample Design Issues for Household Surveys in Developing Countries*, in *Household Surveys in Developing and Transition Countries: Technical Report*, United Nations, New York.



## الفصل السابع

# تقدير أخطاء المعاينة

## لبيانات الاستقصاء

### ٧ - ١ مقدمة

١ - يلقي هذا الفصل نظرة عامة موجزة على مختلف الطرق المستخدمة لتقدير أخطاء المعاينة لبيانات استقصاء الأسر المعيشية، التي تولدها مختلف تصاميم العينات، التي تتراوح من التصاميم المعيارية التي يمكن أن توجد في أي كتاب مدرسي تمهيدي معني بنظرية المعاينة (مثل كوكران (١٩٧٧))، إلى التصاميم الأكثر تعقيداً التي تُستخدَم لاستقصاءات الأسر المعيشية الواسعة النطاق. لتصاميم العينات المعيارية تُقدّم المعادلات إلى جانب الأمثلة الرقمية لتوضيح تقدير أخطاء المعاينة، وإقامة فترات ثقة، وحساب آثار التصميم وأحجام العينات الفعلية. ثم تُعرض طرق تقدير أخطاء المعاينة للتصاميم الأكثر تعقيداً. ثم تناقش فضائل ونقائص كل طريقة وتُعرض أمثلة رقمية لتوضيح تنفيذ الإجراءات. وسيقدّم مثال، بناءً على بيانات من استقصاء حقيقي، لتوضيح حقيقة أن البرامج الإحصائية المعيارية تقدر أخطاء المعاينة في تقديرات الاستقصاء، ممّا يؤدي إلى استنتاجات خاطئة عن المَعْلَمَات التي هي موضع اهتمام الاستقصاء. ولتجنّب هذه المشكلة، يوصي الفصل بقوة باستخدام برامج إحصائية حاسوبية خاصة تأخذ في الحسبان التام الطبيعة المركّبة للتصاميم التي تُستخدم عموماً في استقصاءات الأسر المعيشية. وترد توصيفات لبضعة برامج حاسوبية من هذا القبيل ومقارنتها.

### ٧ - ١ - ١ تقدير أخطاء المعاينة لبيانات الاستقصاء المركّب

٢ - تقدمت في الآونة الأخيرة الأهداف التحليلية لاستقصاءات الأسر المعيشية المصمّمة جيداً فتجاوزت الجداول الموجزة الأساسية الشاملة لأعداد أو مجاميع المَعْلَمَات التي هي موضع اهتمام. والمحللون مهتمون الآن أيضاً بتوليد الفرضيات واختبار بناء النماذج. فبدلاً من مجرد تقدير نسبة السكان الذين يعيشون في الفقر أو الحاصلين على شهادة الدراسة الثانوية أو أعلى منها، يريد المحلّلون الآن أن يقيّموا أثر السياسات، أو أن يستطلعوا الطريقة التي بها يتأثر متغير في استجابة رئيسية، أو الأداء الأكاديمي لطفل في المدرسة، مثلاً، أو مستوى الفقر في أسرة، بعوامل مثل المنطقة، والحالة الاجتماعية - الاقتصادية، ونوع الجنس، والسن.

٣ - تتطلب الإجابة على هذا النوع من الأسئلة تحليل بيانات على مستوى الأسرة المعيشية أو الشخص. ويجب أن تشمل نشرة نتائج هذه التحليل، بالضرورة، مقاسات ملائمة لإحكام أو دقة التقديرات المستمدة من بيانات الاستقصاء. والمعلومات عن إحكام تقديرات الاستقصاء مطلوبة لاستخدام نتائج الاستقصاء استخداماً صحيحاً وتنفيذها، وكذلك لتقييم وتحسين تصاميم العينات وإجراءاتها. وإن مراقبة تصاميم العينة وتقييمها على جانب

كبير من الأهمية في حالة برامج الاستقصاء الوطنية الكبيرة، التي تُصمّم في كثير من الأحيان لتكون المصدر الوحيد للمعلومات المفصلة عن مواضيع متنوعة تنوعاً كبيراً.

٤ - وإن واحداً من المقاسات الرئيسية للإحكام في الاستقصاءات بالعيّنة هو التباين في العيّنة (هذا المفهوم في الفصل الثالث)، وهو مؤشر للتباين قُدم باختيار عيّنة بدلاً من تعداد جميع السكان، على افتراض أن المعلومات التي جُمعت في الاستقصاء صحيحة. وتباين العيّنت مقياس لتنوع توزيع العيّنت بواسطة التقدير. ويُستخدم الخطأ المعياري، أو الجذر التربيعي للتباين لقياس خطأ العيّنة. ولأيّ استقصاء بعينه، يمكن تقييم مقدّر خطأ المعاينة هذا واستخدامه لبيان دقة التقديرات.

٥ - يتوقف شكل تقدير التباين، وكيفية تقييمه، على تصميم العيّنة التي يقوم عليها. ففي التصاميم المعيارية تُقيّم هذه المقدّرات في الغالب باستخدام معادلات بسيطة. غير أنه في تصاميم العيّنت المركّبة المستخدمة في استقصاء الأسر المعيشية، التي تنطوي في الغالب على التقسيم الطبقي وتكوين المجموعات والمعاينة ذات الاحتمالات غير المتساوية غالباً ما تكون أشكال هذه المقدّرات معقدة ومن الصعب تقييمها. ويتطلب حساب أخطاء المعاينة في هذه الحالة إجراءات تأخذ في الحسبان تعقيد تصميم العيّنة التي ولدت البيانات، وهذه بدورها تتطلب في الغالب برنامجاً حاسوبياً ملائماً.

٦ - في كثير من البلدان النامية، يكون تحليل بيانات استطلاع الأسر المعيشية في كثير من الأحيان محصوراً بتحليل الجداول الأساسية، مع تقدير المتوسطات والنسب والمجاميع، لكن لا توجد أيّ إشارة إلى إحكام هذه التقديرات أو دقتها. حتى في مكاتب الإحصاء الوطنية، التي توجد فيها بنية تحتية واسعة النطاق لجمع البيانات الإحصائية وتجهيزها، غالباً ما يجد المرء قلةً في الدراية بالتحليل المفصل للبيانات على المستوى الجزئي. وغالباً ما يُفاجأ مصمّمو الاستقصاءات أو المطلون عندما يعرفون، مثلاً، أن جميع العناصر يُدخل أوجّه ترابط بين العناصر تقلل دقة التقديرات المتصلة بالعيّنات العشوائية البسيطة، التي هم معتادون على تحليلها؛ أو أن استخدام مُعاملات الترجيح في التحليل يزيد أخطاء المعاينة بوجه عام؛ أو أن برامج الحاسوب العادية التي يستخدمونها بشكل روتيني في عملهم لا تبين بشكل ملائم هذه النواقص في الدقة.

٧ - يحاول هذا الفصل أن يصلح هذا الوضع بإلقاء نظرة عامة على طرق حساب تقديرات خطأ المعاينة في أنواع التصاميم المركّبة التي تُستخدم عادةً في استقصاءات الأسر المعيشية في البلدان النامية، وكذلك برامج الحاسوب الإحصائية المستخدمة في تحليل هذه الاستقصاءات. وتُقدّم بضعة أمثلة عديدة لتوضيح إجراءات التباين التي نوقشت.

## ٧ - ١ - ٢ نظرة عامة

٨ - يعطي الفرع ٧ - ٢ تعريفاً بحسب المبدأ الأول لتباين المعاينة بموجب المعاينة العشوائية البسيطة، بما في ذلك أمثلة عديدة توضح حساب تباين المعاينة وبناء فترات الثقة. وترد تعاريف القياسات الأخرى لأخطاء المعاينة في الفرع ٧ - ٣. ويعطي الفرع ٧ - ٤ معادلات لحساب تباين المعاينة بموجب مختلف التصاميم العادية، مثل المعاينة الطبقيّة ومعاينة المجموعات. وتُقدّم بضعة أمثلة رقمية لتسهيل فهم المفاهيم. ويناقش الفرع ٧ - ٥ خصائص مشتركة في تصاميم استقصاءات الأسر المعيشية، وكذلك مضمون وهيكل بيانات الاستقصاء اللازمة لتقدير ملائم لأخطاء المعاينة. وتُقدّم أيضاً الاستمارة العامة للتقديرات التي هي موضع اهتمام في استقصاءات الأسر المعيشية. ويعطي الفرع ٧ - ٦ مبادئ توجيهية موجزة بشأن عرض معلومات عن أخطاء المعاينة؛ ويصف الفرع ٧ - ٧ طرق عملية لحساب أخطاء المعاينة في تصاميم أكثر تعقيداً. غالباً ما تتطلب هذه الطرق إجراءات

خاصة واستخدام برامج حاسوبية متخصصة. وتناقش مزالق استخدام البرامج الحاسوبية الإحصائية العادية لتحليل بيانات الاستقصاء في الفرع ٧ - ٨، باستخدام مثال يستند إلى بيانات مستمدة من استقصاء للتغطية بالتحسين أُجري في بوروندي في سنة ١٩٨٩. وتُستعرض بعض البرامج الحاسوبية المتاحة للعموم لتقدير أخطاء المعاينة في بيانات استقصاءات الأسر المعيشية وتُقارَن في الفرعين ٧ - ٩ و ٧ - ١٠. وينتهي الفصل ببعض الملاحظات الختامية في الفرع ٧ - ١١.

## ٧ - ٢ تباين العينات بموجب المعاينة العشوائية البسيطة

٩ - يمكن تعريف تباين التقدير في المعاينة بأنه معدل الانحراف التربيعي عن معدل قيمة التقدير، حيث يؤخذ المعدل لجميع العينات الممكنة. والمعاينة العشوائية البسيطة، كما أُشير في الفصل الثالث، هي أبسط تقنيات المعاينة الأولية، لكنها نادراً ما تُستخدم في الاستقصاءات الكبيرة لأن تنفيذها ذو كفاءة قليلة جداً وهو باهظ التكلفة.

١٠ - لتسهيل فهم مفهوم تباين المعاينة، ننظر في مجموعة سكانية صغيرة قوامها خمس أسر معيشية ( $N=5$ )، تُختار منها عينة صغيرة مؤلفة من أسرتين معيشيتين (الحجم  $n=2$ ) بمعاينة عشوائية بسيطة بدون تبديل. افترض أن المتغير الذي هو موضع اهتمام هو النفقات الشهرية للأسرة على الغذاء وأن نفقات كل واحدة من الأسر الأربعة الأخرى هي كما ترد في الجدول ٧ - ١ أدناه.

الجدول ٧ - ١

النفقات الشهرية بالدولار على الغذاء لكل أسرة معيشية

الأسرة (١)	النفقات على الغذاء بالدولار ( $Y_1$ )
١	١٠
٢	٢٠
٣	٣٠
٤	٤٠
٥	٥٠

١١ - لاحظ، أولاً، أنه لأننا نعلم قيمة المتغير الذي هو موضع اهتمام لجميع الأسر المعيشية في مجموعتنا السكانية، نستطيع حساب قيمة المعلمة المتصل بمتوسط النفقات الشهرية للأسرة المعيشية على الغذاء، وهو

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i = \frac{10 + 20 + 30 + 40 + 50}{5} = \frac{150}{5} = 30$$

وإن تقدير المعاينة العشوائية البسيطة بلا تبديل لمعدل النفقات الشهرية على الغذاء هو

$$\hat{Y} = \frac{1}{2} \sum_{i \in S} Y_i$$

حيث يؤخذ الجمع من الوحدات المختارة لإدراجها في العينة. من الواضح أن التقدير الذي حُصلَ عليه يتوقف على العينة المختارة. يرينا الجدول ٧ - ٢ الوارد أدناه جميع العينات الممكنة، والتقدير الذي يستند إلى كل عينة، والانحرافات في تقدير كل عينة عن متوسط المجموعة السكانية، ومربع الانحرافات. لاحظ أن  $\hat{Y}_{ave}$  يشير إلى معدل جميع التقديرات بناءً على أساس العينات. لاحظ أن  $\bar{Y}$  هو رمز لمتوسط السكان، بينما  $\hat{Y}$  هو رمز لتقدير متوسط السكان المسمى متوسط العينة (انظر الجدول ٧ - ٢).

الجدول ٧ - ٢

حساب تباين المعاينة الحقيقي لـ  $\hat{Y}$ ، مَعْلَمَة المعدّل

العيّنة ١	وحدات العيّنة	تقدير العيّنة ( $\hat{Y}_i$ )	$\hat{Y}_i - \hat{Y}_{ave}$	$(\hat{Y}_i - \hat{Y}_{ave})^2$
١	(١,٢)	١٥	١٥ -	٢٢٥
٢	(١,٣)	٢٠	١٠ -	١٠٠
٣	(١,٤)	٢٥	٥ -	٢٥
٤	(١,٥)	٣٠	صفر	صفر
٥	(٢,٣)	٢٥	٥ -	٢٥
٦	(٢,٤)	٣٠	صفر	صفر
٧	(٢,٥)	٣٥	٥	٢٥
٨	(٣,٤)	٣٥	٥	٢٥
٩	(٣,٥)	٤٠	١٠	١٠٠
١٠	(٤,٥)	٤٥	١٥	٢٢٥
المعدّل		٣٠	صفر	٧٥٠

لاحظ أن معدل التقديرات بناءً على جميع العيّنات الممكنة هو:

$$\hat{Y}_{ave} = \frac{1}{10} \sum_{i=1}^{10} \hat{Y}_i = \frac{15+20+25+30+25+30+35+35+40+45}{10} = \frac{300}{10} = 30 = \bar{Y}$$

١٢ - بعبارة أخرى، معدل قيمة التقديرات لكل العيّنات الممكنة يساوي متوسط السكان. وإن تقديراً هذه صفاته يُوصف بأنه غير متحيز للمعلّمة التي يقدرها.

١٣ - تباين المعاينة الحقيقي للمعدل المقدر للنفقات الشهرية على الغذاء من المعاينة العشوائية البسيطة بلا تبديل للحجم  $n=2$  من هذه المجموعة السكانية هو

$$Var(\hat{Y}) = \frac{1}{10} \sum_{i=1}^{10} (\hat{Y}_i - \hat{Y}_{ave})^2 = \frac{750}{10} = 75$$

١٤ - تتبع مشكلة النهج الوارد أعلاه من حقيقة أنه ليس من العملي أن نُختار جميع العيّنات الممكنة من المجموعة السكانية. ففي الواقع العملي اختيرت عيّنة واحدة فقط، وقيم المجموعة السكانية غير معروفة. والنهج الأكثر عمليةً هو استخدام معادلات لحساب التباين. وهذه المعادلات موجودة لجميع تصاميم العيّنات العادية.

١٥ - بموجب المعاينة العشوائية البسيطة بدون تبديل، يُعطى تباين المعاينة للمتوسط المقدر ( $\hat{Y}$ )، بناءً على حجم العيّنة  $n$ ، في التعبير التالي

(١ - ٧)

$$Var(\hat{Y}) = \left(1 - \frac{n}{N}\right) \frac{\delta^2}{n}$$

$$\delta^2 = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N-1} \quad \text{حيث}$$

هو مقياس لتباين الخصائص التي هي موضع اهتمام (تباين السكان الذي مقداره  $Y$ ). الرمز  $\delta^2$  في العادة غير معروف، ويجب تقديره من العينة (انظر المعادلة ٧ - ٢ أدناه). يمكن أن نرى بوضوح من المعادلة الواردة أعلاه أن تباين المعاينة يتوقف على العوامل التالية:

(أ) التباين السكاني للخاصية التي هي موضع اهتمام؛

(ب) حجم المجموعة السكانية؛

(ج) حجم العينة؛

(د) تصميم العينة وطريقة التقدير.

١٦ - نسبة المجموعة السكانية المدرجة في العينة،  $m/N$ ، تُسمى كسر المعاينة (ويشار إليها بالرمز  $f$ )؛ والعامل  $[1-(m/N)]$ ، أو  $1-f$ ، الذي هو نسبة المجموعة السكانية غير المدرجة في العينة المتناهية لتصحيح المجموعة السكانية ( $fpc$ ). يمثل الرمز  $fpc$  التعديل الذي أُجري على الخطأ العادي في التقدير لحساب أن العينة اختيرت بلا تبديل من المجموعة السكانية المتناهية. لكن تجدر ملاحظة أنه عندما يكون كسر المعاينة صغيراً يمكن تجاهل الرمز  $fpc$  في الواقع العملي يمكن تجاهل  $fpc$  إذا كان لا يزيد عن ٥ في المائة (كوكران، ١٩٧٧).

١٧ - تشير المعادلة الواردة أعلاه إلى أن التباين في المعاينة يتناسب تناسباً عكسياً مع حجم العينة. وكلما زاد حجم العينة نقص التباين في المعاينة؛ وفي حالة تعداد السكان أو العد الكامل (حيث  $n=N$ )، لا يوجد تباين في المعاينة. لاحظ أن عدم الاستجابة يُنقص حجم العينة فعلياً وبذلك يزيد التباين في المعاينة.

١٨ - يمكن أن يُرى أن التقدير غير المتحيز للتباين في المعاينة للمتوسط المقدّر يُعطى على الشكل التالي

(٢ - ٧)

$$v(\hat{Y}) = (1 - \frac{n}{N}) \frac{s^2}{n}$$

حيث يكون  $s^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y})^2}{n-1}$  تقديراً للتباين السكاني،  $\delta^2$ ، بناءً على أساس العينة. ويُشار إليه بأنه تباين في العينة. وفترة الثقة البالغة ٩٥ في المائة لمتوسط السكان (انظر الفقرة ٣٠ من الفصل الثالث) تُعطى على الشكل التالي:

(٣ - ٧)

$$\hat{Y} \pm 1.96 \sqrt{v(\hat{Y})}$$

١٩ - فيما يتعلق بنسبة السكان، يُعطى التقدير المستند إلى العينة والتباين في المعاينة بالشكل التالي، على التوالي:

(٤ - ٧)

$$\hat{p} = \frac{\text{عدد الوحدات مع الخاصية}}{n}$$

(٥ - ٧)

$$v(\hat{p}) = (1 - \frac{n}{N}) \frac{\hat{p}(1-\hat{p})}{n-1} \text{ و}$$

٢٠ - يوجز الجدول ٧ - ٣ الوارد أدناه تقديرات مختلف الخصائص السكانية وتباينات التقديرات بحسب المعاينة العشوائية البسيطة بدون تبديل.

الجدول ٧ - ٣

التقديرات وتبايناتها لخصائص سكانية مختارة

المعلّمة	التقدير	تباين التقدير
متوسط السكان ( $\hat{Y}$ )	$\hat{Y} = \frac{1}{n} \sum_{i \in \text{العينة}} Y_i$	$v(\hat{Y}) = (1 - \frac{n}{N}) \frac{s^2}{n}$
مجموع السكان	$\hat{T} = N\hat{Y}$	$v(\hat{T}) = N^2 v(\hat{Y})$
نسبة السكان من هذه الفئة	$\hat{P} = \frac{\text{عدد الوحدات المدرجة في العينة حسب الفئة}}{n}$	$v(\hat{P}) = (1 - \frac{n}{N}) \frac{\hat{P}(1 - \hat{P})}{n-1}$

٢١ - بوجه العموم، تُعطى فترة الثقة البالغة  $(1 - \alpha)$  لمتوسط السكان بالشكل التالي:

(٦ - ٧)

التباين المقدّر للتقدير  $\pm z_{1-\alpha/2} \sqrt{\text{التقدير}}$

حيث  $z_{1-\alpha/2}$  هو النسبة المئوية البالغة  $(1 - \alpha/2)$ -th من التوزيع العادي المعياري.

٢٢ - يوضّح المثال التالي تقدير التباين في المعاينة بناءً على العينة المختارة.

### المثال ١

انظر في عينة عشوائية بسيطة مقدارها  $n = 20$  أسرة معيشية مأخوذة من مجموعة سكانية كبيرة مقدارها  $N = 20,000$  أسرة معيشية. البيانات المجموعة مقدمة في الجدول ٧ - ٤ أدناه، حيث المتغير  $Y$  يشير إلى النفقات الأسبوعية للأسرة على الغذاء والمتغير  $Z$  يبين إن كانت الأسرة تملك جهاز تلفزيون أم لا ( $z=1$  إن كانت تملك، وصفرًا إن كانت لا تملك).

الجدول ٧ - ٤

نفقات الأسرة الأسبوعية على الغذاء وملكية جهاز تلفزيون للأسرة المدرجة في العينة

الأسرة المعيشية (i)	$Y_i$	$Z_i$	i	$Y_i$	$Z_i$
١	٥	صفر	١١	٧	١
٢	١٠	١	١٢	٨	١
٣	٥	صفر	١٣	٩	١
٤	٩	١	١٤	١٠	١
٥	٥	١	١٥	٨	١
٦	٦	١	١٦	٨	صفر
٧	٧	صفر	١٧	٥	صفر
٨	١٥	١	١٨	٧	صفر
٩	١٢	١	١٩	١٢	١
١٠	٨	صفر	٢٠	٤	صفر
المجموع		١٦٠	١٢		

تقدير النفقات الشهرية للأسرة المتوسطة من السكان على الغذاء هو:

$$\hat{Y} = \frac{1}{20} \sum_{i=1}^{20} Y_i = \frac{5+10+\dots+12+4}{20} = \frac{160}{20} = 8$$

التباين المقدر للمتوسط المقدر هو:

$$v(\hat{Y}) = \left(1 - \frac{20}{20\,000}\right) \left\{ \frac{(5-8)^2 + (10-8)^2 + \dots + (12-8)^2 + (4-8)^2}{19} \right\} = 7.87$$

فترة الثقة البالغة ٩٥ في المائة للمتوسط السكاني هي:

$$8 \pm 1.96 \times \sqrt{7.87} = (2.50, 13.50)$$

تقدير مجموع النفقات الشهرية للأسرة من السكان على الغذاء هو:

$$\hat{Y} = N\hat{Y} = 20\,000 \times 8 = 160\,000$$

التباين المقدر للمجموع المقدر هو:

$$v(\hat{Y}) = 20\,000^2 \times 7.87 = 3\,148\,000\,000$$

فترة الثقة البالغة ٩٥ في المائة للمتوسط السكاني هي:

$$160\,000 \pm 1.96 \times \sqrt{3\,148\,000\,000} = (50\,030, 269\,970)$$

النسبة السكانية المقدرة من الأسر التي تملك جهاز تلفزيون هي:

$$\hat{P} = \frac{1}{20} \sum_{i=1}^{20} Z_i = \frac{12}{20} = 0.6$$

التباين المقدر للنسبة المقدرة من الأسر التي تملك جهاز تلفزيون هو:

$$v(\hat{P}) = \left(1 - \frac{20}{20\,000}\right) \frac{0.6(1-0.6)}{19} = 0.0126$$

فترة الثقة البالغة ٩٥ في المائة للمتوسط السكاني هي:

$$0.6 \pm 1.96 \times \sqrt{0.0126} = (0.38, 0.82)$$

### ٣ - ٧ قياسات أخرى لخطأ المعاينة

٢٣ - بالإضافة إلى تباين المعاينة، توجد قياسات أخرى لخطأ المعاينة. وتشمل هذه القياسات الخطأ العادي، ومكافئ التباين، وأثر التصميم. هذه القياسات متصلة بالجبر بمعنى أنه يمكن استمداد التعبير عن أي واحد منها من القياسات الأخرى التي تستخدم عمليات جبرية بسيطة.

### ٧ - ٣ - ١ الخطأ المعياري

٢٤ - الخطأ المعياري في التقدير هو الجذر التربيعي لتباين المعاينة المتعلق به. تفسير هذا القياس أسهل لأنه يعطي إشارة إلى خطأ المعاينة باستخدام نفس المعيار الذي يستخدمه التقدير، بينما التباين يقوم على أساس مربع الفوارق.

٢٥ - ثمة سؤال يبرز مراراً في تصميم الاستقصاءات، وهو كم يجب أن يكون حجم الخطأ العادي لكي يعتبر مقبولاً؟ يتوقف الجواب على هذا السؤال على حجم التقدير. فالخطأ العادي الذي مقداره ١٠٠، مثلاً، يُعتبر صغيراً لدى تقدير الدخل السنوي، لكنه يعتبر كبيراً لدى تقدير متوسط معامل الترجيح للسكان. وكذلك الخطأ المعياري البالغ  $\sqrt{3\ 148\ 000\ 000}=56\ 107$  للمجموع المقدر البالغ ١٦٠ ٠٠٠ الذي حُصِلَ عليه من المثال ١ الوارد أعلاه يمكن اعتباره كبيراً جداً.

### ٧ - ٣ - ٢ مكافئ التباين

٢٦ - مكافئ تباين التقدير هو نسبة الخطأ العادي فيه إلى معدل قيمة التقدير نفسه. لذلك، يعطينا مكافئ التقدير مقاساً لخطأ المعاينة بالنسبة إلى الخاصية الجاري قياسها. ويعبر عنه عادة بنسبة مئوية.

٢٧ - مكافئ التباين مفيد في مقارنة دقة تقديرات الاستقصاء التي تتباين أحجامها أو نطاقاتها. غير أنه لا يفيد في تقديرات الخصائص التي تكون قيمتها صفراً أو رقماً سالباً، بما في ذلك تقديرات التغير، مثلاً، أو التغير في متوسط الدخل على مدى سنتين.

### ٧ - ٣ - ٣ أثر التصميم

٢٨ - يُعرّف أثر التصميم (المشار إليه بالرمز  $deff$ ) بأنه نسبة تباين المعاينة للتقدير بموجب تصميم معين إلى تباين المعاينة للتقدير بناءً على عينة عشوائية بسيطة من نفس الحجم. ويمكن أن يُعتقد بأنه هو العامل الذي يجب أن يُضرب به تباين التقدير القائم على أساس عينة عشوائية بسيطة من نفس الحجم، لكي يأخذ في الحسبان تعقيدات التصميم الفعلي للعينة نتيجة لعوامل مثل التقسيم الطبقي والتجميع والترجيح.

٢٩ - بعبارة أخرى، يوجد في التقدير القائم على بيانات مأخوذة من عينة مركبة حجمها  $n$  نفس التباين الذي ليوجد في تقدير محسوب من بيانات مأخوذة من عينة عشوائية بسيطة حجمها  $n/deff$ . لهذا السبب، تُسمى نسبة  $n/deff$  أحياناً الحجم الفعال للعينة للتقدير القائم على بيانات مأخوذة من تصميم مركب. للاطلاع على مناقشة عامة لحسابات "الحجم الفعال للعينة"، انظر كيش (١٩٩٥) وبوتهوف، وودبري ومانتون (١٩٩٢) والمراجع المشار إليها فيهما. وانظر أيضاً مختلف فروع الفصل الثالث للاطلاع على مناقشة أكثر تفصيلاً لآثار التصميم واستخدامها في تصميم العينات.

### ٧ - ٤ حساب تباين المعاينة لتصاميم معيارية أخرى

٣٠ - للتصاميم البسيطة والتقديرات الخطية البسيطة، كالتوسطات والنسب والمجاميع، يمكن في العادة استمداد معادلات يمكن استخدامها لحساب تباينات التقديرات. غير أنه في حالة أنواع التصميمات المركبة والتقديرات المرتبطة باستقصاءات الأسر المعيشية، يكون ذلك في الغالب صعباً أو مستحيلًا. في هذا الفرع،

نقدم أمثلة لتوضيح حساب تباين المعاينة لتصاميم عيّنات المجموعات المصنفة طبقياً والوحيدة المرحلة. أما معدلات وأمثلة حسابات التباين لتصاميم العيّنات العادية الأخرى فتوجد في الكتب المدرسية (مثل كوكران، ١٩٧٧؛ وكيش، ١٩٦٥).

### ٧ - ٤ - ١ المعاينة الطبقيّة

٣١ - يرد وصف مفصل للمعاينة الطبقيّة في الفصل الثالث وفي المرفق الأول. وفي هذا الفرع، سوف نُعنى بتقدير التباين فقط بموجب التصميم. انظر، مثلاً، تصميماً طبقياً فيه  $H$  طبقات، وتقديرات عيّنة لمتوسط السكان للطبقات المحددة بـ  $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_H$  وتقديرات المعاينة للتباينات السكانية للطبقات المحددة بـ  $S_1^2, S_2^2, \dots, S_H^2$ . يكون مقدّر متوسط السكان بموجب هذا التصميم كما يلي:

(٧ - ٧)

$$\hat{Y}_{st} = \sum_{h=1}^H \hat{Y}_h$$

حيث  $\hat{Y}_h$  هو التقدير المستند إلى العيّنة لـ  $H, \dots, h=1, \dots$ . ويعطى تباين التقدير كما يلي:

(٨ - ٧)

$$v(\hat{Y}_{st}) = \sum_{h=1}^H v(\hat{Y}_h)$$

وبالمعاينة الطبقيّة العشوائية، يعطى التقدير وتباينه المقدّر كما يلي:

(٩ - ٧)

$$\hat{Y}_{st} = \sum_{h=1}^H \frac{N_h}{N} \bar{y}_h = \sum_{h=1}^H W_h \bar{y}_h$$

حيث  $\bar{y}_h$  هو متوسط العيّنة للطبقة  $h$ ، ويكون  $Nh$  حجم السكان في الطبقة  $h$ ، و

$$W_h = \frac{N_h}{N}, h=1, \dots, H$$

ويعطى التباين المقدّر لهذا التقدير بموجب المعاينة الطبقيّة العشوائية كما يلي:

(١٠ - ٧)

$$v(\hat{Y}_{st}) = \sum_{h=1}^H W_h^2 v(\bar{y}_h) = \sum_{h=1}^H \left( \frac{N_h}{N} \right)^2 \left( 1 - \frac{n_h}{N_h} \right) \frac{s_h^2}{n_h}$$

حيث  $n_h$  هو حجم العيّنة في الطبقة  $h$ ، و  $s_h^2$  هو تباين العيّنة، وتقدير يستند إلى العيّنة لـ  $H, \dots, h=1, \dots$ .

### المثال ٢

نطبق الآن هذه النتائج على مثال للتصميم الطبقي يشمل ثلاث طبقات ومعلّمات كما وردت في الجدول ٧ - ٥ أدناه. افترض أننا مهتمون بتقدير متوسط السكان بناءً على عيّنة شاملة يبلغ حجمها ١٥٠٠.

الجدول ٧ - ٥

## مثال بيانات لتصميم عيّنة طبقية

المعلّمة	السكان	الطبقة الأولى (العاصمة)	الطبقة الثانية (إقليم - حضري)	الطبقة الثالثة (إقليم - حضري)
الحجم	$N = 1,000,000$	$N_1 = 300,000$	$N_2 = 500,000$	$N_3 = 200,000$
التباين	$S^2 = 75,000$	$S_1^2 = ?$	$S_2^2 = ?$	$S_3^2 = ?$
المتوسط	$\bar{Y} = ?$	$\bar{Y}_1 = ?$	$\bar{Y}_2 = ?$	$\bar{Y}_3 = ?$
تكلفة الوحدة	N/A	$C_1 = 1$	$C_2 = 4$	$C_3 = 16$
حجم العيّنة بموجب التخصيص المثالي <sup>أ</sup>	$n = 1,500$	$n_1 = 857$	$n_2 = 595$	$n_3 = 48$
متوسط العيّنات	N/A	$\bar{y}_1 = 4,000$	$\bar{y}_2 = 2,500$	$\bar{y}_3 = 1,000$
تباين العيّنات	N/A	$s_1^2 = 90,000$	$s_2^2 = 62,500$	$s_3^2 = 10,000$

ملاحظة: N/A تعني "لا ينطبق".

انظر الفصل الثالث.

تقدير متوسط السكان هو:

$$\hat{\bar{Y}}_{st} = \frac{300,000}{1,000,000} \times 4,000 + \frac{500,000}{1,000,000} \times 2,500 + \frac{200,000}{1,000,000} \times 1,000 = 2,650$$

التباين المقدّر للتقدير الوارد أعلاه هو:

$$v(\hat{\bar{Y}}_{st}) = \left( \frac{300,000}{1,000,000} \right)^2 \left( 1 - \frac{857}{300,000} \right) \left( \frac{90,000}{857} \right) + \left( \frac{500,000}{1,000,000} \right)^2 \left( 1 - \frac{595}{500,000} \right) \left( \frac{62,500}{595} \right) + \left( \frac{200,000}{1,000,000} \right)^2 \left( 1 - \frac{48}{200,000} \right) \left( \frac{10,000}{48} \right) = 43.98516$$

فترة الثقة البالغة ٩٥ في المائة لمتوسط السكان هي:

$$2,650 \pm 1.96 \times \sqrt{43.98516} = (2,637, 2,663)$$

لاحظ أن التباين المقدّر للمتوسط المقدّر بموجب المعاينة العشوائية البسيطة يُعطى بما يلي:

$$v(\hat{\bar{Y}}_{SRS}) = \left( 1 - \frac{1,500}{1,000,000} \right) \times \frac{75,000}{1,500} = 49.925$$

لذلك، يكون أثر هذا التصميم الطبقي  $\frac{43.98516}{49.925} = 0.88$  وحجم العيّنة الفعلي هو  $\frac{1,500}{0.88} = 1,705$ 

ويعني هذا أن تباين التقدير القائم على أساس عيّنة طبقية عشوائية حجمها ١٥٠٠ هو نفس التباين الموجود في التقدير القائم على أساس عيّنة عشوائية بسيطة حجمها ١٧٠٥.

٣٢ - يرد وصف مفصل لتقنية معاينة المجموعات في الفصل الثالث. ونقدم في هذا الفرع مثلاً بسيطاً لتوضيح حساب أخطاء المعاينة للحالة الخاصة — حالة معالجة المجموعات ذات المرحلة الواحدة.

## المثال ٣

افتراض أننا مهتمون بتقدير نسبة الطلاب الذين هم في سن المدرسة في الإقليم، الذين حُصِّنوا ضد شلل الأطفال. افترض، لأجل البساطة، أنه يوجد ما مجموعه ٥٠٠ منطقة تعداد متساوية الحجم في الإقليم، ويوجد في كل منطقة تعداد ٢٥ طفلاً في سن المدرسة. تعمل مناطق التعداد بمثابة المجموعات في هذا المثال. افترض أننا اخترنا ١٠ مناطق تعداد بمعاينة عشوائية بسيطة بدون تبديل من الـ ٥٠٠ منطقة تعداد في الإقليم، وأنه قيست نسبة الأطفال المحصنين الذين هم في سن المدرسة لكل منطقة معاينة مدرجة في منطقة التعداد المدرجة في العيّنة، وكانت نتائجها كالنتائج الواردة في الجدول ٧ - ٦ أدناه.

الجدول ٧ - ٦

نسب الأطفال المحصنين الذين هم في سن المدرسة في ١٠ مناطق تعداد، باعتبارها المتغير الذي هو موضع اهتمام

منطقة التعداد المدرجة في العيّنة (i)	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
نسبة العيّنة ( $\hat{P}_i$ )	$\frac{8}{25}$	$\frac{10}{25}$	$\frac{12}{25}$	$\frac{14}{25}$	$\frac{15}{25}$	$\frac{17}{25}$	$\frac{20}{25}$	$\frac{20}{25}$	$\frac{21}{25}$	$\frac{23}{25}$

في هذا المثال، يكون تقدير نسبة الأطفال المحصنين الذين هم في سن المدرسة:

$$\hat{P} = \frac{160}{250} = 0.64 \text{ أو } 64\%$$

علاوةً على ذلك، يكون تباين العيّنة:

$$s_p^2 = \frac{1}{10-1} \sum_{i=1}^{10} (\hat{P}_i - \hat{P})^2 = 0.040533$$

لذلك، يكون تباين النسبة المقدرة:

$$v(\hat{P}) = \left(1 - \frac{10}{500}\right) \times \frac{0.040533}{10} = 0.003972$$

لاحظ أنه بموجب المعاينة العشوائية البسيطة، يكون تباين النسبة المقدرة:

$$v(\hat{P}_{SRS}) = \left(1 - \frac{250}{12500}\right) \times \frac{0.64(1-0.64)}{250-1} = 0.0009078$$

لذلك يكون أثر تصميم معاينة هذه المجموعة = 4.38  $\frac{0.003972}{0.0009078}$ 

$$\text{ويكون حجم العيّنة الفعلي} = \frac{250}{4.38} = 57$$

ويعني هذا أن التقدير القائم على أساس عيّنة المجموعة التي حجمها ٢٥٠ هو نفس التباين للمجموعة القائمة على أساس عيّنة عشوائية بسيطة حجمها ٥٧.

## ٧ - ٥ المَعْلَمَاتُ المشتركة لتصاميم وبيانات عيِّنة استقصاء الأسر المعيشية

### ٧ - ٥ - ١ انحرافات تصاميم استقصاء الأسر المعيشية عن المعاينة العشوائية البسيطة

٣٣ - كما ذُكر آنفاً، نادراً ما تستخدم المعاينة العشوائية البسيطة في الواقع لاستقصاءات الأسر المعيشية الكبيرة الحجم لأن تنفيذها باهظ التكلفة. غير أن من الأهمية بمكان أن يُفهم هذا التصميم فهماً كاملاً لأنه يشكّل الأساس النظري لتصاميم عيّنات أكثر تعقيداً. وتنحرف معظم تصاميم عيّنات استقصاء الأسر المعيشية عن المعاينة العشوائية البسيطة بسبب وجود واحد أو أكثر من المعالم الثلاثة التالية:

- (أ) التقسيم الطبقي في مرحلة واحدة أو أكثر من مراحل المعاينة؛
- (ب) تجميع الوحدات في مرحلة واحدة أو أكثر من مراحل المعاينة، مما يخفض التكاليف لكنه يزيد تباين التقديرات بسبب الترابط بين الوحدات في العيِّنة نفسها؛
- (ج) الترجيح للتعويض عن نواقص في العيِّنة مثل عدم تساوي احتمالات الاختيار، وعدم الاستجابة، وعدم التغطية (انظر الفصل السادس للاطلاع على التفاصيل).

٣٤ - يشار إلى تصميم العيِّنة بأنه مركَّب إذا كان فيه واحد أو أكثر من المَعْلَمَاتُ المذكورة أعلاه. ومعظم تصاميم استقصاءات الأسر المعيشية مركَّبة، وبذلك تنتهك افتراضات المعاينة العشوائية البسيطة. لذلك من شأن تحليل بيانات استقصاءات الأسر المعيشية كما لو كانت مولدة بتصميم عيِّنة عشوائية بسيطة سيؤدي إلى أخطاء في التحليل وفي الاستنتاجات القائمة على هذه البيانات. علاوةً على ذلك، كما ذُكر أعلاه، لا يمكن التعبير عن التقديرات التي هي موضع اهتمام في معظم استقصاءات الأسر المعيشية كدالاتٍ خطية للملاحظات، لذلك لا يمكن أن يكون ثمة أيّ معادلات مغلقة الشكل للتباينات. تتناول الفروع التالية مسألة طرق تقدير التباين لتصاميم استقصاءات الأسر المعيشية التي تأخذ التعقيدات المذكورة أعلاه في الحسبان.

### ٧ - ٥ - ٢ إعداد ملفات البيانات لتحليلها

٣٥ - تكون بيانات الاستقصاءات التي تُجمع في البلدان النامية أحياناً غير قابلة للتحليل الذي يتجاوز التواتر الأساسي والجدولة. وتوجد بضعة أسباب لعدم القابلية هذا. أولاً، ربما يكون توثيق العيِّنة للاستقصاء محدوداً جداً أو غير تقني. ثانياً، ربما لا تكون ملفات البيانات بالشكل والهيكل والمعلومات اللازمة التي تسمح بإجراء أيّ تحليل متقدم. ثالثاً، ربما لا توجد البرامج الحاسوبية الملائمة والدراية الفنية.

٣٦ - لتحليل بيانات العيِّنة الاستقصائية تحليلاً ملائماً، يجب أن تتضمن قاعدة البيانات المرتبطة بالعملية جميع المعلومات التي تعكس عملية اختيار العيِّنة (انظر الفصل الخامس للاطلاع على مناقشة مفصلة). وينبغي أن تشمل قاعدة البيانات، على وجه الخصوص، علامات تعريف ملائمة لطبقات تقسيم العيّنات، و وحدات المعاينة الأولى، و وحدات المعاينة الثانية، إلخ. في بعض الأحيان، تحتاج الطبقات الفعلية و وحدات المعاينة الأولى المستخدمة في اختيار العيِّنة للاستقصاء إلى تعديل لأغراض تتعلق بتقدير التباين. هذه التعديلات تكون ضرورية لجعل تصميم العيِّنة الفعلي ينسجم مع واحد من خيارات تصميم العيِّنة المتاحة في واحد على الأقل من البرامج الحاسوبية للتحليل الإحصائي (انظر الفرع ٧ - ٩). الطبقات و وحدات المعاينة الأولى التي تُنشأ لتقدير التباين تُسمّى أحياناً الطبقات الكاذبة، أو طبقات التباين أو وحدات المعاينة الأولى الكاذبة أو وحدات المعاينة التباينية.

ويجب إدراج متغيرات تصميم العينة ذات الصلة، وكذلك المتغيرات التي تُنشأ لأغراض تقدير التباين، في ملف البيانات، إلى جانب الوثائق ذات الصلة التي تبين كيف تعرّف هذه المتغيرات وكيف تُستعمل. ويلزم لتقدير التباين مجموعة لا تقل عن ثلاثة متغيرات: معامل ترجيح العينة، والطبقة (أو الطبقة الكاذبة)، ووحدة المعاينة الأولية (أو وحدة المعاينة الأولية الكاذبة). تلخص هذه المتغيرات الثلاثة تصميم العينة ومن شأن إدراجها في مجموعة بيانات الاستقصاء أن يمكّن من إجراء التحليل الملائم للبيانات، وأن يفسّر التعقيدات في تصميم العينة.

٣٧ - علاوة على ذلك، ينبغي تطوير معاملات ترجيح العينات لكل وحدة معاينة في ملف البيانات. وينبغي أن تعكس معاملات الترجيح هذه احتمال الاختيار لكل وحدة معاينة، وأن تعوض كذلك عن عدم الاستجابة في الاستقصاء والنواقص الأخرى في العينة. تلزم معاملات ترجيح متغيرات التصميم ورقاع تعريفها لتقدير قابلية تقديرات الاستقصاء للتغير تقديراً ملائماً. وكما ذُكر في الفصل السادس وفي الفروع السابقة من هذا الفصل، معاملات ترجيح العينات هامة ليس فقط لتوليد تقديرات ملائمة للاستقصاء، وإنما أيضاً لتقدير أخطاء المعاينة التي تقع في هذه التقديرات. لذلك، من الأساسي أن تُدرج كل المعلومات عن معاملات الترجيح في ملفات البيانات. وعلى وجه الخصوص، كلما حدث تعديل بسبب عدم استجابة، أو تعديل لما بعد التقسيم الطبقي، أو أنواع أخرى من التعديلات يجب أن تتضمن وثائق الاستقصاء وصفاً لإجراءات التعديل هذه.

### ٧ - ٥ - ٣ أنواع تقديرات الاستقصاء

٣٨ - في معظم استقصاءات الأسر المعيشية، أكثر تقديرات الاستقصاء، التي هي موضع اهتمام، شيوعاً تأتي في شكل مجاميع ونسب. افترض تصميمياً طبقياً ذا ثلاث مراحل مع وحدات معاينة أولية في المرحلة الأولى، ووحدات معاينة ثانوية في المرحلة الثانية وأسر معيشية في المرحلة الثالثة. يمكن التعبير عن التقدير الاستقصائي لمجموع ما كما يلي:

(٧ - ١١)

$$\hat{Y} = \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} \sum_{k=1}^{l_j} W_{hijk} Y_{hijk}$$

حيث  $W_{hijk}$  تساوي معامل الترجيح النهائي للأسرة المعيشية (لـ  $j = 1, \dots, m_i$ ) رقم  $k^{th}$  المختارة في وحدة المعاينة الثانوية (لـ  $i = 1, \dots, n_h$ ) رقم  $i^{th}$  في الطبقة (لـ  $h = 1, \dots, H$ ) رقم  $h^{th}$ ؛ و

$Y_{hijk}$  تساوي قيمة المتغير  $Y$  للأسرة المعيشية رقم  $k^{th}$  المختارة في وحدة المعاينة الثانوية رقم  $i^{th}$  في وحدة المعاينة الأولية رقم  $i^{th}$  في الطبقة رقم  $h^{th}$ .

٣٩ - على المستوى الأساسي جداً، تتناسب معاملات الترجيح المرتبطة بوحدات العينة تناسباً عكسياً مع احتمالات اختيار الوحدات لإدراجها في العينة. غير أنه غالباً ما تُستخدم طرق أكثر تقدماً لحساب معاملات الترجيح المراد تطبيقها في التحليل. وقد وُصفت بعض هذه الطرق في الفصل السادس وفي المراجع المشار إليها فيه.

٤٠ - يعرف تقدير الاستقصاء للنسبة كما يلي:

(٧ - ١٢)

$$\hat{R} = \frac{\hat{Y}}{\hat{X}}$$

حيث  $\hat{Y}$  و  $\hat{X}$  تقديران لمجموعي المتغيرين  $Y$  و  $X$ ، على التوالي، حُسبا كما هو مبين في المعادلة (٧ - ١٢) أعلاه.

٤١ - في حالة المعاينة متعددة المراحل، تكون المتوسطات والنسب مجرد حالات خاصة لمقدّر النسب. ففي حالة المتوسط، المتغير  $X$  في مقام الكسر هو حساب متغير يُعرّف بأنه يساوي ١ لكل عنصر بحيث يكون المقام مجموع مُعاملات الترجيح. وفي حالة النسبة، يُعرّف مقام الكسر  $X$  أيضاً بأنه يساوي ١ لجميع العناصر؛ وإن المتغير  $Y$  في البسط متغير ذو حدّين يعرّف بأنه إما صفر وإما واحد، وذلك يتوقف على ما إذا كانت الوحدة المراقبة تجهز الخاصة الجاري تقدير نسبتها. في معظم استقصاءات الأسر المعيشية يعرف المقام في مقدر النسبة بأنه يعرّف بأشكال مختلفة — كمجموع السكان، مجموع الإناث، مجموع الذكور، مجموع سكان الريف، مجموع السكان في إقليم معين أو مقاطعة معيّنة، إلخ.

## ٧ - ٦ مبادئ توجيهية لتقديم معلومات عن أخطاء المعاينة

### ٧ - ٦ - ١ اتخاذ قرار بشأن ما يوضع في التقرير

٤٢ - في الاستقصاءات الوطنية الكبيرة، التي توجد فيها متغيرات عديدة ومجالات في موضع الاهتمام وأهداف متعددة وغالباً ما تكون متنافسة، ليس من العملي عرض كل تقدير إلى جانب خطأ المعاينة المرتبط به. فمن شأن هذا ليس فقط أن يزيد حجم المنشور وإنما يربح أيضاً أن يجعل العرض مزدحماً بالنتائج الفنية. وفي ضوء التباين المتوقع في تقديرات أخطاء المعاينة نفسها، ربما يؤدي عرض نتائج متغيرات بالغة العدد إلى تشوش وشعور بمفهوم غير متسق فيما يتعلق بالنوعية الإجمالية لبيانات الاستقصاء التي جمعت. وسيكون أكثر فائدة بكثير أن تُعرض المعلومات المتعلقة بخطأ المعاينة لأهم الخصائص، التي هي موضع اهتمام، مسبقاً، وأن يوضع الباقي في تذييل.

٤٣ - من المهم، لدى عرض معلومات عن أخطاء المعاينة، أن نبقى نُصب أعيننا أثرها المحتمل على تفسير نتائج الاستقصاء والقرارات السياسية التي يمكن أن تُستمدّ من هذا التفسير. وينبغي دائماً أن يُنظر إلى المعلومات المتعلقة بأخطاء المعاينة باعتبارها مجرد عنصر من العناصر التي يتكوّن منها مجموع خطأ المعاينة، وليست دائماً هي العنصر الأهم. ففي بعض أوضاع الاستقصاء، ربما يكون لأخطاء غير أخطاء المعاينة (انظر الفصل الثامن) أثرٌ أهم من آثار أخطاء المعاينة على النوعية الإجمالية لبيانات الاستقصاء. لهذا السبب، يوصى بأن تشمل المعلومات المقدمة عن أخطاء المعاينة مناقشة للمصادر الرئيسية للأخطاء التي هي من غير أخطاء المعاينة وبعض عمليات التقييم النوعي للأثر على النوعية الإجمالية لبيانات الاستقصاء. ولأن أخطاء المعاينة تصبح أكثر أهمية بكثير على المستويات المنخفضة للتقسيم، يوصى أيضاً بإدراج بعض الملاحظات التحذيرية بشأن درجة تقسيم بيانات الاستقصاء.

٤٤ - ينبغي بوجه عام أن تشمل المعلومات المقدمة عن أخطاء المعاينة من التفاصيل ما يكفي لتسهيل تفسير نتائج الاستقصاء تفسيراً صحيحاً وتلبية متطلبات منظور مستخدمي البيانات بكامله، من مستخدمي البيانات العامة إلى راسمي السياسة (الذين يكمن اهتمامهم في استخدام نتائج الاستقصاء لصياغة السياسة) والمحلل الفني الذي (يعمل في مواصلة التحاليل وتقديم تقارير عن النتائج) إلى الخبير الإحصائي في المعاينة (المعني بالكفاءة الإحصائية للتصميم مقارنةً بالبدائل الأخرى، وبجوانب التصميم التي يمكن استخدامها في تصميم الاستقصاءات في المستقبل).

## ٧ - ٦ - ٢ كيف تقدم تقارير عن معلومات أخطاء المعاينة

٤٥ - يمكن أن تعرض أخطاء المعاينة في ثلاثة أشكال مختلفة، هي:

- (أ) بالقيم المطلقة للأخطاء المعيارية؛  
 (ب) بالأخطاء المعيارية النسبية ( الجذور التربيعية للتباينات النسبية)؛  
 (ج) بفترات الثقة.

٤٦ - يتوقف الخيار بين هذه الأشكال الثلاثة للعرض على طبيعة التقدير. ففي المؤسسات التي تتباين فيها التقديرات من حيث الحجم ووحدات القياس، يمكن أن تنطبق قيمة الأخطاء المعيارية نفسها على التقديرات عندما يُعبر عنها بالقيم النسبية؛ ومن ثمّ ربما يكون أكثر كفاءة لعرض الأخطاء المعيارية النسبية. غير أن الأخطاء المعيارية المطلقة، بوجه عام، أيسر فهماً وأوثق صلةً بالتقدير، لا سيما في حالات النسب المئوية والنسب والمعدلات. ويستدعي استخدام فترات الثقة اختيار مستوى للثقة (ولنقل ٩٠ أو ٩٥، مثلاً، أو ٩٩ في المائة). وحيث إن هذا المستوى يختلف بحسب أهداف الاستقصاء ومتطلبات الدقة في التقديرات، من المهم تحديد مستوى الثقة المستخدم في عرض معلومات أخطاء المعاينة ثم المحافظة على مستوى الثقة هذا طول الوقت في تحديد أهمية النتائج. والنسبة التي هي أكثر ما تكون استخداماً، في الواقع، كما ذكر أعلاه، هي فترة الثقة البالغة ٩٥ في المائة (انظر الفقرة ٣٠ من الفصل الثالث، والفقرة ٢٢ من الفصل السابع)، أي:

(٧ - ١٣)

التقدير  $\pm 1.96 \times$  الخطأ المعياري

٤٧ - للاطلاع على مزيد من التفاصيل حول موضوع عرض المعلومات عن أخطاء المعاينة، بما في ذلك مبادئ توجيهية محددة لمختلف فئات المستخدمين وعدد الأشكال التوضيحية، انظر "الأمم المتحدة (١٩٩٣)" والمراجع المشار إليها فيه.

## ٧ - ٦ - ٣ قاعدة تقريبية في الإبلاغ عن الأخطاء المعيارية

٤٨ - ثمة قاعدة تقريبية تُستخدم في أحيان كثيرة في الإبلاغ عن الأخطاء المعيارية، وهي تقرير الخطأ المعياري إلى رقمين هامين ثم توضع تقديرات الأرقام العشرية الملازمة بنفس عدد الخانات العشرية التي في الخطأ المعياري. وذلك مثلاً:

١ - إذا كان تقدير النقطة ٧٣ ٤٥٦ والخطأ المعياري ٢ ٣٤٥، عندئذٍ نبلغ عن تقدير النقطة بـ ٧٣ ٥٠٠ والخطأ المعياري بـ ٢ ٣٠٠.

٢ - إذا كان تقدير النقطة ١,٥٤٣٢٨ والخطأ العادي ٠,٠١٣٥٦، عندئذٍ نبلغ عن تقدير النقطة بـ ١,٥٤٣ والخطأ العادي بـ ٠,٠١٤.

٤٩ - يبدو أن التفكير العام وراء هذه القاعدة التقريبية يتصل بالإحصاءات التكنولوجية. فوجود رقمين هامين في الخطأ المعياري والعدد المشابه من الأرقام في التقدير النقطي يضمن ألا يكون ثمة أثر كبير جداً لخطأ التقريب في الإحصاءات التكنولوجية الناتجة الذي ذكر أنه يوجد في عدد كبير من الأرقام غير ذات الصلة. غير أنه تجدر ملاحظة أن هذه القاعدة التقريبية لا تعمل بالضرورة في أوضاع لا تكون فيها الإحصاءات التكنولوجية موضع اهتمام رئيسي.

## ٧ - ٧ طرق تقدير التباين في استقصاءات الأسر المعيشية

٥٠ - نصف في هذا الفرع بإيجاز بعض الطرق التقليدية لتقدير التباينات أو أخطاء المعاينة المستندة إلى بيانات الاستقصاء. ويمكن تقسيم طرق تقدير أخطاء المعاينة لاستقصاءات الأسر المعيشية في أربع فئات عامة، هي:

(أ) الطرق المضبوطة؛

(ب) طريقة المجموعة النهائية؛

(ج) تقريبات ضبط الخطية؛

(د) تقنيات التكرار.

سنناقش الآن بإيجاز كل واحدة من هذه الفئات بدورها. يمكن للقراء المهتمين أن يحصلوا على مزيد من التفاصيل من مراجع مثل كيش وفرانكل (١٩٧٤)، وفولتر (١٩٨٥)، وليتونين وباخينين (١٩٩٥).

### ٧ - ٧ - ١ الطرق المضبوطة

٥١ - أعطى الفرعان ٧ - ٢ و ٧ - ٤ بضعة أمثلة على طرق الرقابة لتقدير التباين في تصاميم العيّنات المعيارية. وتشكّل هذه الطرق أفضل نهج لتقدير التباين عندما تنطبق عليها. غير أن تطبيقها على حساب تباينات المعاينة في التقديرات المستندة إلى بيانات استقصاء الأسر المعيشية يكون معقداً بسبب بضعة عوامل. أولاً، تصاميم العيّنات المستخدمة لمعظم استقصاءات الأسر المعيشية أكثر تعقيداً من المعاينة العشوائية البسيطة (انظر الفرع ٧ - ٥ - ١ أعلاه). ثانياً، التقديرات التي هي موضع اهتمام ربما لن تكون في شكل الدالات الخطية البسيطة للقيم الملاحظة، ولذلك نجد أن تباين العيّنة لا يمكن في أحيان كثيرة التعبير عنه بمعادلات ذات شكل مغلق، مثل معادلة متوسط العيّنة بموجب المعاينة العشوائية البسيطة أو بموجب المعاينة الطبقيّة. علاوة على ذلك، تتوقف طرق الرقابة على تصميم العيّنات الذي هو قيد النظر في تطبيق معين، والتقدير الذي هو موضع الاهتمام، وإجراءات الترجيح المستخدمة.

٥٢ - نناقش في الفروع التالية طرق تقدير التباين لتصاميم عيّنات تستخدم عادةً في استقصاءات الأسر المعيشية. هذه الطرق مصمّمة للتغلب على نواقص طرق الرقابة.

### ٧ - ٧ - ٢ طريقة المجموعة النهائية

٥٣ - يمكن أن تستخدم طريقة المجموعة النهائية لتقييم التباين (انظر هانسن وهرويتز ومادو، ١٩٥٣، الصفحات ٢٥٧ - ٢٥٩) لتقدير التباينات في تقديرات الاستقصاء بناءً على عيّنة ولدها تصميم عيّنة معقد. وبموجب هذه الطريقة تتكوّن المجموعة النهائية من العيّنة الكاملة التي أنشئت من وحدة معاينة أولية، بغض النظر عن المعاينة في مراحل لاحقة من تصميم متعدد المراحل. وتُحسب تقديرات التباين باستخدام مجاميع ما بين وحدات المعاينة الأولية دون الحاجة إلى حساب عناصر التباين في كل مرحلة من مراحل الاختيار.

٥٤ - افترض أن عيّنة من  $n_b$  وحدة من وحدات المعاينة الأولية قد اختيرت من الطبقة رقم  $h$  (بأي عدد من المراحل داخل وحدات المعاينة الأولية). عندئذٍ يُعطى تقدير المجموع من الطبقة رقم  $h$  بالمعادلة:

(١٤ - ٧)

$$\hat{Y}_{hi} = \sum_{j=1}^{m_i} W_{bijk} Y_{bijk} \quad \text{حيث} \quad \hat{Y}_b = \sum_{i=1}^{n_b} \hat{Y}_{hi}$$

لاحظ أن التقدير على مستوى وحدة المعاينة الأولية  $\hat{Y}_{hi}$  هو تقديراً لـ  $\frac{\hat{Y}_b}{n_b}$ . وهكذا يُعطى تباين التقديرات الفردية على مستوى وحدة المعاينة الأولية بما يلي:

(١٥ - ٧)

$$v(\hat{Y}_{hi}) = \frac{1}{n_b - 1} \sum_{i=1}^{n_b} \left( \hat{Y}_{hi} - \frac{\hat{Y}_b}{n_b} \right)^2$$

وتباين مجموعهما،  $\hat{Y}_b$ ، والمجموع على مستوى الطبقة، مقدراً من عينة عشوائية حجمها  $n_b$  يُعطى تقديراً مجموع السكان للطبقة  $h$  بما يلي:

(١٦ - ٧)

$$v(\hat{Y}_b) = \frac{n_b}{n_b - 1} \sum_{i=1}^{n_b} \left( \hat{Y}_{hi} - \frac{\hat{Y}_b}{n_b} \right)^2$$

٥٥ - لاحظ أن التعامل الجبري البسيط ينتج التعبير المكافئ التالي لتقدير تباين مجموع السكان للطبقة  $h$

(١٧ - ٧)

$$v(\hat{Y}_b) = \frac{n_b}{n_b - 1} \left\{ \sum_{i=1}^{n_b} \hat{Y}_{hi}^2 - \frac{\left( \sum_{i=1}^{n_b} \hat{Y}_{hi} \right)^2}{n_b} \right\}$$

٥٦ - وأخيراً، بالمعينة المستقلة في جميع الطبقات، يُحصل على تقدير التباين لمجموع السكان الإجمالي بأخذ مجموع تباينات المجاميع على مستوى الطبقة، أي:

(١٨ - ٧)

$$v(\hat{Y}) = \sum_{b=1}^H v(\hat{Y}_b)$$

لاحظ أنه يُستخدم في بعض الأحيان معامل تصحيح للمجموعة السكانية المحدودة  $(1 - n_b/N_b)$  في المعادلة الواردة أعلاه.

٥٧ - المعادلة (١٨ - ٧) ملحوظة بمعنى أن تباين المجموع المقدّر هو دالة مجاميع وحدات المعاينة المرجحة ترجيحاً ملائماً  $\hat{Y}_{hi}$  فقط، دون أي إشارة إلى هيكل المعاينة وطريقتها في وحدات المعاينة الأولية. هذا يسهّل معادلة تقدير التباين تسهلاً كبيراً لأن حساب عناصر التباين المعزّوة إلى المراحل الأخرى من مراحل المعاينة في وحدات المعاينة الأولية غير مطلوب. هذه السمة تعطي طريقة المجموعة النهائية مرونة كبيرة في تناول تصاميم عينات مختلفة وهي، في الواقع، واحدة من نقاط القوة الرئيسية للطريقة، وهي سبب رئيسي لانتشار استخدامها في عمل الاستقصاءات.

٥٨ - الآن يُعطى مقدر التباين للنسبة،  $\hat{R} = \frac{\hat{Y}}{\hat{X}}$ ، بما يلي:

(١٩ - ٧)

$$v(\hat{R}) = \frac{1}{\hat{X}^2} \left\{ v(\hat{Y}) + R^2 v(\hat{X}) - 2\hat{R} \text{cov}(\hat{Y}, \hat{X}) \right\}$$

حيث تُحسب  $v(\hat{Y})$  و  $v(\hat{X})$  وفقاً لتباين مجموع مقدر، و

$$(٢٠ - \text{v}) \quad \text{cov}(\hat{Y}, \hat{X}) = \sum_{h=1}^H \left\{ \frac{n_h}{n_b-1} \sum_{i=1}^{n_h} \left( \hat{X}_{hi} - \frac{\hat{X}_h}{n_b} \right) \left( \hat{Y}_{hi} - \frac{\hat{Y}_h}{n_b} \right) \right\}$$

أو بما يكافئه،

$$(٢١ - \text{v}) \quad \text{cov}(\hat{Y}, \hat{X}) = \frac{n_b}{n_b-1} \left\{ \sum_{i=1}^{n_b} \hat{X}_{bi} \hat{Y}_{bi} - \frac{\left( \sum_{i=1}^{n_b} \hat{X}_{bi} \right) \left( \sum_{i=1}^{n_b} \hat{Y}_{bi} \right)}{n_b} \right\}$$

٥٩ - لاحظ أن المعادلة الواردة أعلاه لتباين نسبة يمكن تبسيطها باستخدام حقيقة أن التباين النسبي للنسبة مساو تقريباً للفرق بين التباينات النسبية للبسط والمقام. تذكّر أن التباين النسبي للتقدير هو نسبة تباينه إلى مُربّعه. هكذا، للنسبة المقدرة  $\hat{R}$ ، يعطى التباين النسبي، المشار إليه  $relvar(\hat{R})$ ، بما يلي:

$$(٢٢ - \text{v}) \quad relvar(\hat{R}) = \frac{v(\hat{R})}{\hat{R}^2}$$

لذلك، يُعطى تقدير لتباين النسبة بما يلي:

$$(٢٣ - \text{v}) \quad v(\hat{R}) = \hat{R}^2 relvar(\hat{R}) = \hat{R}^2 \{ relvar(\hat{Y}) - relvar(\hat{X}) \}$$

٦٠ - يمكن تخطيط طريقة المجموعة النهائية لحساب أخطاء المعاينة لمجاميع ونسب مقدرة بالخطوات التالية:

- الخطوة ١ - لكل طبقة على حدة، احسب التقدير المرجح  $\hat{Y}_{bi}$  للخاصية التي هي موضع اهتمام،  $Y$ ، لكل وحدة معاينة أولية (وفقاً لإجراءات الترجيح المحددة في الفصل السادس).
- الخطوة ٢ - احسب مربع القيمة لقيمة كل وحدة معاينة أولية من الخطوة ١ أعلاه.
- الخطوة ٣ - احسب مجموع القيم من الخطوة ٢ على جميع وحدات المعاينة الأولية في الطبقة.
- الخطوة ٤ - احسب مجموع المجاميع المقدرة لوحدات المعاينة الأولية من الخطوة ١ على جميع وحدات المعاينة الأولية.
- الخطوة ٥ - خذ مربع حاصل الخطوة ٤ واقسم على  $n_b$ ، عدد وحدات المعاينة الأولية في الطبقة.
- الخطوة ٦ - اطرح حاصل الخطوة ٥ من حاصل الخطوة ٣، واضرب هذا الفرق بالمعامل  $n_b/(n_b-1)$ . هذا هو التباين المقدر للخاصية على مستوى الطبقة.
- الخطوة ٧ - اجمع حاصل الخطوة ٦ في جميع الطبقات للحصول على خطأ المعاينة المقدر للخاصية التي هي موضع الاهتمام.

- الخطوة ٨ - اجمع الجذر التربيعي لحاصل الخطوة ٧ للحصول على خطأ المعاينة المقدر للخاصية التي هي موضع الاهتمام.
- ٦١ - لحساب خطأ المعاينة المقدر للنسب، كالنسب المقدرة، نمضي في الخطوات التالية:
- الخطوة ٩ - احسب التباين النسبي للبسط،  $\bar{Y}$ ، بتقسيم حاصل الخطوة ٧ على مربع تقدير المقام.
- الخطوة ١٠ - كرر الخطوة ٩ للحصول على التباين النسبي للمقام،  $\bar{X}$ .
- الخطوة ١١ - اطرح حاصل الخطوة ١٠ من حاصل الخطوة ٩.
- الخطوة ١٢ - اضرب حاصل الخطوة ١١ بمربع النسبة المقدرة،  $\hat{R}$ . هذا هو التباين المقدر لـ  $\hat{R}$ .
- الخطوة ١٣ - احسب الجذر التربيعي لحاصل الخطوة ١٢ للحصول على خطأ المعاينة المقدر لـ  $\hat{R}$ .

#### المثال ٤

ننظر الآن في مثال افتراضي لتوضيح طريقة المجموعة النهائية لتقدير التباين. افترض أننا مهتمون بتقدير مجموع النفقات الأسبوعية على الغذاء للأسر المعيشية في المدينة ألف. نصمم استقصاءً مستخدمين تصميمًا طبقيًا لمجموعة ذات ثلاث مراحل ويشتمل على ثلاث طبقات تُختار فيه وحدتا معاينة أوليتان من كل طبقة، وتُختار أسرَتان معيشيتان من كل وحدة معاينة أولية مدرجة في العينة. ثم تُسجل النفقات الأسبوعية على الغذاء لكل أسرة مدرجة في العينة في الاستقصاء. يعرض الجدول ٧ - ٧ البيانات التي حُصلَ عليها من هذا الاستقصاء، بما في ذلك مُعاملات الترجيح ( $W_{hij}$ ) والنفقات الأسبوعية على الغذاء بالدولار ( $Y_{hij}$ )، لكل أسرة مدرجة في العينة.

الجدول ٧ - ٧

النفقات الأسبوعية للأسرة المعيشية على الغذاء، بحسب الطبقة

الطبقة	وحدة المعاينة الأولية	الأُسرة	معامل الترجيح ( $W_{hij}$ )	النفقات بالدولار ( $Y_{hij}$ )	$W_{hij} * Y_{hij}$
١	١	١	١	٣٠	٣٠
		٢	١	٢٨	٢٨
	٢	١	٣	١٢	٣٦
		٢	٣	١٥	٤٥
٢	١	١	٥	٦	٣٠
		٢	٥	٧	٣٥
		٢	٢	١٦	٣٢
	٢	١	٦	٧	٤٢
		٢	٢	١٨	٣٦
		٢	٦	٧	٤٢
٣	١	١	٦	٨	٤٨
		٢	٤	١٣	٥٢
	٢	٤	١٥	٦٠	
المجموع			٤٢	٤٧٤	٤٧٤

٦٢ - عملاً بالفصل السادس، يُعطى هنا تقدير لمجموع النفقات الأسبوعية للأسرة المعيشية في المدينة على الغذاء بما يلي:

$$\hat{Y} = \sum_{b=1}^3 \sum_{i=1}^2 \sum_{j=1}^2 W_{bij} Y_{bij} = 474$$

ويعطى أيضاً تقدير لمتوسط النفقات الأسبوعية للأسرة المعيشية على الغذاء بما يلي:

$$\hat{\bar{Y}} = \frac{\sum_{b=1}^3 \sum_{i=1}^2 \sum_{j=1}^2 W_{bij} Y_{bij}}{\sum_{b=1}^3 \sum_{i=1}^2 \sum_{j=1}^2 W_{bij}} = \frac{474}{42} = 11$$

(مقرباً إلى أقرب دولار)

٦٣ - ننفذ الآن الخطوات في طريقة المجموعة النهائية لتقدير التباين في أعمدة الجدول ٧ - ٨ أدناه. تتفق أرقام الأعمدة مع الخطوات التي أوجزناها أعلاه.

الجدول ٧ - ٨

### تنفيذ الخطوات في طريقة المجموعة النهائية لتقدير التباين

الطبقة	وحدة المعاينة الأولية	الخطوة ١	الخطوة ٢	الخطوة ٣	الخطوة ٤	الخطوة ٥	الخطوة ٦	الخطوة ٧
١	١	٥٨	٣٣٦٤	٩٩٢٥	١٢٩	٩٦٦٠,٥٠	٥٢٩	
	٢	٨١	٦٥٦١	-	-	-	-	
٢	١	٦٥	٤٢٢٥	٨٨٤٩	١٣٣	٨٨٤٤,٥	٩	
	٢	٦٨	٤٦٢٤	-	-	-	-	
٣	١	٩٠	٨١٠٠	٢٠٦٤٤	٢٠٢	٢٠٤٠٢	٤٨٤	
	٢	١١٢	١٢٥٤٤	-	-	-	-	
المجموع								١٠٢٢

٦٤ - تقديرات التباين على مستوى الطبقة هي ٥٢٩ للطبقة ١، و ٩ للطبقة ٢، و ٤٨٤ للطبقة ٣. يُحصل على التقدير الإجمالي للتباين في المجموع المقدر لإيرادات الأسرة المعيشية الأسبوعية (الخطوة ٧ من مخططنا) بجمع كل التقديرات على مستوى الطبقة، ويبلغ مجموعها ١٠٢٢.

### ٧ - ٧ - ٣ تقريبات ضبط الخطية

٦٥ - معظم التقديرات التي هي موضع اهتمام في استقصاءات الأسر المعيشية غير خطية. في طريقة الضبط الخطي تُضبط هذه التقديرات غير الخطية "ضبطاً خطياً" باستخدام توسيع لسلسلة تايلور. وينطوي هذا على التعبير عن التقدير بمصطلحات توسيع لسلسلة تايلور ثم تقريب تباين التقدير بتباين المرتبة الأولى الفرع الخطي من ذلك التباين، باستخدام الطرق المضبوطة التي نوقشت في أجزاء سابقة.

٦٦ - افترض أننا نود تقدير تباين التقدير  $z$  لمعلمة  $Z$  وافترض أن  $z$  دالة غير خطية لتقديرات بسيطة  $y_1, y_2, \dots, y_m$  للمعلمات  $Y_1, Y_2, \dots, Y_m$ ، بمعنى أن

(٢٤ - ٧)

$$z = f(y_1, y_2, \dots, y_m)$$

فإذا افترضنا أن  $z$  قريبة من  $Z$ ، يكون توسيع سلسلة تايلور إلى أحكام الدرجة الأولى في  $Z-z$  ما يلي:

(٢٥ - ٧)

$$z = Z + \sum_{i=1}^m d_i (y_i - Y_i)$$

حيث تكون  $d_i$  المشتقات الجزئية لـ  $z$  فيما يتعلق بالـ  $y_i$ ، بمعنى أن،

$$d_i = \frac{\partial z}{\partial y_i}$$

وهي دالة للتقدير الأساسي  $y_i$ ، هذا يعني أن التباين  $z$  يمكن تقريبه بتباين الدالة الخطية في المعادلة (٢٤ - ٧) أعلاه، التي نعرف كيف نحسبها من الطرق المضبوطة التي قدّمت في الفروع السابقة، بمعنى:

(٢٦ - ٧)

$$v(z) = v\left(\sum_{i=1}^m d_i y_i\right) = \sum_{i=1}^m d_i^2 v(y_i) + \sum_{i \neq j} d_i d_j \text{cov}(y_i, y_j)$$

٦٧ - تنطوي المعادلة (٢٦ - ٧) على مصفوفة تباينات مشتركة هي  $m \times m$  لـ  $m$  تقديرات أساسية  $y_1, y_2, \dots$  بمصطلحات تباين  $m$  ومصطلحات التباين المشترك المائلة، ومقدارها  $(m-1)/2$ ، التي يمكن تقييمها من الطرق المضبوطة للإحصاءات الخطية التي نوقشت في الفروع السابقة.

### المثال ٥ (تباين نسبة)

لتوضيح نهج الضبط الخطي ننظر في تقدير التباين للنسبة التالية:

(٢٧ - ٧)

$$z = r = \frac{y}{x}$$

لاحظ أنه لهذه الحالة  $\frac{\partial r}{\partial x} = -\frac{y}{x^2} = -\frac{r}{x}$  و  $\frac{\partial r}{\partial y} = \frac{1}{x}$ ، ولذلك،

(٢٨ - ٧)

$$v(r) = \frac{1}{x^2} \{v(y) + r^2 v(x) - 2r \text{cov}(y, x)\}$$

وهو التعبير المألوف لتباين النسبة التي توجد في معظم الكتب الدراسية المتعلقة بالمعاينة.

٦٨ - يُستخدم الضبط الخطي على نطاق واسع في الواقع العملي لأنه يمكن تطبيقه على جميع تصاميم المعاينة تقريباً وعلى أية إحصائية يمكن ضبطها خطياً، بمعنى التعبير عنها كأداة خطية للإحصاءات المألوفة، كالتوسطات أو المجاميع، مع مكافئات آتية من الاشتقاقات الجزئية اللازمة في توسيع سلسلة تايلور. وعندما يُضبط تباين التقدير غير الخطي ضبطاً خطياً، يمكن تقريبه باستخدام الطرق المضبوطة الوارد وصفها أعلاه (انظر كوكران (١٩٧٧) ولوهر (١٩٩٩) للاطلاع على التفاصيل التقنية المتعلقة بعملية الضبط الخطي، بما في ذلك أمثلة).

## ٧-٧-٣-١ الميزات

٦٩ - نظراً إلى أن طريقة الضبط الخطي لتقدير التباين مستخدمة منذ زمن بعيد، طُوِّرت نظريته تطويراً جيداً وهي قابلة للتطبيق على فئة من فئات تصميم المعاينة أوسع من الفئات التي تنطبق عليها طرق التكرار (يرد وصفها أدناه). فإذا عُرفت الاشتقاقات الجزئية وكانت مصطلحات التربيع والرتب الأعلى منه في توسيع سلسلة تايلور ضئيلة الحجم، عندئذ سينتج الضبط الخطي تقديراً تقريبياً لتباين ما يقرب من جميع التقديرات الخطية التي هي موضع اهتمام، كالنسب ومكافئات التراجع.

## ٧-٧-٣-٢ أوجه القصور

٧٠ - لا يعمل الضبط الخطي جيداً إلا إذا كانت الافتراضات الواردة أعلاه عن الاشتقاقات الجزئية والرتب الأعلى منها صحيحة. وإلا فربما يسفر العمل عن تحيزات خطيرة في التقديرات. ومن الصعب أيضاً، بوجه عام، تطبيق هذه الطريقة على وظائف مركبة تتعلق بمُعاملات الترجيح ووضع معادلة منفصلة لكل نوع من أنواع المقدرات، وقد يحتاج هذا إلى برامج خاصة. لا يمكن تطبيق هذه الطريقة على إحصائيات كالمتوسط والنسب المئوية الأخرى التي ليست دالات سلسلة لمجاميع السكان أو متوسطات السكان.

٧١ - علاوة على ذلك، من الصعب تطبيق التعديلات المرتبطة بعدم الاستجابة وعدم التغطية مع نهج الضبط الخطي، الذي يتوقف على تصميم العينة، والتقدير الذي هو موضع اهتمام، وترجيح الإجراءات. وهو يتطلب أيضاً إدراج المعلومات المتعلقة بتصميم العينة ( الطبقات، ووحدات المعاينة الأولية، ومُعاملات الترجيح) في ملف البيانات.

## ٧-٧-٤ التكرار

٧٢ - يشمل نهج التكرار فئة من الطرق تنطوي على أخذ عيّنات فرعية متكررة، أو مكرّرات، من البيانات، وإعادة حساب تقدير الاستقصاء المرجّح لكل مكرّر، وإعادة حساب العينة الكاملة، ثم حساب التباين كدالة لانحرافات تقديرات هذه المكرّرات عن تقدير العينة الكاملة. ويمكن تلخيص هذا النهج بالخطوات التالية:

- الخطوة ١ - احذف العيّنات الفرعية المختلفة من العينة الكاملة لتشكّل عيّنات مكررة.
- الخطوة ٢ - أبرز مُعاملات ترجيح مكرّرة لتكرار عملية التقدير لكل عينة مكررة.
- الخطوة ٣ - أبرز تقديراً من العينة الكاملة ومن كل مجموعة مُعاملات الترجيح المكرّرة.
- الخطوة ٤ - احسب تباين التقدير باعتباره مربع انحرافات تقديرات المكرّرات عن تقدير العينة الكاملة.

٧٣ - افترض، مثلاً، أن  $k$  مكرّرات قد أنشئت من عينة، ولكل منها تقديرات مقدارها  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$  لمعلمة  $\theta$ . وافترض أن التقدير المستند إلى العينة الكاملة هو  $\hat{\theta}_0$ . عندئذ يُعطى تقدير التباين المستند إلى التكرار بما يلي:

(٧ - ٢٩)

$$\text{var}(\hat{\theta}) = \frac{1}{c} \sum_{r=1}^k (\hat{\theta}_r - \hat{\theta}_0)^2$$

حيث  $c$  مقدار ثابت يتوقف على طريقة التقدير. وتختلف طرق التكرار فيما يتعلق بقيمة المقدار الثابت وفي الطريقة التي تشكّل بها المكرّرات (انظر الفرع ٧ - ٧ - ٥ لاستعراض موجز لأكثر تقنيات التكرار المستعملة شيوعاً).

### ٧ - ٧ - ٤ - ١ هيكل ملف البيانات

٧٤ - أيّ ما كانت تقنية التكرار، يظل هيكل ملف البيانات ثابتاً على حاله، كما هو مبين في الجدول ٧ - ٩ أدناه.

### الجدول ٧ - ٩ هيكل ملف البيانات لنهج التكرار

السجّل	البيانات	مُعامل ترجيح العيّنة الكاملة	مُعاملات ترجيح المكرّرات			
			١	٢	٣	...
١	البيانات ١	$W_1$	$W_{11}$	صفر	$W_{13}$	...
٢	البيانات ٢	$W_2$	صفر	$W_{22}$	$W_{23}$	...
٣	البيانات ٣	$W_3$	$W_{31}$	$W_{32}$	صفر	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
$N$	البيانات $n$	$W_n$	$W_{n1}$	$W_{n2}$	$W_{n3}$	صفر

ملاحظة: تشير النقاط إلى مسلسلات

### ٧ - ٧ - ٤ - ٢ الميزات

٧٥ - الميزة الرئيسية لنهج التكرار المتصل بنهج الضبط الخطّي هي أن التكرار يستخدم نفس طريقة التقدير الأساسية، بغض النظر عن الإحصائية الجاري تقديرها (لأن تقدير التباين دالّة للعيّنة، لا للتقدير)، بينما تقريب الضبط الخطّي يجب وضعه تحليلاً لكل إحصائية، ويمكن أن يكون ذلك عملاً مُضنياً في استقصاءات الأسر المعيشية الكبيرة، التي يكون فيها عدد كبير من الخصائص التي هي موضع اهتمام. علاوةً على ذلك، استعمال تقنيات التكرار شيء مريح، وهي تنطبق على كل الإحصائيات تقريباً الخطية منها وغير الخطية. باستخدام التكرار، يمكن حساب التقديرات بسهولة للمجموعات السكانية الفرعية ويمكن أن تنعكس آثار عدم الاستجابة وعدم التغطية في مُعاملات ترجيح المكرّرات.

### ٧ - ٧ - ٤ - ٣ أوجه القصور

٧٦ - تعتمد تقنيات التكرار على كثافة استخدام الحاسوب، ويعود معظم السبب في ذلك إلى أنها تحتاج إلى حساب مجموعة من مُعاملات ترجيح المكرّرات، التي هي مُعاملات لترجيح التحليل يعاد حسابها لكل مكرّر يُختار لكي يمثل كل مكرّر على الوجه الملائم نفس المجموعة السكانية التي مثلها العيّنة الكاملة. يمكن أيضاً أن يكون تشكيل المكرّرات معقداً بالقيود التي تُفرض على تصميم العيّنة (انظر الفرع ٧ - ٧ - ٥ أدناه)، وهكذا يمكن أحياناً أن يؤدي إلى الإفراط في تقدير أخطاء المعاينة.

٧٧ - ننهي مقارنتنا العامة بين نهجَي تقدير أخطاء المعاينة، نهج الضبط الخطّي ونهج التكرار. غير أن التحقيقات التجريبية (انظر كيش وفرانكل، ١٩٧٤) أظهرت أن الفوارق بين النتائج التي تنتجها الطريقتان، في حالة العيّنات الكبيرة والإحصائيات العديدة، فروق ضئيلة جداً.

## ٧-٧-٥ بعض تقنيات التكرار

٧٨- أكثر تقنيات التكرار المستخدمة شيوعاً هي:

- (أ) المجموعات العشوائية؛
- (ب) التكرار المعاد المتوازن؛
- (ج) التكرار بطريقة المطواة (JK1 و JK2 و JKn)؛
- (د) سير الحذاء.

سنناقش الآن بإيجاز كل واحدة من هذه التقنيات بدورها.

## ٧-٧-٥-١ المجموعات العشوائية

٧٩- تستتبع تقنية المجموعة العشوائية تقسيم العينة الكاملة إلى مجموعات  $k$  بطريقة تحافظ معها على تصميم الاستقصاء، بمعنى أن كل مجموعة تمثل نسخة صغيرة عن الاستقصاء، تصوّر تصميم العينة. فإذا كانت العينة الكاملة، مثلاً، عينة عشوائية بسيطة حجمها  $n$ ، يمكن إذن تشكيل المجموعات العشوائية بتقسيم الملاحظات  $n$  عشوائياً إلى مجموعات  $k$ ، يبلغ حجم كل واحدة منها  $n/k$ . وإذا كانت عينة مؤلفة من مجموعات، تقسم وحدات المعاينة الأولية على المجموعات  $k$ ، بطريقة تضمن معها أن تحتفظ كل وحدة معاينة أولية بملاحظات، وبذلك تظل كل مجموعة عشوائية عينة مجموعات. وإذا كانت العينة عينة طبقية متعددة المراحل، يمكن تشكيل المجموعات العشوائية باختيار عينة وحدات معاينة أولية من كل طبقة. لاحظ أن مجموع عدد المجموعات العشوائية التي ستشكّل في هذه الحالة لا يمكن أن يتجاوز عدد وحدات المعاينة الأولية المدرجة في العينة في أصغر الطبقات.

٨٠- يمكن استخدام طريقة المجموعة العشوائية بسهولة لتقدير أخطاء المعاينة للإحصاءات الخطية، كالتوسطات والمجاميع ودالاتها السلسلة، وكذلك الإحصائيات غير الخطية كالنسب والنسب المئوية. وليس من الضروري الحصول على برامج حاسوبية خاصة لتقدير أخطاء المعاينة، التي هي مجرد انحرافات معيارية عن التقديرات التي تقوم على أساس المجموعات العشوائية التي شكّلت من العينة الكاملة. غير أن إنشاء مجموعات عشوائية يمكن أن يكون صعباً في تصاميم العينات المركبة، لأن كل مجموعة عشوائية يجب أن تحافظ على هيكل تصميم الاستقصاء نفسه. فإذا كان لدينا تصميم توجد فيه وحدتا معاينة أوليتان لكل طبقة، مثلاً، لا يمكن تشكيل أكثر من مجموعتين عشوائيتين، وبوجه عام، يؤدي العدد الصغير من المجموعات العشوائية إلى تقديرات غير دقيقة لأخطاء المعاينة. وتقضي القاعدة التقريبية العامة بوجود ١٠ مجموعات عشوائية على الأقل للحصول على تقدير أكثر استقراراً لخطأ المعاينة.

## ٧-٧-٥-٢ التكرار المعاد المتوازن

٨١- يفترض التكرار المعاد المتوازن تصميماً فيه وحدتا معاينة أوليتان في كل طبقة. وينطوي تشكيل مكرّر على تقسيم كل طبقة إلى وحدتي معاينة أوليتين واختيار واحدة من الوحدتين من كل طبقة، وفقاً لنمط محدد لتمثل الطبقة بكاملها. يمكن تعديل التقنية لتعمل مع تصاميم أخرى بجمع وحدات المعاينة الأولية في طبقات كاذبة تضم كل واحدة منها وحدتي معاينة أوليتين.

## ٧-٧-٥-٣ المطواة

٨٢ - المطواة، شأنها في ذلك شأن التكرار المعاد المتوازن، تعميمٌ لطريقة المجموعة العشوائية التي تسمح بتراكب المجموعات المكررة. توجد ثلاثة أنواع من تقنيات المطواة: JK1، JK2، وJKn.

٨٣ - الـ JK1 هي الوحدة المنفردة النمطية للمطواة لتصاميم العيّنات العشوائية البسيطة. غير أنه يمكن استخدامها لتصاميم أخرى إذا جُمعت وحدات المعاينة في وحدات فرعية عشوائية تحاكي كل واحدة منها العيّنة الكاملة.

٨٤ - الـ JK2 تشبه التكرار المعاد المتوازن بمعنى أنها تفترض وحدتي معاينة أوليّتين لكل تصميم طبقة. وفي حالة وحدات المعاينة الأولية التي تمثل نفسها، يمكن إنتاج أزواج من وحدات المعاينة الثانوية. يمكن تعديل الـ JK2، شأنها في ذلك شأن التكرار المعاد المتوازن، لإنتاج تصاميم أخرى بتجميع وحدات المعاينة الأولية في طبقات كاذبة في كل طبقة منها وحدتا معاينة أوليّتان. ثم تُسقط إحدى الوحدتين بطريقة عشوائية من كل طبقة بدورها لكي تُشكّل المكرّرات.

٨٥ - الـ JK<sub>n</sub> هي الوحدة المنفردة النمطية للمطواة للتصاميم الطبقيّة. لإنشاء مكرّرات، تُسقط كلُّ وحدة معاينة أوليّة بدورها من كل طبقة. ويُعاد ترجيح وحدات المعاينة الأولية الباقية في الطبقة لتقدير مجموع الطبقة. ويساوي عدد المكرّرات عدد وحدات المعاينة الأولية (أو وحدات المعاينة الأولية الكاذبة).

## ٧-٧-٥-٤ سير الحذاء (الاعتماد على النفس)

٨٦ - تبدأ تقنية الاعتماد على النفس باختيار العيّنة الكاملة التي تنتج من جديد جميع المزايا الهامة للسكان ككل. ثم تُعامل العيّنة الكاملة كما لو كانت هي مجموع السكان، وتؤخذ منها عيّنات فرعية. ويُحصّل على تقدير خطأ المعاينة، كما حصل من قبل، بأخذ الانحراف القياسي للتقديرات بناءً على العيّنات المكرّرة عن التقديرات القائمة على العيّنة الكاملة.

٨٧ - تعمل طريقة الاعتماد على النفس جيداً مع تصاميم العيّنة العامة ومع الوظائف غير السلسلة، كالنقاط المتوية، مثلاً. غير أنها تحتاج إلى حسابات أكثر مما تحتاج إليه تقنيات التكرار.

٨٨ - يحدّد الجدول ٧-١٠، الوارد أدناه، في قيمة المقدار الثابت  $c$  في معادلة التباين (المعادلة (٧-٢٨)) التي تتفق مع مختلف طرق التكرار.

الجدول ٧-١٠

## قيم العامل الثابت في معادلة التباين لمختلف تقنيات التكرار

تقنية التكرار	قيمة العامل الثابت في المعادلة (٧-٢٨)
المجموعة العشوائية	$k(k-1)$
التكرار العشوائي المتوازن	$k$
JK1	١
JK2	٢
JK <sub>n</sub>	$k/(k-1)$
الاعتماد على النفس	$k-1$

## المثال ٦ (المطواة)

نعطي الآن مثالا عددياً بسيطاً ينطبق على طريقة المطواة لتقدير التباين. افترض أن لدينا عيّنة حجمها ٣. نستطيع الآن أن ننشئ ثلاث عيّنات فرعية كل واحدة منها بحجم ٢ بإسقاط وحدة واحدة كل مرة من العيّنة الكاملة. يعرض الجدول ٧ - ١١ أدناه قيم المتغير (Y). للعيّنات الفرعية الثلاث، يشير الرمز "X" إلى أيّ وحدات من العيّنة مشمولة بالعيّنة الفرعية.

الجدول ٧ - ١١

تطبيق طريقة المطواة لتقدير التباين على عيّنة صغيرة وعيّناتها الفرعية

العيّنة الفرعية (g)			
وحدة العيّنة	Y	١	٢
١	٥	X	X
٢	٧	X	X
٣	٩	X	X
مجموع العيّنة	٢١		
متوسط العيّنة	٧	٦	٧

$$s^2 = \frac{(5-7)^2 + (7-7)^2 + (9-7)^2}{3-1} = 4 \text{ تباين العيّنة:}$$

$$\frac{s^2}{n} = \frac{4}{3} \text{ تقدير تباين متوسط العيّنة (تجاهل):}$$

$$\frac{\bar{y}_1 + \bar{y}_2 + \bar{y}_3}{3} = \frac{6+7+8}{3} = 7 \text{ متوسط متوسطات العيّنة الفرعية:}$$

يُعطى تقدير تباين المطواة لمتوسط العيّنة بما يلي:

$$V_f(\bar{y}) = \frac{n-1}{n} \sum_{g=1}^3 (\bar{y}_g - \bar{y})^2 = \frac{3-1}{3} [(6-7)^2 + (7-7)^2 + (8-7)^2] = \frac{4}{3}$$

هذا هو بالضبط نفس التباين المقدر لمتوسط العيّنة، المحسوب أعلاه.

## المثال ٧ (تشكيل المكرّرات)

يستخدم الجدول ٧ - ١٢ بيانات العيّنة ٤ (الفرع ٧ - ٧ - ٢) لتوضيح تكوين العيّنات المكرّرة لمختلف طرق التكرار وكذلك حساب التباينات بطريقة المطواة.

الجدول ٧ - ١٢  
العينة الكاملة: النفقات بحسب الطبقة

الطبقة	وحدة المعاينة الأولى	الأسرة المعيشية	معامل الترجيح $W_{hij}$	النفقات $Y_{hij}$	$W_{hij} * Y_{hij}$
١	١	١	١	٣٠	٣٠
١	١	٢	١	٢٨	٢٨
١	٢	١	٣	١٢	٣٦
١	٢	٢	٣	١٥	٤٥
٢	١	١	٥	٦	٣٠
٢	١	٢	٥	٧	٣٥
٢	٢	١	٢	١٦	٣٢
٢	٢	٢	٢	١٨	٣٦
٣	١	١	٦	٧	٤٢
٣	١	٢	٦	٨	٤٨
٣	٢	١	٤	١٣	٥٢
٣	٢	٢	٤	١٥	٦٠
المجموع			٤٢		٤٧٤

$$\hat{Y}_0 = \frac{\sum_{b=1}^3 \sum_{i=1}^2 \sum_{j=1}^2 W_{bij} Y_{bij}}{\sum_{b=1}^3 \sum_{i=1}^2 \sum_{j=1}^2 W_{bij}} = \frac{474}{42} = 11 = \text{المتوسط المقدر بناءً على العينة الكاملة}$$

الجدول ٧ - ١٣  
طريقة المطواة (أسقط وحدة المعاينة الأولى ٢ من الطبقة ١)

الطبقة	وحدة المعاينة الأولى	الأسرة المعيشية	معامل الترجيح $W_{hij}$	النفقات $Y_{hij}$	$W_{hij} * Y_{hij}$
١	١	١	١	٣٠	٣٠
١	١	٢	١	٢٨	٢٨
٢	١	١	٥	٦	٣٠
٢	١	٢	٥	٧	٣٥
٢	٢	١	٢	١٦	٣٢
٢	٢	٢	٢	١٨	٣٦
٣	١	١	٦	٧	٤٢
٣	١	٢	٦	٨	٤٨
٣	٢	١	٤	١٣	٥٢
٣	٢	٢	٤	١٥	٦٠
المجموع			٣٦		٣٩٣

$$\hat{Y}_1 = \frac{393}{36} = 11 = \text{المتوسط المقدر بناءً على أساس العينة المكررة الواردة أعلاه}$$

يمكننا أن نواصل هذه العملية، بإسقاط وحدة معاينة أولية كل مرة من كل طبقة. يمكن تشكيل ٦ عيّنات مكرّرة بهذه الطريقة. يبين الجدول ٧ - ١٤ أدناه تقديرات متوسط الإيرادات الأسبوعية للأسرة المعيشية بناءً على كل واحدة من العيّنات المكرّرة الست.

الجدول ٧ - ١٤  
التقديرات بناءً على أساس المكرّرات

المكرّرة $j$	وحدة المعاينة الأولى المحذوفة	التقدير $\hat{Y}_j$	$\hat{Y}_j - \hat{Y}_0$	$(\hat{Y}_j - \hat{Y}_0)^2$
١	وحدة المعاينة الأولى ٢، الطبقة ١	١١	صفر	صفر
٢	وحدة المعاينة الأولى ١، الطبقة ١	١٠	١ -	١
٣	وحدة المعاينة الأولى ٢، الطبقة ٢	١٠	١ -	١
٤	وحدة المعاينة الأولى ١، الطبقة ٢	١٢	١	١
٥	وحدة المعاينة الأولى ٢، الطبقة ٣	١٢	١	١
٦	وحدة المعاينة الأولى ١، الطبقة ٣	١٣	٢	٤
المجموع				٨

تقدير المطواة لتباين المتوسط المقدّر يُعطى بواسطة:

$$\text{var}_{JK}(\hat{Y}) = \sum_{h=1}^H \left\{ \frac{n_h - 1}{n_h} \sum_{j=1}^{n_h} (\hat{Y}_j - \hat{Y}_0)^2 \right\} = \frac{1}{2} \times 8 = 4$$

(لاحظ أنه لهذا المثال،  $H=3$  و  $n_h=2$  لكل  $h$ )

٨٩ - نختتم هذا الفرع بإعطاء مثال آخر لتشكيل عيّنات مكرّرة، باستخدام طريقة التكرار المعاد المتوازن. النتائج المبينة في الجدول ٧ - ١٥ أدناه هي لنمط حذف وحدات المعاينة الأولى المحددة في عنوان الجدول.

الجدول ٧ - ١٥

طريقة التكرار المعاد المتوازن (أسقط وحدة المعاينة الأولى ٢ من الطبقة ١ والطبقة ٣؛ ووحدة المعاينة الأولى ١ من الطبقة ٢)

الطبقة	وحدة المعاينة الأولى	الأسرة المعيشية	معامل الترجيح $W_{hij}$	النفقات $Y_{hij}$	$W_{hij} * Y_{hij}$
١	١	١	١	٣٠	٣٠
١	١	٢	١	٢٨	٢٨
٢	٢	١	٢	١٦	٣٢
٢	٢	٢	٢	١٨	٣٦
٣	١	١	٦	٧	٤٢
٣	١	٢	٦	٨	٤٨
المجموع				١٨	٢١٦

$$\hat{Y}_{1,BRR} = \frac{216}{18} = 12 = \text{المتوسط المقدّر بناءً على عيّنة التكرار المعاد المتوازن أعلاه}$$

## ٧ - ٨ مزالق استخدام البرامج الحاسوبية الإحصائية القياسية لتحليل بيانات استقصاءات الأسر المعيشية

٩٠ - التحاليل الملائمة لبيانات استقصاءات الأسر المعيشية تتطلب حساب أخطاء المعاينة في التقديرات بطريقة تأخذ في الحسبان تَعَقُّدُ التصميم الذي ولَّدَ هذه البيانات. ويشمل هذا التقسيم إلى طبقات والتقسيم إلى مجموعات، والمعاينة ذات الاحتمالات غير المتساوية، وعدم الاستجابة، والتعديلات الأخرى لمُعَامِلَاتِ الترجيح في المعاينة (انظر الفصل السادس للاطلاع على التفاصيل المتعلقة بوضع مُعَامِلَاتِ الترجيح وتعديلها). وإن البرامج الحاسوبية الإحصائية العادية لا تفسر هذه التعقيدات لأنها تفترض في الأحوال النمطية أن عناصر العينة اختيرت من المجموعة السكانية بمعاينة عشوائية بسيطة. وإن التقديرات النقطية للمُعَامِلَاتِ السكانية، كما أظهر الفصل السادس، تتأثر بمُعَامِلَاتِ ترجيح العينة المرتبطة بكل ملاحظة. وتتوقف مُعَامِلَاتِ الترجيح هذه على احتمالات الاختيار والسمات الأخرى لتصميم الاستقصاء، مثل التقسيم إلى طبقات والتقسيم إلى مجموعات. وإن هذه البرامج الحاسوبية، عندما تتجاهل مُعَامِلَاتِ ترجيح العيّنات، تسفر عن تقديرات نُقْطِيَّةٍ متحيزة. وإجراء تحليل مرجح بهذه البرامج الحاسوبية يقلل التحيز في التقديرات النقطية إلى حد ما، لكن حتى مع ذلك غالباً ما يكون تقدير أخطاء المعاينة في التقديرات النقطية أقل كثيراً من حجمه الفعلي لأن إجراء تقدير التباين في الأحوال النموذجية لا يأخذ في الحسبان سمات التصميم الأخرى، مثل التقسيم إلى طبقات والتقسيم إلى مجموعات. وهذا يعني أن الاستنتاجات المستمدة من هذه التحاليل ستكون مضللة. فربما يُعَلَنُ خطأً أن الفوارق بين المجموعات، مثلاً، ذات أهمية، أو ربما تُرْفَضُ الافتراضات بطريق الخطأ. ويمكن، مثلاً، أن يكون للاستنتاجات المستمدة من تحليلات بيانات الأسر المعيشية آثار كبيرة على تخصيص الموارد ورسم السياسة على المستويين الوطني والإقليمي.

٩١ - نستخدم الآن مثلاً من بروغان (٢٠٠٤) لتوضيح حقيقة أن استخدام البرامج الحاسوبية الإحصائية المعيارية يمكن أن يؤدي إلى تقديرات نُقْطِيَّةٍ متحيزة، وأخطاء معيارية وفترات ثقة غير ملائمة، واختبارات للأهمية مضللة. يستند المثال إلى مجموعة بيانات من استقصاء بالعينة للتغطية بتحسين التسمم بالكزاز، أُجْرِيَ في بوروندي في سنة ١٩٨٩. كان من أهداف الاستقصاء مقارنة إيجابية المصل (المعرفة بأنها مضاد سُمِّيَّة الكزاز بما لا يقل عن ٠,٠١ وحدات دولية/مليتر مع تاريخ التطعيم ضد التسمم بالكزاز. للاطلاع على مزيد من المعلومات عن منهجية هذا الاستقصاء ونتائجه المنشورة انظر بروغان (٢٠٠٤) والمراجع المشار إليها فيه. يعرض الجدول ٧ - ١٦ أدناه تقديرات للنسبة المئوية من النساء اللاتي أخذن المصل الإيجابي والخطأ المعياري وفترة الثقة المتصلين به.

٩٢ - لاحظ أن التقديرات النقطية هي نفسها لكل البرامج التي تستخدم مُعَامِلَاتِ الترجيح، لكن هناك فرقاً بين التقديرات المرجحة وغير المرجحة. علاوة على ذلك، الأخطاء المعيارية التي تنتجها البرامج الحاسوبية الملائمة تساوي ما يقرب من ضعف الأخطاء التي تنتجها البرامج الحاسوبية المعيارية التي تفترض المعاينة العشوائية البسيطة. بعبارة أخرى، البرامج الحاسوبية المعيارية تستغل تباينات تقديرات الاستقصاء إلى حد خطير. وربما يترك هذا أثراً كبيراً على رسم السياسة. فإذا كان التخطيط جارياً لتدخُّل ما، مثلاً، بناءً على إيجابية المصل بنسبة ٦٥ في المائة أو أقل، عندئذٍ سيُنْفَذُ التدخل نتيجةً للتحليل القائم على أساس البرامج الحاسوبية الخاصة، لكنه ربما لن ينفذ بناءً على تحليل يستخدم البرامج الحاسوبية المعيارية. ويكشف الجدول ٧ - ١٦ أن البرامج الحاسوبية التي تقدر تباينات تقديرات الاستقصاء تقديراً ملائماً تُحَقِّقُ نفس النتائج تقريباً. ونلقي في الفرع التالي نظرة عامة موجزة على بعض البرامج الحاسوبية الإحصائية المتاحة للجمهور، التي تُسْتخدَمُ لتحليل بيانات استقصاءات الأسر المعيشية.

الجدول ٧ - ١٦

استخدام برامج حاسوبية مختلفة لتقدير تباينات تقديرات الاستقصاء، مع نسبة النساء اللائي أبدن إيجابية مصل من بين النساء اللائي وضعن مواليدهن مؤخراً، بوروندي، ١٩٨٨ - ١٩٨٩

البرنامج الحاسوبي	النسبة المئوية لمن أبدن إيجابية مصل	الخطأ المعياري	فترة ثقة بنسبة ٩٥ في المائة
SAS 8.2 متوسطات بلا معاملات ترجيح	٧٤,٩	٢,١	(٧٩,٠ .٧٠,٨)
SAS 8.2 متوسطات بمُعَامِلَاتِ ترجيح	٦٧,٢	٢,٣	(٧١,٧ .٦٢,٧)
SAS 8.2 متوسطات استقصاء	٦٧,٢	٤,٣	(٧٥,٦ .٥٨,٨)
سودان 8.0	٦٧,٢	٤,٣	(٧٥,٦ .٥٨,٨)
ستاتا 7.0	٦٧,٢	٤,٣	(٧٥,٦ .٥٨,٨)
إي بي أي إنفو 6.04d	٦٧,٢	٤,٣	(٧٥,٦ .٥٨,٨)
وسفار 4.1	٦٧,٢	٤,٣	(٧٥,٦ .٥٨,٨)

## ٧ - ٩ برامج حاسوبية لتقدير أخطاء المعاينة

٩٣ - طرق تقدير أخطاء المعاينة المذكورة أعلاه مستخدمة منذ مدة طويلة في البلدان المتقدمة النمو، ونُفِّدَ معظمها بواسطة خوارزميات حاسوبية معدة خصيصاً، طوّرتها وكالات الإحصاء الحكومية، ومعاهد أكاديمية، ومنظمات استقصاء خاصة. وأدى التقدم الذي تحقق مؤخراً في تكنولوجيا الحاسوب إلى وضع بضعة برامج حاسوبية لتنفيذ هذه التقنيات. وكثير من هذه البرامج الحاسوبية متاح الآن لاستخدامه في الحواسيب الشخصية. وما البرامج الحاسوبية إلا نهج واحد فقط من النهج العامة لتقدير التباينات التي بحثت في الفرع ٧ - ٧. وينتج معظم هذه البرامج الحاسوبية التقديرات التي هي أكثر ما تكون استعمالاً، مثل المتوسطات، والنسب، والحصص، ومكافئات التراجع الخطي. وتشتمل بعض البرامج الحاسوبية أيضاً على عمليات تقريب لسلسلة عريضة من المقدّرات، مثل مكافئات التراجع الحسابي.

٩٤ - في هذا الفرع نلقي نظرة عامة موجزة على بعض البرامج الحاسوبية المتاحة للجمهور لتقدير أخطاء المعاينة لبيانات استقصاء الأسر المعيشية. وهذه ليست قائمة جامعة مانعة لكل البرامج ومجموعات البرامج المتاحة. وإنما قصرنا انتباهنا على عدد قليل من مجموعات البرامج الحاسوبية المتاحة في الوقت الحاضر على الحواسيب الشخصية لكي يستخدمها محلّ بيانات الاستقصاء العام. سنستعرض كل مجموعة برامج بإيجاز، ونحدد تصاميم المعاينة التي تنطبق هنا وطرق تقدير التباين. ونبرز أيضاً ميزات ونواقص استخدام هذه البرامج. ولم تُبَدَل أية محاولة لتفصيل الإجراءات التقنية والحاسوبية التي تقوم عليها هذه البرامج. هذه المعلومات يمكن الحصول عليها من مواقع منتجي هذه البرامج على شبكة الإنترنت أو من المراجع المذكورة في آخر هذا الفصل.

٩٥ - البرامج الستة المستعرضة هنا هي: CENVAR (سِنْفَار) و EPI INFO (إي بي أي إنفو)، و PC CARP (بي سي كارب)، و STATA (ستاتا)، و SUDAAN (سودان)، و WESVAR (وسفار). السودان (شاه، وبارنويل، وبيبلر، ١٩٩٨)، وستاتا (ستاتا كوربوريشن، ١٩٩٦)، وبي سي كارب (فولر وآخرون، ١٩٨٩)، وسِنْفَار كلها تستخدم طريقة الضبط الخطّي لتقدير أخطاء المعاينة للإحصاءات غير الخطية. أما وسفار فيستخدم طرق التكرار فقط. وتنفذ النصوص الحديثة من السودان تقنية التكرار المعاد المتوازن (BRR) وتقنية المطواة (Jackknife) أيضاً. وكذلك وضع برنامج SAS (أس أي أس) وبرنامج SPSS (أس بي أس أس) وهما غير مستعرضين هنا) وحدات جديدة

لتحليل بيانات الاستعراضات. وتعرض البرامج القائمة على التكرار كثيراً من الطرق الأساسية ما عدا طريقة سير الحذاء (Bootstrap). نحن نقدم مقارنة موجزة فقط للسمات العامة لهذه البرامج الحاسوبية. فالمقارنة المتقنة تحتاج إلى مقارنات أوسع نطاقاً في استقصاءات معاينة ذات أحجام مختلفة وإحصائيات أكثر كثيراً، وهذه تقع خارج نطاق الاستعراض المحدود الذي أجريناه هنا.

٩٦ - يمكن العثور على روابط الإنترنت الموصلة إلى البرامج الحاسوبية الإحصائية المستعرضة في هذا الفصل، وغيرها كثير، على الموقع التالي: [www.fas.harvard.edu/~stats/survey-soft/survey-soft.html](http://www.fas.harvard.edu/~stats/survey-soft/survey-soft.html).

٩٧ - يعرض بروغان (٢٠٠٤) مقارنة مفصلة بين بضعة برامج حاسوبية إحصائية منها البرامج المستعرضة هنا، بناءً على بيانات مأخوذة من استعراض الأسر المعيشية المُجرى في بوروندي والمذكور أعلاه.

٩٨ - نلقي الآن نظرات عامة موجزة على البرامج الحاسوبية. باستطاعة القراء المهتمين الحصول على مزيد من التفاصيل من الكتب الدليلية الحديثة أو من المواقع المذكورة أدناه على شبكة الإنترنت.

### سِنْفَار CENVAR

مكتب الولايات المتحدة لتعداد السكان؛ اتصل بمركز البرامج الدولية  
United States Bureau of the Census; International Programs Center

Washington, D.C. 20233-8860

البريد الإلكتروني: [IMPS@census.gov](mailto:IMPS@census.gov)

الموقع الإلكتروني على الإنترنت: [www.census.gov/ipc/www/imps](http://www.census.gov/ipc/www/imps)

٩٩ - سِنْفَار عنصر من عناصر نظام برامج حاسوبية إحصائية صمّمه مكتب الولايات المتحدة لتعداد السكان بغية تجهيز وإدارة وتحليل بيانات استقصاء معقد اسمه نظام التجهيز المتكامل على الحاسوب الصغير (IMPS). وهو ينطبق على معظم تصاميم العيّنات، مثل المعاينة العشوائية البسيطة، ومعاينة المجموعات المتعددة المراحل. يستخدم سِنْفَار التقريب بطريقة الضبط الخطي لتقدير التباين.

١٠٠ - من بين التقديرات المنتجة ببرنامج سِنْفَار المتوسطات، والنسب، والمجاميع للعيّنة بكاملها، وكذلك فئات فرعية محددة في تخطيط جدولي. بالإضافة إلى أخطاء المعاينة، يعرض البرنامج مكافئات التنوّع، وآثار التصميم، وأحجام العيّنات غير المرجحة.

### إي بي أي إنفو EPI INFO

مراكز الولايات لمكافحة الأمراض والوقاية منها

مكتب برنامج علم الأمراض الوبائية، ميلستوب C08

Epidemiology Program Office, Mailstop C08

Centers for Disease Control and Prevention,

Atlanta, GA 30333

البريد الإلكتروني: [EpiInfo@cdcl.cdc.gov](mailto:EpiInfo@cdcl.cdc.gov)

الموقع الإلكتروني: <http://www.cdc.gov/epiinfo/>

١٠١ - إي بي أي إنفو نظام برامج حاسوبية إحصائية صمّمته مراكز الولايات المتحدة لمكافحة الأمراض والوقاية منها بغية تجهيز وإدارة وتحليل بيانات علم الأمراض الوبائية، بما في ذلك بيانات الاستطلاعات المركّبة (عنصر CSAMPLE). وتوجد وثائق ذات صلة على شبكة الإنترنت في هذا البرنامج ويمكن طبعها فصلاً ففصلاً. وهي مصمّمة خصيصاً لمعاينة المجموعات المتعددة المراحل المقسمة إلى طبقات بواسطة نموذج المعاينة النهائية للمجموعات.

١٠٢ - ينتج إي بي أي إنفو تقديرات لخطأ المعاينة للمتوسطات والنسب لمجموع العيّنة وكذلك للفئات الفرعية المحددة بتخطيط ذي شعبتين. ويشمل الناتج المطبوع مكرّرات غير مرجحة، ونسباً أو متوسطات مرجحة، وأخطاءً معيارية، وحدود فترة ثقة بنسبة ٩٥ في المائة، وآثار تصميم.

### بي سي كارب PC CARP

جامعة ولاية أيوا Iowa State University

المختبر الإحصائي

Statistical Laboratory

219 Snedecor Hall

Ames, IA 50011

الموقع الإلكتروني: <http://cssm.iastate.edu/software/pccarp.html>

١٠٣ - بي سي كارب برنامج حاسوبي إحصائي وضعته جامعة ولاية أيوا لتقدير الأخطاء المعيارية في المتوسطات، والحصص، والخمسات في المائة، والنسب، والفروق في النسب، وتحليل الجداول الاحتمالية ذات الاتجاهين. البرنامج مصمّم لتناول عيّنات المجموعات المتعددة المراحل والمقسمة في طبقات. يستخدم بي سي كارب نهج الضبط الخطي لتقييم التباين.

### ستاتا STATA

شركة ستاتا Stata Corporation

702 University Drive East

College Station, TX 77840

البريد الإلكتروني: [stata@stata.com](mailto:stata@stata.com)

الموقع الإلكتروني: <http://www.stata.com>

١٠٤ - ستاتا برنامج حاسوبي تحليلي إحصائي صمّم لتقييم أخطاء المعاينة للمتوسطات، والمجاميع، والنسب، والحصص، والتراجع الخطي، والتراجع الحسابي، وإجراءات تحليل بروبيت (probit). وتشمل القدرات الإضافية تقدير المجموعات الخطية للمعلّات والاختبارات الافتراضية، وكذلك تقييم الخمسات في المائة، وتحليل الجدول الاحتمالي، والتعويض عن البيانات المفقودة، والتحليل الأخرى. تستخدم ستاتا نهج الضبط الخطي لتقدير التباينات.

**سودان SUDAAN**

معهد مثلث البحوث Research Triangle Institute  
 مركز البرامج الحاسوبية الإحصائية Statistical Software Center  
 Research Triangle Institute  
 3040 Cornwallis Road  
 Research Triangle Park, NC 27709-2194  
 البريد الإلكتروني: SUDAAN@rti.org

الموقع الإلكتروني: <http://www.rti.org/patents/sudaan.html>

١٠٥ - سودان برنامج حاسوبي إحصائي لتحليل البيانات المترابطة، بما في ذلك بيانات الاستقصاء المركب. وهو ينطبق على تصاميم واسعة التنوع، بما في ذلك المعاينة العشوائية البسيطة، وتصاميم المعاينة المتعددة المراحل. وهو يوفر تسهيلات لتقدير سلسلة من الإحصائيات وما يرتبط بها من أخطاء معاينة، بما في ذلك المتوسطات، والحصص، والنسب، والخمسات في المائة، والجداول المتعددة، والنسب الشاذة، ونماذج تراجع المخاطر الخطية والسوقية والنسبية؛ وتحليل الجداول الاحتياطية. يستخدم البرنامج نهج الضبط الخطي لتقدير التباينات.

**وسفار WESVAR**

وستات إنك. Westat, Inc.  
 1650 Research Blvd  
 Rockville, MD 20850-3129

البريد الإلكتروني: WESVAR@westat.com

الموقع الإلكتروني: <http://www.westat.com/wesvar/>

١٠٦ - وسفار برنامج حاسوبي إحصائي صمّمته وستات إنك. لتحليل بيانات الاستقصاءات المركبة، بما في ذلك تحليل الجداول الاحتياطية، والتراجع، والتراجع الحسابي. وهو ينطبق على معظم تصاميم المعاينة، لكنه صمّم خصيصاً لعينات المجموعات المتعددة المراحل والمقسمة إلى طبقات، بناءً على نموذج معاينة المجموعات النهائية.

١٠٧ - يستخدم وسفار تقنيات التكرار لتقدير التباينات، بما في ذلك المطواة، ونصف العينة المتوازن، وتعديل فاي لطريقة نصف العينة المتوازن. وهو يتطلب إنشاء نص جديد من مجموعة البيانات في شكل وسفار خاص وتحديد معاملات ترجيح مكررة.

**٧ - ١٠ مقارنة عامة للبرامج الحاسوبية**

١٠٨ - البرامج الحاسوبية التي استُعرضت هنا فيها سمات كثيرة مشتركة. فجميع البرامج تحتاج إلى تحديد معاملات الترجيح، والطبقات، ووحدات المعاينة لكل عنصر في عينة. وهي لا تتناول كلها كل تصميم عينة قابل للفهم بطريقة غير متحيزة. فوحدات المعاينة الأولية في معظم تصاميم العينات المتعددة المراحل والمقسمة إلى طبقات، مثلاً، تُختار باحتمالات متناسبة مع الحجم وبلا بديل. وثمة برنامج واحد فقط في القائمة، وهو سودان، فيه سمات مصممة بموجب نموذج اختيار عينة مجموعة نهائية (انظر الفرع ٧ - ٧ أعلاه). علاوة على ذلك،

في السودان أيضاً سمات مصمّمة للقضاء على التباينات للتصاميم التي تستخدم اختيار وحدات المعاينة الأولى بلا بديل. وإن ستاتا هو البرنامج الوحيد الذي فيه سمات تقدير لتفسير الاختيار الطبقي والمتعدد المراحل المستخدم في تصميمه.

١٠٩ - جميع البرامج تقدر تباينات المعاينة والإحصاءات ذات الصلة (آثار التصميم، والترابط داخل الفئة، إلخ.) للمتوسطات، والمجاميع والحصص لمجموع العيّنة، وللصفات الفرعية لمجموع العيّنة، وللفاوت بين الفئات الفرعية. ومعظمها يقدر تباينات المعاينة لإحصاءات التراجع والتراجع الحسابي. وجميعها تقدر الإحصاءات القائمة على تباينات المعاينة التي تنتجها.

١١٠ - سنفار، وإي بي أي إنفو، وبي سي كارب، وسفار متاحة إما بالمجان وإما برسم اسمي. وينبغي للمستخدمين المهتمين أن يستخدموا عناوين البريد الإلكتروني ومعلومات الاتصال الأخرى الموفرة للحصول على مزيد من المعلومات عن كيفية الحصول على البرامج الحاسوبية والوثائق المتصلة بها.

## ٧ - ١١ ملاحظات ختامية

١١١ - ألقى هذا الفصل نظرة عامة موجزة على إجراءات حساب أخطاء المعاينة للتقديرات القائمة على أساس تصاميم العيّنة المعيارية وتصاميم أكثر تعقيداً تستخدم في استقصاءات الأسر المعيشية. ومما يذكر أن حساب أخطاء المعاينة جانب هو غاية في الأهمية من جوانب تحليل النتائج المستمدة من استقصاءات الأسر المعيشية وتقديم تقارير عنها. ومن ناحية مثالية، يجب حساب جميع الخصائص في برنامج الجدولة في الاستقصاء. غير أنه من ناحية عملية، تُعيّن مجموعة خصائص رئيسية، هي موضع اهتمام، لحساب أخطاء المعاينة في كل مجال. ويجب أن تكون الخصائص المختارة هي الخصائص التي تُعتبر هامة فنياً للاستقصاء، لكنها يجب أيضاً أن تتضمن خياراً تمثلياً للبنود التي توجد لها صفات إحصائية، وهي الصفات التي يُعتقد بأنها مجمعة تجمّعاً كبيراً (مثل ذلك أنها منوعات تشير إلى الإثنية أو الحصول على خدمات)؛ والخصائص التي يُعتقد بأن لها آثاراً منخفضة على التجمّع (كالوضع العائلي، مثلاً). بالإضافة إلى ذلك، يجب أن يسترشد الخيار بسمات أخرى، كالخصائص التي تضم نسبة عالية أو منخفضة من السكان، أو مجالات هامة تكون موضع اهتمام.

١١٢ - دعا هذا الفصل أيضاً إلى استخدام برامج حاسوبية لتقدير أخطاء المعاينة لبيانات الاستقصاء. وقد قدمنا أمثلة لأوضاع ارتكبت فيها أخطاء فاحشة في تقدير أخطاء المعاينة عندما تستخدم برامج حاسوبية إحصائية معيارية. وإن استخدام برامج حاسوبية إحصائية معيارية، بوجه عام، لتحليل بيانات استقصاء الأسر المعيشية سوف يقدر التباين الحقيقي لتقديرات الاستقصاء بأقل من قدره الحقيقي. ويمكن أن تؤدي هذه التقديرات الأصغر من الواقع للخطأ المعياري إلى التوصل إلى استنتاجات مضلّة فيما يتعلق بنتائج الاستطلاع، وإلى خطأ في التقييم، كأن تعلن، مثلاً، أنه توجد فروق كبيرة بين متوسطي مجموعتين، أو أن ترفض افتراضاً ويكون الرفض غير صحيح.

١١٣ - قدّم الفصل أيضاً قائمة ببعض البرامج الحاسوبية الإحصائية المتاحة للجمهور، إلى جانب معلومات أساسية للاتصال، وألقى نظرة على تطبيقها. وإن قلة المعرفة أو الدراية بتقدير أخطاء المعاينة واحد من العوائق التي تعوق إجراء تحليل متقدم للبيانات في البلدان النامية. فكثير من المطلقين لا يعرفون الحاجة إلى استخدام برامج حاسوبية متخصصة أو إذا عرفوا، يفضلون أن يتلقوا تدريباً على استخدام برامج حاسوبية جديدة.

١١٤ - ينبغي توكيد أن الفصل يمثل مقدمة فقط لميدانٍ واسعٍ ومتمنامٍ من ميادين تقدير تباين بيانات استقصاء مركّب. ويُشجّع القارئ على النظر في بعض المراجع المشار إليها في نهاية الفصل للاطلاع على معالجة أكثر تفصيلاً وأكثر منهجية للموضوع. وللإطلاع على استعراضٍ أوسع نطاقاً لهذه البرامج الحاسوبية وغيرها، بما في ذلك رمز الحاسوب وإنتاجه لبعض البرامج الحاسوبية المتاحة، انظر بروغان (٢٠٠٤) والمراجع المشار إليها فيه.

١١٥ - وأخيراً، يجب الاعتراف بأن التقدم السريع للتكنولوجيا يؤدي إلى جعل كثير من البرامج الحاسوبية تصبح بعد وقت قصير نسبياً قديمة أو تتحسن تحسناً يفوق الخصائص المذكورة في هذا الاستعراض. بل إن من الممكن أن تصبح بعض الخصائص قديمة في نفس الوقت الذي يُنشر فيه هذا الكتاب. لذلك، من الأهمية بمكان أن نتذكّر أنه يجب الحصول على أدقّ المعلومات عن البرامج الحاسوبية، في وقت استخدامها، من الكتب الدليلية المتعلقة بها أو من مواقعها الإلكترونية.

### المراجع ومواد أخرى للقراءة

- An, A. and D. Watts (2001). New SAS procedures for analysis of sample survey data. *SUGI Paper*, No. 23, Cary, North Carolina: SAS Institute, Inc. Available from <http://support.sas.com/rnd/app/papers/survey.pdf>.
- Binder D.A. (1983). On the variances of asymptotically normal estimators from complex surveys, *International Statistical Review*, Vol. 51, pp. 279-92.
- Brick J.M., and others J. (1996), *A User's Guide to WesVarPC*, Rockville, Maryland: Westat, Inc..
- Brogan, Donna (2005). *Sampling error estimation for survey data*. Household Sample on Surveys in Developing and Transition Countries. In *Studies in Methods*, No. 96. Sales No. E.05.XVII.6.
- Burt, V.L., and S.B. Cohen (1984). A comparison of alternative variance estimation strategies for complex survey data. *Proceedings of the American Statistical Association Survey Research Methods Section*.
- Carlson, B.L., A.E. Johnson and S.B. Cohen (1993). An Evaluation of the Use of Personal Computers for Variance Estimation with Complex Survey Data, *Journal of Official Statistics* 9, No. 4, pp. 795-814
- Dippo, C.S., R.E. Fay and D.H. Morganstein (1984). Computing variances from complex samples with replicate weights. *Proceedings of the American Statistical Association Survey Research Methods Section*.
- Fuller, Wayne, and others (1989). PC CARP: USERS MANUAL, Ames, Iowa: Statistics Laboratory, Iowa State University. Available from <http://cssm.iastate.edu/software>.
- Hansen, M.H., W.N. Hurwitz and W.G. Madow (1953). *Sample Survey Methods and Theory*, Vol. I, *Methods and Applications*. New York: Wiley, Sect. 10.16.
- Hansen M.H., W.G. Madow and B.J. Tepping (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association* vol. 78, No. 384, pp. 776-793
- Kist, Leslie (1965). *Survey Sampling*. New York: John Wiley and sons.
- \_\_\_\_\_ (1995). Leslie Kish: Selected Papers, Steven Heeringa and Graham Kalton, eds. Hoboken, New Jersey: John Wiley & sons, Inc.
- Kish, L., and M.R. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society: services B*, vol. 36, pp. 1-37

- Landis J.R., and others (1982). A statistical methodology for analyzing data from a complex survey: the First National Health and Nutrition Examination Survey. *Vital and Health Statistics*, vol. 2, No. 92. , Washington, D.C.: Department of Health, Education and Welfare.
- Lehtonen, R., and E. J. Pahkinen (1995). *Practical Methods for Design and Analysis of Complex Surveys*. New York: Wiley
- Lepkowski J.M., J.A. Bromberg, and J.R. Landis (1981), A program for the analysis of multivariate categorical data from Complex Sample Surveys. *Proceedings of the American Statistical Association Statistical Computing Section*.
- Levy, Paul S., and Stanley Lemeshow (1999). *Sampling of Populations: Methods and Applications*. 3<sup>rd</sup> ed. New York: John Wiley & Sons.
- Lohr, Sharon (1999). *Sampling: Design and Analysis*. Pacific Grove, California: Duxbury Press.
- Potthoff, R.F., M.A. Woodbury, and K.G. Manton, (1992). "Equivalent sample size" and "equivalent degrees of freedom" refinements for inference using survey weights under superpopulation models. *Journal of the American Statistical Association*, vol. 87, pp. 383-396.
- Rust K. (1985), "Variance Estimation for Complex Estimators in Sample Surveys", *Journal of Official Statistics* 1(4), 381-397.
- Rust, K.F., and J.N.K. Rao (1996). Variance estimation for complex surveys using replication techniques, *Statistical Methods in Medical Research*, vol. 5, pp. 283-310.
- Shah, Babhai V. (1998) Linearization methods of variance estimation. In *Encyclopedia of Biostatistics*, vol. 3, Peter Armitage and Theodore Colton, eds. New York: John Wiley and sons, pp. 2276-2279.
- Shah B.V., B.G. Barnwell and G.S. Bieler, (1996). *SUDAAN User's Manual: Release 7.0*, Research Triangle Park, North Carolina: Research Triangle Institute.
- Tepping B.J. (1968), Variance estimation in complex surveys, *Proceedings of the American Statistical Association Social Statistics Section*, pp. 11-18.
- United Nations (1993). *Sampling errors in household surveys*. UNFPA/UN/INT-92-P80-15E. New York: Statistical Division, Department for Economic and Social Information and Policy Analysis, and National Household Survey Capability Programme.
- Wolter, K.M. (1985), *Introduction to Variance Estimation*. New York: Springer-Verlag.
- Woodruff, R.S. (1971), A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association*, vol. 66, No. 334, pp. 411-414

## الفصل الثامن

# أخطاء غير أخطاء المعاينة في استقصاءات الأسر المعيشية

### ٨ - ١ مقدمة

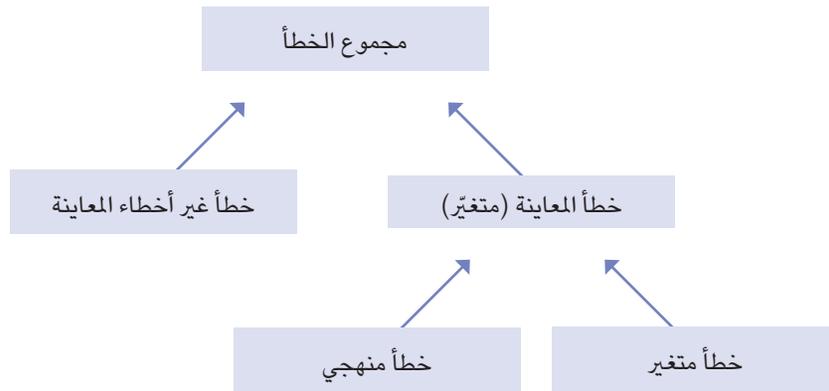
- ١ - تحتاج أخطاء المعاينة وكذلك الأخطاء التي هي من غير أخطاء المعاينة إلى التحكم بها وتخفيضها إلى مستوى لا يهزم وجودها فيه فائدة النتائج النهائية للاستقصاء. في الفصول السابقة المتعلقة بتصميم العينات ومنهجية التقدير أولي اهتمام كثير للتركيز على أخطاء المعاينة، واهتمام أقل، إلى حد ما، على المصادر الأخرى للتباين في الاستقصاءات، مثل عدم الاستجابة وعدم التغطية، التي تشكل فئة الأخطاء المسماة الأخطاء من غير أخطاء المعاينة. هذه الأخطاء الأخيرة خطيرة جداً، لا سيما عندما تكون غير عشوائية بسبب التحيز الذي تُدخله في تقديرات الاستقصاء.
- ٢ - جميع بيانات الاستقصاء عرضة للخطأ من مصادر مختلفة. وإنما التمييز الأساسي العام هو بين الأخطاء الناشئة من عملية القياس وأخطاء المعاينة، أي الأخطاء في تقدير القيم السكانية الناشئة من قياس عينة من السكان.
- ٣ - افترض في الفصول السابقة أن كل وحدة  $Y_i$  في مجموعة سكانية مرتبطة بقيمة  $y_i$  تُعتبر القيمة الحقيقية للوحدة من حيث الخاصية  $y$ . وافترض أيضاً أنه حيثما كانت  $Y_i$  في العينة، فإن قيمة  $y$  التي يُبلغ عنها أو تلاحظ فيها هي  $y_i$ . وهذا يصح في بعض الأوضاع لا في الأوضاع كلها. ففي البلدان التي يوجد فيها تسجيل حيوي وشامل للوقائع الحياتية بواسطة شهادات الميلاد، يمكن الحصول بسهولة على القيم "الحقيقية" عندما تشير  $y_i$  إلى السن. أما في أوضاع أخرى، كالأوضاع التي تشتمل على تقدير نوعي لصحة المرء ربما يكون الحصول على القيم الحقيقية أو حتى تعريفها أصعب كثيراً. فالشخص المريض، مثلاً، رجلاً كان أم امرأة، ربما يقيّم نفسه بأنه في صحة جيدة، وذلك يتوقف على الظروف التي يمر بها.
- ٤ - في الواقع العملي لإجراء الاستقصاءات، لا يصح افتراض أن القيمة التي يُبلغ عنها أو تُلاحظ في الوحدة  $Y_i$ ، هي دائماً  $y_i$ ، بغض النظر عن هوية الشخص الذي يبلغ عنها أو الظروف التي يُحصل فيها على هذه القيمة. فالخبرة الفعلية في إجراء الاستقصاءات تعطي أمثلة عديدة تبين أخطاءً في القياس أو الملاحظة وكذلك أخطاءً من الإجابة الخاطئة، أو عدم الاستجابة، أو أسباب أخرى، كلما أُجري استقصاء.
- ٥ - بالإضافة إلى أخطاء الاستجابة، تكون الاستقصاءات عرضة لأخطاء التغطية والتجهيز، إلخ. وإن جودة مقدّر العينة لمعلمة من معالم السكان دالة لمجموع الخطأ في الاستقصاء، وهو يشمل خطأ المعاينة والخطأ الذي من غير أخطاء المعاينة كليهما. وينشأ خطأ المعاينة، كما أُشير من قبل، نتيجة لأخذ عينة احتمالية بدلاً من إجراء تعداد كامل للسكان، وليس له سبب آخر. أما الخطأ الذي من غير أخطاء المعاينة، من جهة أخرى، فيرتبط بصورة

رئيسية بإجراءات جمع البيانات وإعدادها. يعرض الشكل ٨ - ١ أدناه العلاقة بين أخطاء المعاينة والأخطاء غير أخطاء المعاينة باعتبارها عناصر في خطأ الاستقصاء بأسره.

٦ - لذلك، ينشأ الخطأ من غير أخطاء المعاينة بصورة رئيسية من تعاريف ومفاهيم غير صحيحة، وأطر معاينة غير دقيقة، واستبيانات غير مرضية، وطرق معيبة لجمع البيانات، والجدولة، والتشفير، وتغطية ناقصة لوحدة العينة. هذه الأخطاء لا يمكن التنبؤ بها وليس من السهل التحكم بها. وربما يزداد هذا الخطأ بزيادة حجم العينة، خلافاً لخطأ المعاينة. ويمكن أن يكون الخطأ من غير أخطاء المعاينة، ما لم يتم التحكم به على الوجه الصحيح، أكثر ضرراً من خطأ المعاينة في الاستقصاءات الكبيرة للأسر المعيشية.

الشكل ٨ - ١

العلاقة بين أخطاء المعاينة وأخطاء غير المعاينة باعتبارها عناصر في خطأ الاستقصاء بأسره



## ٨ - ٢ التحيز والخطأ المتغير

٧ - كما يُظهر الجدول ٨ - ١، يمكن تصنيف الأخطاء كأخطاء متغيرة، أو كتحيّز. تنشأ الأخطاء المتغيرة بالدرجة الأولى من أخطاء المعاينة، وإن كان الخطأ من غير أخطاء المعاينة يساهم هو أيضاً في الخطأ المتغير، ويكون ذلك بصورة رئيسية بواسطة عمليات إعداد البيانات كالتشفير والكتابة. خلافاً لذلك، يأتي التحيز بصورة رئيسية نتيجةً لخطأ من غير أخطاء المعاينة ويعود منشؤه إلى عوامل مثل التعاريف غير الصحيحة، وإجراءات القياس الخاطئة، والإجابات الخاطئة، وعدم الاستجابة، والنقص في تغطية المجموعات السكانية المستهدفة، إلخ. ويمكن أيضاً أن تُعزى بعض أنواع التحيز إلى خطأ المعاينة: وتنشأ هذه الأخطاء من حسابات تباينات المعاينة باستخدام تقدير للتباين لا يعكس بطريقة صحيحة تصميم العينة، وبذلك يُسفر عن إفراط في، أو انتقاص من، تقديرات أخطاء المعاينة.

الجدول ٨ - ١

تصنيف أخطاء الاستقصاءات

خطأ المعاينة	الأخطاء المتغيرة
الخطأ من غير أخطاء المعاينة	
خطأ المعاينة	التحيز
الخطأ من غير أخطاء المعاينة	

٨ - يشير التحيز أساساً إلى أخطاء منهجية تؤثر في استقصاء يُجرى بموجب تصميم محدد للاستقصاء بالعينة مع حدوث نفس الخطأ الثابت. وكما جاء ضمناً في الفقرة السابقة، تشكل أخطاء المعاينة في العادة معظم الأخطاء المتغيرة في الاستقصاء، بينما ينشأ التحيز في معظمه من مصادر غير متعلقة بالمعاينة. وهكذا، ينشأ التحيز من أخطاء في التصميم الأساسي للاستقصاء والإجراءات، بينما يحدث الخطأ المتغير بسبب عدم تطبيق تصاميم الاستقصاء وإجراءاته تطبيقاً متسقاً.

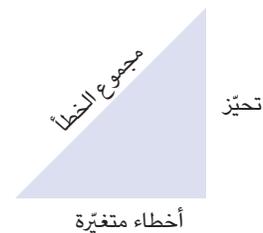
٩ - المصطلح الإحصائي لمجموع خطأ الاستقصاء هو متوسط مربع الخطأ، وهو مساوٍ للتباين زائداً مربع التحيز (انظر الشكل ٨ - ٢). ونقول، لمجرد الجدل، إنه إذا كان التحيز صفرًا، يكون متوسط مربع الخطأ ببساطة نفس تباين التقدير. غير أن التحيز لا يكون صفرًا في استقصاءات الأسر المعيشية. إلا أن من المستحيل تقريباً قياس مجموع التحيز، كما أشير من قبل، ويعود جزء من السبب في ذلك إلى أن حسابه يستدعي معرفة القيمة السكانية الحقيقية، التي تكون في معظم الحالات غير معروفة. ومصادر التحيز كثيرة جداً وطبيعته معقدة جداً، ونادراً ما تُجرى محاولات لتقديره كاملاً.

١٠ - المثلث الوارد في الشكل ٨ - ٢ أدناه يصف مجموع الخطأ والعناصر التي يتكوّن منها. فالضلع القائم للمثلث يمثل التحيز، بينما تمثل قاعدة المثلث الأخطاء المتغيرة. ويعكس كون الوتر مقاساً لمجموع الخطأ مفهوم أن جذر متوسط مربع الخطأ (بمعنى أن مجموع الخطأ) يساوي الجذر التربيعي لتباين المعاينة زائداً مربع التحيز. ومن ثم،

$$(٨ - ١) \quad \text{جذر متوسط مربع الخطأ} = \sqrt{VE^2 + Bias^2}$$

الشكل ٨ - ٢

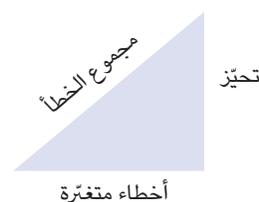
مجموع خطأ الاستقصاء والعناصر التي يتكوّن منها



١١ - فإذا انخفض خطأ التباين أو التحيز، ينخفض مجموع الخطأ وفقاً لذلك. يصور الشكل ٨ - ٣ وضعاً انخفض فيه الخطأ المتغير والتحيز كلاهما انخفاضاً كبيراً. ونتيجة لذلك، انخفض مجموع الخطأ انخفاضاً كبيراً، كما يُرى من طول الوتر مقارنة بطوله في الشكل ٨ - ٢.

الشكل ٨ - ٣

مجموع خطأ الاستقصاء المخفض



١٢ - القصد من تحقيق تصميم جيد للاستقصاء بالعيّنة، مشفوعاً باستراتيجية تنفيذ جيدة، هو تخفيض الخطأ المتغيّر والتحيزّ كليهما، بغية الحصول على نتائج معاينة دقيقة نسبياً.

١٣ - تتحقق الدقة العالية، بوجه عام، من الحصول على عيّنات كبيرة ومصمّمة جيداً، من وجوه أخرى، بينما لا يمكن تحقيق الدقة إلاّ إذا انخفض الخطأ المتغيّر والتحيزّ كلاهما إلى الحد الأدنى. وينطوي هذا القول على التصميم الدقيق، وربما لا يكون، مع ذلك، دقيقاً جداً إن كان فيه تحيزّ كبير. ومن المهم الاعتراف، في هذا السياق، بأن تقديرات الأخطاء المعيارية التي غالباً ما تكون مشمولة بتقارير استقصاءات الأسر المعيشية تستقل من تقدير مجموع خطأ الاستقصاء لأن هذه التقديرات لا تأخذ في الحسبان أثر التحيزّ.

١٤ - من الناحية العملية، يمكن زيادة تقسيم الأخطاء من غير أخطاء المعاينة إلى عنصر متغيّر وأخطاء منهجية، بيمر ولايبيرغ (٢٠٠٣). والأخطاء المنهجية بوجه عام أخطاء غير مُعوّضة، ولذلك تميل إلى الاتفاق (تسير في معظمها في نفس الاتجاه)، بينما الأخطاء المتغيّرة أخطاءً مُعوّضة تميل إلى عدم الاتفاق (يلغي بعضها بعضاً).

## ٨ - ٢ - ١ العنصر المتغيّر

١٥ - ينشأ العنصر المتغيّر من عناصر الخطأ من عوامل الصدفة (العوامل العشوائية) التي تؤثر في عيّنات مختلفة وفي تكرار الاستقصاء. وفي حالة عملية القياس، في وسعنا أن نتصور أن منظور الإجراءات بأسره، الذي يتراوح من اختيار الشخص المقابل إلى جمع البيانات وإعادة البيانات يمكن تكراره باستخدام الإجراءات المحدّدة نفسها، بموجب الشروط المحدّدة نفسها، وباستقلال، دون أن يؤثر تكرار أي شيء في شيء آخر. تتأثر نتائج التكرار بعوامل عشوائية، وكذلك بعوامل منهجية، تنشأ من أحوال بموجبها تُجرى التكرارات وتؤثر في التكرار بنفس الطريقة.

١٦ - عندما تقع الأخطاء المتغيّرة بسبب أخطاء معاينة فقط، يساوي مربع الخطأ المتغيّر تباين المعاينة. ويكون انحراف متوسط قيمة الاستقصاء عن القيمة السكانية الحقيقية هو التحيزّ. ويمكن أن تنشأ الأخطاء المتغيّرة والتحيزّات إما من عمليات المعاينة وإما من عمليات غير المعاينة. ويقاس الخطأ المتغيّر انحراف التقدير عن قيمته المتوقعة ويضم تباين المعاينة وتباين غير المعاينة كليهما. والفرق بين قيمة التقدير المتوقعة وقيمته الحقيقية هو مجموع التحيزّ ويشمل تحيزّ المعاينة وتحيزّ غير المعاينة كليهما.

١٧ - يمكن تقدير الأخطاء المتغيّرة على أساس مقارنات مصمّمة بشكل ملائم بين عمليات تكرار (مكررات) عمليات الاستقصاء بموجب نفس الشروط. ويتوقف انخفاض الأخطاء المتغيّرة على عمل المزيد من شيء ما، مثل زيادة حجم العيّنة أو استخدام عدد أكبر من الأشخاص المقابليين. من جهة أخرى، لا يمكن تخفيض التحيزّ إلاّ بتحسين إجراءات الاستقصاء كإدخال تدابير مراقبة الجودة، مثلاً، في مختلف مراحل عملية الاستقصاء.

## ٨ - ٢ - ٢ الخطأ المنهجي (التحيزّ)

١٨ - يحدث الخطأ المنهجي عندما يكون ثمة ميلٌ، مثلاً، إما إلى الإقلال في المقدار المبلّغ عنه في الاستقصاء وإما إلى المبالغة فيه. ففي بعض المجتمعات التي لا توجد فيها شهادات ميلاد لتسجيل المواليد، يوجد ميل لدى الرجال إلى أن يذكروا في البلاغ عن أعمارهم أعماراً أكبر من أعمارهم الحقيقية. ومن البديهي أن تؤدي هذه الممارسة إلى تحيزّ منهجي — مبالغة في تقدير متوسط العمر لدى الذكور من السكان.

### ٨ - ٢ - ٣ تحيُّز المعاينة

١٩ - ربما تنشأ تحيُّزات المعاينة إما من إجراء معاينة باحتمال محدد يكون غير كافٍ أو خطأً، وإما من طرق تقدير خاطئة. يشمل الأول عيوباً في الأطر، وإجراءات اختيار خاطئة، أو تعداداً جزئياً أو غير كامل لوحات مختارة. يعرض الفصلان الثالث والرابع من هذا الكتاب الدليلي مناقشة مفصلة للظروف العديدة التي يمكن أن ينشأ تباين المعاينة بموجبها من عدم كفاية تنفيذ تصميم العينة — حتى وإن كان تنفيذاً قريباً من الكمال.

### ٨ - ٢ - ٤ مزيد من المقارنة بين التحيُّز والخطأ المتغيِّر

٢٠ - يصعب قياس التحيُّزات بوجه عام، ولذلك نؤكد ضرورة التحكم بها بقوة. ولا يمكن إجراء تقديراتها إلا بمقارنة نتائج الاستقصاء مع مصادر بيانات خارجية موثوقة. ومن جهة أخرى، يمكن تقدير الخطأ المتغيِّر بواسطة مقارنات الأقسام الفرعية للعينة أو بواسطة تكرار الاستقصاء بموجب نفس الشروط. ويمكن تخفيض التحيُّز بتحسين إجراءات الاستقصاء.

٢١ - يقول فيرما (١٩٩١)، إن الأخطاء الآتية من بعض المصادر — منها التغطية، وعدم الاستجابة، واختيار العينات — يبدو أنها تكون بصورة رئيسية على شكل تحيُّز. ومن جهة أخرى، قد تبدو أخطاء التفسير وقيد البيانات إلى حد كبير أخطاء متغيِّرة.

٢٢ - مع أن الأخطاء المنهجية والمتغيِّرة على السواء تخفُّص الدقة والموثوقية بوجه عام، فالتحيُّز أكثر ضرراً بتقديرات مثل متوسط السكان، والحصص، والمجاميع. فهذه التقديرات الخطية مجاميع ملاحظات في العينة. وكما لوحظ من قبل، يمكن تخفيض الأضرار المتغيرة من غير أخطاء المعاينة، شأنها في ذلك شأن أخطاء المعاينة، بزيادة حجم العينة. أما التقديرات غير الخطية، مثل مكافئات الترابط، والأخطاء المعيارية، وتقديرات التراجع، والأخطاء المتغيِّرة والمنهجية، يمكن أن تؤدي إلى تحيُّز خطير (بيمر ولايبيرغ، ٢٠٠٣). والغرض الرئيسي، كما هو الحال في كثير من استقصاءات الأسر المعيشية، هو توفير مقاسات وصفية للسكان، مثل المتوسطات، ومجموع السكان، والحصص؛ لذلك ينصبُّ التوكيد في هذه الحالات على تخفيض الخط المنهجي.

٢٣ - باختصار، ينشأ التحيُّز من نواقص في تصميم الاستقصاء الأساسي وإجراءاته. وإن قياسه أصعب من قياس الخطأ المتغيِّر، ولا يمكن تقديره إلا على أساس المقارنة مع مصادر أكثر موثوقية من خارج الاستقصاء العادي، أو بمعلومات يُحصَل عليها باستخدام إجراءات محسَّنة.

### ٨ - ٣ مصادر الخطأ غير أخطاء المعاينة

٢٤ - الأسباب المختلفة والعديدة للخطأ من غير أخطاء المعاينة موجودة منذ المرحلة الأولى، عندما يكون العمل جارياً على تخطيط الاستقصاء وتصميمه، حتى المرحلة النهائية، عندما يجري إعداد البيانات وتحليلها.

٢٥ - يمكن اعتبار برنامج استقصاء الأسر المعيشية مجموعة من القواعد القوية التي تحدد عمليات مختلفة. تصف هذه القواعد، مثلاً، السكان المستهدفين المراد تغطيتهم، وتحدد مفاهيم المواضيع والتعاريف التي ستستخدم في الاستبيان، وتضع طرق جمع البيانات والقياسات التي ستجرى. وإذا نُفِّذت عمليات الاستقصاء وفقاً للقواعد التي وضعت لها من الممكن نظرياً الحصول على قيمة حقيقية للخاصية التي هي قيد الدرس. غير أن الخطأ من غير أخطاء المعاينة يجعل من هذا وضعاً مثالياً يستحيل تحقيقه.

- ٢٦ - بوجه عام، ربما تنشأ أخطاء غير أخطاء المعاينة من واحد أو أكثر من العوامل التالية:
- (أ) عدم كفاية و/أو عدم اتساق تحديد البيانات فيما يتعلق بأهداف الاستقصاء؛
- (ب) تكرار الوحدات أو حذفها نتيجة لعدم دقة تعريف حدود وحدات المعاينة في المناطق؛
- (ج) التعريف غير الكامل أو غير الصحيح لخصوصيات وحدات المعاينة<sup>١</sup> أو الطرق الخاطئة في إجراء المقابلات؛
- (د) طرق غير ملائمة في إجراء المقابلات، أو الملاحظات، أو القياس، باستخدام استبيانات أو تعاريف أو تعليمات غامضة؛
- (هـ) قلة المقابلين الميدانيين المدربين والذين لديهم خبرة، بما في ذلك قلة الإشراف الميداني الجيد؛
- (و) عدم كفاية تدقيق البيانات الأساسية لتصحيح الأخطاء البديهية؛
- (ز) وقوع أخطاء في عمليات إعداد البيانات، كالتشفير، والكتابة، والتحقق، والجدولة، إلخ؛
- (ح) وقوع أخطاء أثناء عرض النتائج المُجدولة أو نشرها.
- غير أن القائمة الواردة أعلاه ليست جامعة مانعة أبداً.

## ٨ - ٤ العناصر التي تتكوّن منها الأخطاء غير أخطاء المعاينة

٢٧ - يعرف بيمر ولايبيرغ (٢٠٠٣) خمسة عناصر تتكوّن منها أخطاء غير المعاينة، وهي: التحديد، والإطار، وعدم الاستجابة، والقياس، وخطأ في الإعداد. يمكننا أن نضيف الخطأ في التقدير كعنصر آخر ينبغي أن يؤخذ في الاعتبار. سنتناقش أنواع الأخطاء هذه بإيجاز فيما يلي أدناه.

### ٨ - ٤ - ١ خطأ التحديد

٢٨ - يحدث خطأ التحديد عندما يكون المفهوم الذي ينطوي عليه السؤال مختلفاً عن الهيكل الذي يقوم عليه السؤال والمراد قياسه. مثلاً، السؤال البسيط الذي نصه "كم طفلاً لدى شخص معين" يمكن أن يحمل تفسيرات مختلفة في بعض الثقافات. ففي الأسر المعيشية التي تضم أسراً ممتدة، ربما لا يفرّق الجيب على السؤال بين أولاده الذين من صلبه وبين أولاد إخوته وأخواته الذين يعيشون معه في الأسرة المعيشية نفسها. وفي استقصاء للإعاقات، يمكن لسؤال عام لشخص إن كانت لديه إعاقة أم لا أن يُفسّر تفسيرات مختلفة، وذلك يتوقف على مدى شدة الإعاقة أو فهم الجيب لمفهوم الإعاقة. فالأشخاص المصابون بإعاقة صغيرة ربما يفهمون أنهم غير معوقين. وما لم يحدث استعراض صحيح وتوجد أسئلة صافية في الاستبيان ربما لا تكشف الردود عن جميع الأشخاص الذين لديهم إعاقات.

### ٨ - ٤ - ٢ خطأ التغطية أو الإطار

٢٩ - في معظم استقصاءات المناطق، تتألف وحدات المعاينة من مجموعات وحدات جغرافية، مثل مناطق تعداد السكان (انظر الفصل الرابع للاطلاع على مناقشة كاملة للأطر). وليس من غير المألوف أن تقع أخطاء في

<sup>١</sup> لاحظ أنه حتى وإن حدث هذا في عملية اختيار العيّنات، فهو مع ذلك نوع من التحيز في غير المعاينة.

رسم حدود مناطق تعداد السكان أثناء رسم خرائط تعداد السكان. لذلك، يمكن أن تُحدَفَ أسر معيشية أو تُكرَّرَ في إطار المرحلة الثانية. ويمكن لهذه الأخطاء أن تُحدث تحيزاً في تقديرات الاستقصاء في اتجاهين. فإذا لم تكن الوحدات المفروض بها أن توجد في الإطار موجودة فيه فعلاً، يكون احتمال الاختيار من الوحدات غير الموجودة صفرًا، ممَّا يسفر عن تقدير أقل من الواقع. ومن جهة أخرى، إذا تكرر تسجيل بعض الوحدات تكون النتيجة إفراطاً في التغطية ومن ثمَّ إفراطاً في التقدير.

٣٠ - لذلك، ربما تسفر الأخطاء المرتبطة بالإطار عن إفراط في التغطية وإقلال للتغطية. والأخير هو النتيجة الأكثر شيوعاً في الاستطلاعات الكبيرة في معظم البلدان الأفريقية.

٣١ - في استقصاءات الأسر المعيشية المتعددة المراحل، تنطوي المعاينة على عدد من المراحل، من بينها اختيار وحدات المناطق في مرحلة واحدة أو أكثر؛ ووضع قائمة بالأسر المعيشية واختيارها؛ ووضع قائمة بالأشخاص الذين هم داخل الأسر المعيشية واختيارهم (انظر الفصل الثالث). ويمكن أن يقع الخطأ في التغطية في أي من هذه المراحل.

٣٢ - من الأهمية بمكان إعادة توكيد أنه ليس من السهل تقدير أحجام أخطاء التغطية أو آثارها، لأن ذلك يحتاج إلى معلومات خارجة ليس فحسب عن العينة وإنما أيضاً، بحكم التعريف، عن إطار المعاينة.

٣٣ - يشير عدم التغطية، كما هو مذكور ضمناً أعلاه، إلى عدم إدراج بعض وحدات مجموعة سكانية معروفة مشمولة بالاستقصاء في إطار المعاينة (انظر الفصل السادس للاطلاع على مناقشة أكثر لأخطاء التغطية، بما في ذلك عدم التغطية). ولأن احتمال اختيار هذه الوحدات يبلغ صفرًا، فهي مستثناة فعلياً من نتيجة الاستقصاء.

٣٤ - من الأهمية بمكان أن نلاحظ أن استثناء قطاعات من مجموعة سكانية أكبر عمداً وبوضوح من السكان المشمولين بالاستقصاء ليس مشاراً إليه هنا. وإنما أهداف الاستقصاء والصعوبات العملية هي التي تحدد هذه الاستثناءات المقصودة. فاستقصاءات الزواج، المدفوعة بمواقف معينة، ربما تستثني الأشخاص الذين هم دون السن القانونية للزواج من الاستقصاء. وغالباً ما يُستثنى الأشخاص المقيمون في مؤسسات بسبب صعوبات الاستقصاء العملية. وربما يُستثنى سكان المناطق التي تنتشر فيها الألغام الأرضية من البلد للمحافظة على سلامة العاملين في الميدان. وينبغي، عند حساب نسب عدم التغطية ألا يشمل أفراد المجموعات المستثناة عمداً وبوضوح في عداد المجموعة السكانية المستهدفة في الاستقصاء، ولا في عداد الأشخاص الذين لم تشملهم التغطية. وينبغي في هذا الصدد أن يكون تعريف المجموعة السكانية المشمولة بالاستقصاء جزءاً من شروط الاستقصاء الأساسية المحددة بوضوح (انظر الفصل الثالث للاطلاع على مناقشة المجموعة السكانية المستهدفة بالاستقصاء).

٣٥ - يشير مصطلح خطأ التغطية الإجمالي إلى مجموع القيم المطلقة لنسب أخطاء عدم التغطية أو نسب أخطاء الإفراط في التغطية. ويشير الصافي من عدم التغطية إلى زيادة عدم التغطية عن نسبة الإفراط في التغطية. فهو إنَّ المجموع الجبري للعملية. فالتغطية الصافية تقيس إجمالي التغطية فقط إذا لم يكن ثمة إفراط في التغطية. ويتفق معظم العاملين في بحوث الاستقصاءات في أن مشكلة إقلال التغطية في معظم الاستقصاءات الاجتماعية أكثر شيوعاً بكثير من مشكلة الإفراط في التغطية. وإن التصحيحات والترجيحات أكثر صعوبة في حالة عدم التغطية منها في حالة عدم الاستجابة، لأنه لا يمكن الحصول على نسب التغطية من العينة نفسها، ولا يُحصل عليها إلا من مصادر خارجية.

٣٦ - ربما تقع أخطاء عدم التغطية بسبب استخدام أطُر معينة لوحدة المعاينة، كما نوقش مناقشة مستفيضة في الفصل الرابع. وإذا لم يتم تحديث الأطُر واستخدمت الأطُر القديمة لتوفير الوقت والمال فربما يؤدي ذلك إلى

تحيزٌ خطير. فإذا لم تُحدَّث قائمة قديمة لوحدات الأسر المعيشية، في استقصاء الأسر المعيشية، مثلاً، منذ وقت إعدادها الأصلي (الذي يمكن أن يكون قد حدث قبل ١٠ سنين من الاستقصاء الحالي)، لن تكون وحدات الأسر المعيشية، التي أضيفت حديثاً في منطقة التعداد المختارة، جزءاً من إطار وحدات الأسر المعيشية في المرحلة الثانية. وكذلك الوحدات السكنية التي هجرها أهلها ستبقى في الإطار كوحدات فارغة. وفي مثل هذه الحالة سيكون هناك حذف وحدات تخص السكان، وإدراج وحدات لا تخص السكان.

٣٧ - يحدث أحياناً عدم تحديد مواقع بعض الوحدات المدرجة في العيّنة أو عدم زيارتها. تبرز هذه المشكلة أيضاً من استخدام قوائم غير كاملة. بالإضافة إلى ذلك، يمكن أن يحول الطقس أو رداءة وسائل النقل دون الوصول إلى وحدات معيّنة أثناء فترة الاستقصاء المحددة.

٣٨ - كما نوقش في الفصل الثالث، الهدف الذي يكمن وراء استقصاء الأسر المعيشية بالعيّنات هو الحصول على نتائج موضوعية لتيسير التوصل إلى استنتاجات صحيحة عن المجموعة السكانية المراد استهدافها من ملاحظة وحدات في العيّنة. لذلك، يمكن تشويه نتائج الاستقصاء إذا كان مدى عدم التغطية مختلفاً بين المناطق والمناطق الفرعية الجغرافية، كنسبة الذكور إلى الإناث، و الفئات العمرية، والطبقات الإثنية والاجتماعية - الاقتصادية.

٣٩ - أخطاء عدم التغطية تختلف عن عدم الاستجابة. فهذا، الذي نوقش من قبل، ينتج عن عدم الحصول على ملاحظات من بعض وحدات العيّنة بسبب الرفض، أو عدم العثور على العنوان، أو عدم وجود المستجيبين في البيت، أو فقدان الاستبيانات، إلخ. ويقاس مدى عدم الاستجابة من نتائج العيّنة بمقارنة العيّنة المختارة مع العيّنة المتحققة. وكما لوحظ أعلاه، لا يمكن تقدير مدى عدم التغطية، خلافاً لذلك، إلا بنوع من الاختبارات الخارجة عن عملية الاستقصاء.

#### ٨ - ٤ - ٢ - ١ الأخطاء الواقعة في تنفيذ العيّنة

٤٠ - أخطاء التنفيذ في المعاينة تشير إلى مفقودات وتشويهات في إطار المعاينة، مثلاً، الخطأ في تطبيق نسب وإجراءات الاختيار. ويوجد مثل آخر في هذا الصدد، وهو الاستعاضة غير الملائمة في الميدان عن أسر مختارة بأسر أخرى يكون الوصول إليها أو التعاون معها أسير.

#### ٨ - ٤ - ٢ - ٢ تخفيض أخطاء التغطية

٤١ - أنجع الطرق لتخفيض خطأ التغطية هي تحسين الإطار باستبعاد الوحدات الخطأ وعمليات التكرار. وأفضل طريقة لتحقيق ذلك هي ضمان تحديث الأطر القديمة بما فيه الكفاية (انظر الفصل الرابع للاطلاع على مناقشة مفصلة). من المهم أيضاً تحديد مواقع وحدات المعاينة للمنطقة والأسر التي توجد فيها بسهولة. وخير طريقة لتحقيق ذلك أن تكون لديك عمليات رسم خرائط جيدة وجاهزة عندما يُراد وضع الإطار الأصلي، الذي هو في العادة إطار آخر عملية تعداد سكان أجريت.

#### ٨ - ٤ - ٣ عدم الاستجابة

٤٢ - يشير عدم الاستجابة، كما ذُكر مراراً في هذا الدليل، إلى عدم الحصول على ردود من بعض وحدات العيّنة. ومن المفيد التفكير في عيّنة السكان بأنها مقسومة إلى طبقتين، تتألف إحداها من جميع وحدات العيّنة التي حُصِلَ منها على ردود، والأخرى من جميع وحدات العيّنة التي لم يُحصَل منها على أي ردود.

٤٣ - لا يكون عدم الاستجابة، في معظم الحالات، منتشرًا بالتساوي في كل وحدات العينة، وإنما يتركز بكثافة بين مجموعات فرعية. ونتيجة لهذا التفاوت في عدم الاستجابة، ينحرف توزع العينات المحققة من مختلف المجموعات عن توزع العينة المختارة. ويرجح أن يثير هذا الانحراف تحيزًا في عدم الاستجابة إذا كانت متغيرات الاستقصاء متصلة هي أيضاً بالمجموعات الفرعية.

٤٤ - يمكن تقدير نسبة عدم الاستجابة بدقة إذا احتُفظ بنتائج تعداد جميع العناصر المؤهلة التي تقع في العينة. وتُعرف نسبة الاستجابة للاستقصاء بأنها نسبة عدد الاستبيانات المملوءة من وحدات العينة إلى مجموع عدد وحدات العينة المؤهلة<sup>٢</sup> (انظر الفصل السادس أيضاً). يوصى بالإبلاغ عن عدم الاستجابة في جميع البلاغات العمومية الصادرة عن بيانات الاستقصاء ويجب أن يكون الإبلاغ إلزامياً في الاستقصاءات الرسمية. وربما يعود عدم الاستجابة إلى عدم وجود الأشخاص المختارين للعينة في البيت، أو رفضهم المشاركة في الاستقصاء، أو لا يستطيعون الإجابة على الأسئلة بسبب العجز. ويمكن أيضاً أن يحدث عدم الاستجابة نتيجة لفقدان المواعيد أو الاستبيانات، أو لعدم القدرة على إجراء الاستقصاء في مناطق معينة بسبب أحوال الطقس، أو تضاريس الأرض، أو انعدام الأمن. وتشير جميع حالات عدم الاستجابة إلى المستجيبين المؤهلين، ويجب أن تستثنى غير المؤهلين، كما تشير الحاشية رقم ٢. ففي استقصاء الخصوبة البشرية، مثلاً، يتألف السكان المستهدفون في مناطق التعداد من النساء اللاتي هن في سن الحمل والولادة فقط، وبذلك تُستثنى الإناث اللاتي يقعن خارج هذه المجموعة العمرية، ويستثنى جميع الذكور.

٤٥ - يوجد، كما لوحظ في الفصل السادس، نوعان من عدم الاستجابة: عدم استجابة الوحدة وعدم الاستجابة بعدم الرد على بند أو بنود في الاستبيان. عدم استجابة الوحدة يشير إلى عدم الحصول على أي معلومات من وحدة عينة محددة، بينما يشير عدم الاستجابة بعدم الرد على بند أو بنود في الاستبيان إلى وضع يُجمع فيه بعض البيانات لكن ليس كل البيانات من الوحدة. ويُستدل على عدم الاستجابة بعدم الرد على بند أو بنود بوجود فجوات في سجلات البيانات المتعلقة بوحدة العينة المستجيبة، وربما تكون هذه الفجوات ناتجة عن رفض الرد، أو حذف المقابلين لبعض المعلومات، أو العجز. وربما يتأثر رفض المستجيب المحتمل للمشاركة في الاستقصاء بعوامل كثيرة من بينها عدم وجود دافع للمشاركة، أو قلة الوقت، أو الحساسية لأسئلة معينة في الدراسة، إلخ. يذكر غروفز وكوبر (١٩٩٥) عدداً من أسباب الرفض، تشمل السياق الاجتماعي للدراسة، وخصائص المستجيب، وتصميم الاستقصاء (بما في ذلك عبء المستجيب)، وخصائص المقابل والتفاعل بين المقابل والمستجيب. وفيما يتعلق بإشارات محددة إلى عدم الاستجابة بعدم الرد على بند أو بنود، ربما يفهم المستجيب من بعض الأسئلة أنها محرجة، و/أو حساسة، و/أو لا صلة لها بالهدف المعلن. وربما يقفز المقابل عن سؤال أو لا يسجل الجواب عليه. بالإضافة إلى ذلك، يمكن رفض الإجابة أثناء التحرير.

٤٦ - يشير حجم عدم الاستجابة في الوحدة (مجموع غير المستجيبين)، بين عوامل أخرى، إلى مدى القبول العام للاستقصاء وتعقده، وتنظيمه، وإدارته، ومن ثم إلى تعقد بنود معينة مسؤول عنها في الاستبيان أو وضوحها أو مقبوليتها، وإلى جودة عمل المقابل في تناول هذه البنود.

٤٧ - عدم الاستجابة يدخل تحيزاً في نتائج الاستقصاء، وربما يكون هذا التحيز خطيراً في أوضاع لا تكون فيها الوحدات غير المستجيبة "ممثلة" للذين استجابوا، كما هي الحالة عادة. فعدم الاستجابة يزيد أخطاء المعاينة وكذلك، بتخفيضه عدد المستجيبين، الأخطاء من غير أخطاء المعاينة.

<sup>٢</sup> ربما يتبين أن بعض الوحدات المعاينة تقع خارج نطاق الاستقصاء ومن ثم غير مؤهلة، مثال ذلك المنازل الشاغرة أو المحكوم عليها بالهدم، أو التي هجرها سكانها.

٤٨ - غالباً ما تسفر الجهود الرامية إلى زيادة الاستجابة عن تعديلات إجرائية، تشمل اختيار العمليات المشمولة بالاستقصاء. فمن أجل زيادة الاستجابة لاستقصاء خصوبة أجري في زامبيا سنة ١٩٧٨، عُيّنَت مُعلّمت لإجراء مقابلات مع النساء لسؤالهن عن استخدام موانع الحمل، إلخ. وظنَّ المسؤولون أنه لو عُيّنَ شباب ذكور لإجراء المقابلات ستكون نسبة المتنوعات عن الإجابة مرتفعة، لأنه يُحظَر على الشباب الذكور أن يسألوا نساءً أكبر منهم سناً، لا سيما أسئلة تتعلق بالجنس، بما في ذلك منع الحمل.

٤٩ - لا يمكن في الممارسة القضاء تماماً على عدم الاستجابة؛ غير أنه يمكن تقليله إلى الحد الأدنى باستخدام تقنيات الإقناع، وتكرّر الزيارات إلى الأسر التي "لم تكن موجودة في البيت"، وطرق أخرى. انظر الفصلين السادس والتاسع للحصول على معلومات أكثر عن التعامل مع عدم الاستجابة بعدم الرد على بند أو بنود في بيانات الاستقصاء.

## ٨ - ٤ - ٤ أخطاء القياس

٥٠ - تقع هذه الأخطاء عندما يخرج ما يُلاحَظ أو يُقاس عن القيم الفعلية لوحدات العينة. وتركز هذه الأخطاء على المضمون الحافظ للاستقصاء، مثل تعريف أهداف الاستقصاء، وقابليتها للتحوّل إلى أسئلة قابلة للاستعمال، والحصول على الإجابات، وتسجيلها، وتشفيرها وتحضيرها. وتتعلق هذه الأخطاء أيضاً بدقة القياس على مستوى الوحدة الفردية.

٥١ - فوضع خطأ في المرحلة الأولى، مثلاً، أو تعاريف أو مفاهيم مضلّة، في بناء الإطار وتصميم الاستبيان سيؤدي إلى تغطية غير كاملة وتفسيرات متباينة من قبل مقابلين مختلفين، ممّا يسفر عن أخطاء في البيانات المجموعة.

٥٢ - عدم كفاية التعليمات المعطاة للموظفين الميدانيين مصدرٌ آخر من مصادر الخطأ. ففي بعض الاستقصاءات تقتضي التعليمات الغامضة وغير الواضحة من المقابلين أن يستخدموا حكمهم في القيام بعملهم الميداني. ويمكن أن يكون المقابلون أنفسهم مصدرًا للخطأ. ففي بعض الأحيان يمكن أن تكون المعلومات التي تجمع بشأن بند معين للوحدة جميعها معلومات خاطئة؛ ويعود معظم السبب في ذلك إلى عدم كفاية التدريب المعطى للعمال الميدانيين.

٥٣ - الإبلاغ عن العمر في أفريقيا مشكلة قياس شائعة بواسطة تقدير الأرقام وتفضيل العدّ على الأصابع. يمكن أن تُعزى هذه الأخطاء وغيرها في القياس إلى المستجيبين أو إلى المقابلين أو إلى كلا الطرفين. وربما يكون في بعض الأحيان تفاعل بين الطرفين يساهم في زيادة هذه الأخطاء. وربما تكون هناك عيوب في جهاز القياس أو تقنيات القياس تسبّب أخطاءً يمكن ملاحظتها.

٥٤ - يمكن أن يقع المستجيبون في الخطأ بواسطة ما يلي:

- عدم فهمهم لسؤال (أو أسئلة) الاستبيان.
- الإجابة بإهمال إجابات غير صحيحة نتيجة لعدم فهمهم أهداف الاستقصاء، مثلاً، فهماً كافياً؛ ويمكن ألاّ يخصّص المستجيبون وقتاً كافياً، على وجه الخصوص، للتفكير في الأسئلة.
- رغبتهم في "التعاون" بالإجابة على الأسئلة، حتى عندما لا يعرفون الجواب الصحيح.
- ميلهم المقصود إلى إعطاء إجابات خاطئة، مثلاً، في الاستقصاءات التي تتناول مسائل حسّاسة، كالدخل، والأمراض التي يشعر المصاب بها بوصمة إن عُرفت.

- ضعف ذاكرتهم في الحالات التي يُقتضى فيها الرجوع إلى فترة طويلة، كما يحدث في حالة جمع معلومات عن سلع غير معمرة لاستقصاءات النفقات.
- ٥٥ - ربما يكون الأثر التراكمي لأخطاء مختلفة آتية من مصادر مختلفة كبيراً، لأن الأخطاء الآتية من مصادر مختلفة ربما لا يلغي بعضها بعضاً. ويمكن أن يكون صافي أثر هذه الأخطاء تحيُّزاً كبيراً.

### ٨ - ٤ - ٥ أخطاء الإعداد

- ٥٦ - تشمل أخطاء الإعداد، فيما تشمله، ما يلي:

- أخطاء التحرير
- أخطاء التشفير
- أخطاء قيد البيانات
- أخطاء البرمجة

- ٥٧ - تنشأ الأخطاء الواردة أعلاه في مرحلة إعداد البيانات. ففي تشفير الإجابات المفتوحة، المتصلة بالخصائص الاقتصادية، مثلاً، يمكن أن يخرج واضعو الرموز عن الإجراءات المقررة في كتب التشفير الدليلية، ومن ثم يضعون رموزاً غير صحيحة للمهن.

### ٨ - ٤ - ٦ أخطاء التقدير

- ٥٨ - تقع معظم أخطاء التقدير نتيجة لعدم تطبيق المعادلات الصحيحة في حساب مُعاملات الترجيح في الاستقصاء. ويمكن أن تنشأ الأخطاء أيضاً من حساب مُعاملات الترجيح حساباً خاطئاً، حتى عند استعمال المعادلة الصحيحة. وتنشأ تقديرات التباين في المعاينة (أخطاء المعاينة) حين لا يكون جهاز تقدير التباين متفقاً مع التصميم الفعلي للعينة، ومن ثم يخلق أخطاءً في فترات الثقة المرتبطة بتقديرات نقاط الاستقصاء. وفي كل حالة من هذا القبيل تكون النتائج متحيّزة.

### ٨ - ٥ تقدير الأخطاء غير أخطاء المعاينة

- ٥٩ - مصادر الأخطاء غير أخطاء المعاينة متعددة ومتنوعة، كما أُشيرَ بإسهاب في هذا الفصل. نتيجة لذلك، يكاد يكون تقدير مجموع الأخطاء غير أخطاء المعاينة، التي تقع في الاستقصاء، مستحيلًا. غير أن من الممكن دراسة وتقدير بعض العناصر التي يتكوّن منها الخطأ غير أخطاء المعاينة، كما يُناقش أدناه.

### ٨ - ٥ - ١ مراجعات الاتساق

- ٦٠ - ينبغي، لدى تصميم أدوات الاستقصاء (الاستبيانات)، بذل عناية خاصة لكي يشمل الاستبيان بنوداً مساعدة معيّنة من المعلومات، تكون بمثابة اختبار لجودة البيانات المراد جمعها. وإذا كان الحصول على هذه البنود الإضافية ميسوراً، يمكن جمعها بالتجوال في جميع الوحدات المشمولة بالاستقصاء؛ وإلا، يمكن جمعها بالتجوال في عينة فرعية من الوحدات فقط.

٦١ - في حالة إجراء استقصاء بعد تعداد للسكان، مثلاً، حيث تتبع الطريقة القانونية، ربما يكون من المساعد أيضاً جمع المعلومات على أساس الواقع، لكي يتسنى حساب عدد الأشخاص الحاضرين مؤقتاً والأشخاص الغائبين مؤقتاً. ومن شأن مقارنة هذين الرقمين أن تعطي فكرة عن جودة البيانات. وكذلك إدراج بنود تؤدي إلى نسب معينة مستقرة نسبياً، كالنسبة بين الجنسين، يمكن أن تكون مفيدة في تقدير جودة بيانات الاستقصاء.

٦٢ - ينبغي أيضاً استخدام التحقق من الاتساق في مرحلة التحضير للاستقصاء. ويمكن إدخال تفقّد مصادر مختلفة لضمان كون الأشخاص الذين يرمز لهم بأنهم أرباب الأسر لا تقل أعمارهم عن سن محدد مسبقاً أو ألا تكون الإناث اللائي لهن تاريخ في الخصوبة دون سن الثالثة عشرة، مثلاً.

### ٨ - ٥ - ٢ فحص العيّنة/التحقّق منها

٦٣ - توجد طريقة لتقدير بعض أنواع الأخطاء غير أخطاء المعاينة في الاستقصاءات والتحكّم بها، وهي مضاعفة العمل بصورة مستقلة في المراحل المختلفة من العملية، بغية تيسير الكشف عن الأخطاء وتصحيحها. لأسباب عملية، لا يمكن إجراء هذا الكشف المكرّر إلا على عيّنة فقط من العمل باستخدام مجموعة أصغر من الموظفين الخبراء المدربين جيداً. وإذا كانت العيّنة مصمّمة تصميماً صحيحاً وإذا أُجري الفحص بكفاءة، يمكن ليس فقط الكشف عن وجود أخطاء، وإنما يمكن أيضاً أخذ فكرة عن حجمها. وإذا أمكن فحص عمل الاستقصاء كاملاً، يمكن اعتبار جودة النتائج النهائية أحسن كثيراً.

٦٤ - فيما يتعلق بفحص العيّنة، يمكن تنفيذ التصحيح على العيّنة التي فُحصت فقط. ويمكن تخفيض أثر وجه القصور هذا إلى حد ما بتقسيم المخرجات في مراحل مختلفة من الاستقصاء، — بمعنى أن يكون ذلك في جداول معبأة، وجداول مشفّرة، وصحائف حساب، إلخ. — إلى مجموعات وفحص عيّنات من كل مجموعة. وفي الحالات التي تكون فيها نسبة الخطأ في مجموعة معيّنة أكبر من المستوى المحدد، يمكن فحص المجموعة بكاملها وتصحيحها من الأخطاء، وبذلك تُحسّن جودة النتائج النهائية.

### ٨ - ٥ - ٣ فحوص ما بعد الاستقصاء أو فحوص لإعادة المقابلات

٦٥ - ثمة فحص هام للعيّنات يمكن استخدامه لتقدير الأخطاء الواردة في الردود، وهو اختيار عيّنة فرعية، أو عيّنة في حالة تعداد السكان، وإعادة تعدادها من قبل موظفين أفضل تدريباً وأكثر خبرة من الذين استُخدموا في التحقّق الرئيسي. ومن الضروري، لكي يكون هذا النهج ناجحاً، أن يُضمّن ما يلي:

- إجراء التعداد بعد الاستقصاء الرئيسي مباشرة لتقليل الخطأ في إعادة التعداد إلى الحد الأدنى.
- اتخاذ خطوات لتقليل آثار التكيف الذي يمكن أن يتركه الاستقصاء الرئيسي على عمل الفحص اللاحق للاستقصاء.

٦٦ - يُصمّم فحص الاستقصاء عادةً لتيسير تقدير التغطية وأخطاء المضمون كليهما. ويفضّل، لتحقيق هذا الغرض، أن تجرى أولاً إعادة تعداد لجميع الوحدات المشمولة بالعيّنة في المراحل العليا، أي التي تكون على مستوى مناطق التعداد والقرى، بغية اكتشاف الأخطاء في التغطية ثم إعادة استقصاء عيّنة فقط من الوحدات النهائية، ويُضمّن بذلك التمثيل الصحيح لمختلف أجزاء السكان التي تكون لها أهمية خاصة من حيث الأخطاء غير أخطاء المعاينة.

٦٧ - توجد ميزة خاصة لفحص الاستقصاء، وهي أنه ييسر إجراء فحص للوحدات يتألف أولاً من المقارنة بين مجموعتي البيانات التي حُصِلَ عليها من التعدادين للوحدات المشمولة بفحص الاستقصاء، ثم تحليل الاختلافات الفردية التي لوحظت. وعند العثور على تفاوتات تُبدلُ جهود لتعيين هذا السبب واكتساب معرفة متعمقة لطبيعة الأخطاء غير أخطاء المعاينة وأنواعها.

٦٨ - إذا لم يكن في الإمكان إجراء فحص للوحدات نتيجةً لضيق الوقت أو لضيق مالي، يمكن اتخاذ إجراء بديل، لكنه أقل فعالية، يُسمَّى الفحص التجميعي. يتألف هذا النهج من مقارنة تقديرات المَعْلَمَات التي توفَّرها البيانات بالمَعْلَمَات المأخوذة من الاستقصاء الرئيسي. ويعطي الفحص التجميعي فكرة فقط عن صافي الخطأ، وهو حصيلة الأخطاء الإيجابية والسلبية. بينما يعطي الفحص للوحدات، خلافاً لذلك، معلومات عن صافي الخطأ والخطأ الإجمالي.

٦٩ - ينبغي أن تُستخدَم في الفحص اللاحق للاستقصاء نفس المفاهيم والتعاريف التي استُخدمت في الاستقصاء الأصلي.

## ٨ - ٥ - ٤ تقنيات مراقبة الجودة

٧٠ - يوجد مجال واسع لتطبيق تقنيات مراقبة الجودة لاستقصاء العمل، بسبب الحجم الكبير والطبيعة المتكررة للعمليات التي ينطوي عليها هذا العمل. يمكن استخدام خرائط ضبط وتقنيات معاينة للقبول في تقييم جودة البيانات وتحسين دقة النتائج النهائية في الاستقصاءات كبيرة الحجم. ويمكن القول للتوضيح إنه يمكن فحص عمل كل كاتب من كتبة إدخال البيانات، بنسبة ١٠٠ في المائة لفترة زمنية ابتدائية، لكن إذا انخفضت نسبة الخطأ إلى ما دون مستوى معين، يمكن التحقق من دقة عَيِّنة فقط من العمل بعد هذه المرحلة.

## ٨ - ٥ - ٥ دراسة أخطاء الذاكرة

٧١ - تنشأ أخطاء الإجابات، كما ذُكر في موقع سابق من هذا الفصل، نتيجةً لعوامل مختلفة، مثل:

- موقف المستجيب من الاستقصاء
- طريقة المقابلة
- مهارة الشخص المقابل
- خطأ الذاكرة

٧٢ - من بين هذه العوامل، يتطلب خطأ الذاكرة انتباهاً خاصاً لأنه يمثل مشاكل خاصة غالباً ما تتجاوز قدرة المستجيب على التحكم، فهو يتصل بطول الفترة المشار إليها في الجواب والفترة الفاصلة بين الفترة المشار إليها في الجواب وتاريخ الاستقصاء. ويمكن معالجة هذه المسألة الأخيرة باختيار فترة يُجاب عنها تكون قبل فترة فاصلة مناسبة من تاريخ الاستقصاء أو فترة هي أقرب ما تكون إلى هذه الفترة الفاصلة.

٧٣ - من طرق دراسة أخطاء الذاكرة طريقة تقوم على جمع وتحليل البيانات المتصلة بأكثر من فترة واحدة مشار إليها في الجواب، في العَيِّنة أو وحدات العَيِّنة المشمولة بالاستقصاء. المشكلة الرئيسية في هذا النهج هي حدوث قدر كبير من آثار التكيف، ربما تكون عائدة إلى أثر البيانات المبلَّغ عنها لفترة واحدة مشمولة بالإجابة على البيانات

المتصلة بالفترة الأخرى. ولمنع حدوث هذا الأثر يمكن جمع البيانات المتصلة بالفترات المختلفة التي هي قيد النظر من وحدات مختلفة في العيّنة. لاحظ أن من الضروري استخدام عيّنات كبيرة لإجراء هذه المقارنة.

٧٤ - ثمة نهج آخر، وهو جمع معلومات إضافية تمكّن من الحصول على تقديرات لفترات مشمولة بالإجابة. فيمكن في الاستقصاء الديمغرافي، مثلاً، جمع بيانات لا عن سن المستجيب فقط، وإنما أيضاً عن اليوم والشهر والسنة التي وُلِد فيها. ومن شأن الاختلاف في المعلومات أن يكشف عن أيّ خطأ ذاكرة موجود في السن المبلّغ عنها.

## ٨ - ٥ - ٦ العيّنات الفرعية المتداخلة

٧٥ - تشتمل طريقة العيّنات الفرعية المتداخلة على اختيار عيّنتين فرعيتين أو أكثر من العيّنة الإجمالية، ينبغي اختيارهما بطريقة متماثلة وتستطيع كل واحدة منهما إعطاء تقدير صحيح للمعلّمة السكانية. هذه الطريقة تساعد على تقديم تقييم لجودة المعلومات، لأنه يمكن استخدام العيّنات الفرعية المتداخلة لضمان الحصول على معلومات عن أخطاء غير أخطاء المعاينة، كالاختلافات الناشئة عن التفاوت في التحيز من قِبَل الموظفين المقابِلين، والطرق المختلفة لأخذ المعلومات، إلخ.

٧٦ - بعد استقصاء العيّنات الفرعية من قِبَل مجموعات مختلفة من المقابِلين، وإعدادها من قِبَل فرق مختلفة من العاملين في مرحلة الجدولة، من شأن المقارنة بين التقديرات المستندة إلى العيّنات الفرعية أن تعطي فحصاً واسع النطاق لجودة نتائج الاستقصاء. فإذا وُجد، مثلاً، في مقارنة التقديرات المستندة إلى أربع عيّنات فرعية قامت باستقصائها وإعدادها مجموعات مختلفة من العاملين في الاستقصاء، أن ثلاثة تقديرات متقاربة جداً بعضها من بعض، بينما يختلف التقدير الرابع اختلافاً كبيراً عنها، وكان هذا الاختلاف أكبر من أن يُعزى إلى خطأ معاينة، عندئذٍ تصبح العيّنة الفرعية التي فيها الفرق الكبير موضع سؤال.

## ٨ - ٦ ملاحظات ختامية

٧٧ - ينبغي إيلاء الأخطاء من غير أخطاء المعاينة الاهتمام اللازم في استقصاءات الأسر المعيشية لأنها يمكن، إن لم يُسيطر عليها، أن تسبب تحيّزات هائلة في نتائج الاستقصاء. ومما يذكر أن معظم الاستقصاءات تولى انتباهاً قليلاً جداً للسيطرة على الأخطاء من غير أخطاء المعاينة، ويُخشى أن يؤدي هذا إلى نتائج ربما تكون غير موثوقة. وأفضل طريقة للتحكم بالأخطاء من غير أخطاء المعاينة هي اتباع الإجراءات الصحيحة في جميع إجراءات الاستقصاء، ابتداءً من التخطيط، واختيار العيّنات، والاستمرار في ذلك حتى تحليل النتائج. وينبغي، على وجه الخصوص، أن يكون التدريب المكثف المُجرى بعناية للموظفين الميدانيين هو الممارسة العادية المتبعة، وأن تُطرح أسئلة الاستقصاء، لا سيما الأسئلة التي ثبتت صحتها في جهود استقصاءات سابقة، طرْحاً كاملاً.

## المراجع ومواد القراءة الأخرى

Biemer, P., and L. Lyberg (2003). *Introduction to Survey Quality*. Wiley Series in Survey Methodology. Hoboken, New Jersey: Wiley.

Biemer, P. and others, eds. (1991). *Measurement Errors in Surveys*. Wiley Series in Probability and Mathematical Statistics, New York: Wiley.

- Cochran, W. (1963). *Sampling Techniques*, New York: Wiley.
- Groves, R., and M. Couper (1995). Theoretical motivation for post-survey non-response adjustment in household surveys, *Journal of Official Statistics*, Vol. 11, No. 1, pp. 93-106.
- Groves, R., and other eds. (2000). *Survey Non-response*, Wiley-Interscience Publication. New York: John Wiley & Sons, Inc..
- Hansen M., W. Hurwitz and M. Bershad, (2003). Measurement Errors in Censuses and Surveys. *Landmark Papers in Survey Statistics*, IASS Jubilee Commemorative Volume.
- Kalton, G., and S. Heeringa. (2003). *Leslie Kish: Selected Papers*. Wiley Series in Survey Methodology, Hoboken, New Jersey: Wiley.
- Kish, L. (1965), *Survey Sampling*, New York: Wiley.
- Murthy, M. (1967). *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.
- Onsembe, Jason (2003). Improving data quality in the 2000 round of population and housing censuses, Addis Ababa, Ethiopia: UNFPA Country Technical Services Team.
- Raj, D. (1972). *The Design of Sample Surveys*. New York: McGraw-Hill Book Company.
- P. Chandhok (1998). *Sample Survey Theory*. London: Narosa Publishing House.
- Shyam, U. (2004). Turkmenistan Living Standards Survey 2003, Technical Report, National Institute of State Statistics and Information, Ashgabad, Turkmenistan.
- Som, R. (1996). *Practical Sampling Techniques*. New York: Marcel Dekker Inc..
- Sukhatme, P. and others (1984). *Sampling Theory of Surveys with Applications*. Ames Iowa and New Delhi: Iowa State University Press and Indian Society of Agricultural Statistics.
- United Nations (1982). National Household Survey Capability Programme: *nonsampling errors in household surveys*: Sources, assessment and control. DP/UN/INT-81-041/2. New York: United Nations, Statistics Division.
- Verma, V. (1991). *Sampling Methods: Training Handbook*. Tokyo Statistical Institute for Asia and the Pacific.
- Whitfold, D. and J. Banda (2001). Post Enumeration Surveys: Are they Worth it? United Nations Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid –Decade Assessment and Future Prospects, New York, 7-10 August.



## الفصل التاسع

# إعداد البيانات لاستقصاءات الأسر المعيشية

### ٩ - ١ مقدمة

١ - يبحث هذا الفصل إعداد البيانات فيما يتعلق بالاستقصاءات الوطنية للأسر المعيشية. يبدأ الفصل بوصف الدورة العادية لاستقصاء الأسر المعيشية، ثم يفحص الاستعدادات المتخذة لإعداد البيانات كجزء لا يتجزأ من عملية التخطيط للاستقصاء.

٢ - لقد تطورت تكنولوجيا المعلومات بسرعة أثناء العقدين الماضيين أو نحو ذلك. وأثر تطورها بدورها تأثيراً كبيراً في تقنيات تصميم وتنفيذ النظم الإحصائية لعمليات الاستقصاء.

٣ - كان التطور الرئيسي في المعدات الإلكترونية هو التحول من الأجهزة الحاسوبية الكبيرة إلى أجهزة الحاسوب الشخصي. أصبح الحاسوب الشخصي بصورة متزايدة أكثر قوة من حيث سرعة الإعداد وسعة التخزين كليهما. ويمكن الآن أن تؤدي الحواسيب الشخصية أنواعاً مختلفة من عمليات الإعداد فيما يتعلق بالعمليات الإحصائية، تتراوح من الاستقصاءات الصغيرة الحجم إلى عمليات إحصائية كبيرة الحجم، كعمليات تعداد السكان، واستقصاءات الأسر المعيشية ذات العينات الكبيرة جداً.

٤ - بالتوازي مع التطورات التي طرأت على الأجهزة الحاسوبية كانت التحسينات الكبيرة التي أُدخلت على نوعية البرامج الحاسوبية لإعداد البيانات الإحصائية. ومكّن هذا التحسين أيضاً من نقل مهام الإعداد من خبراء الحاسوب إلى المتخصصين في الموضوع نفسه.

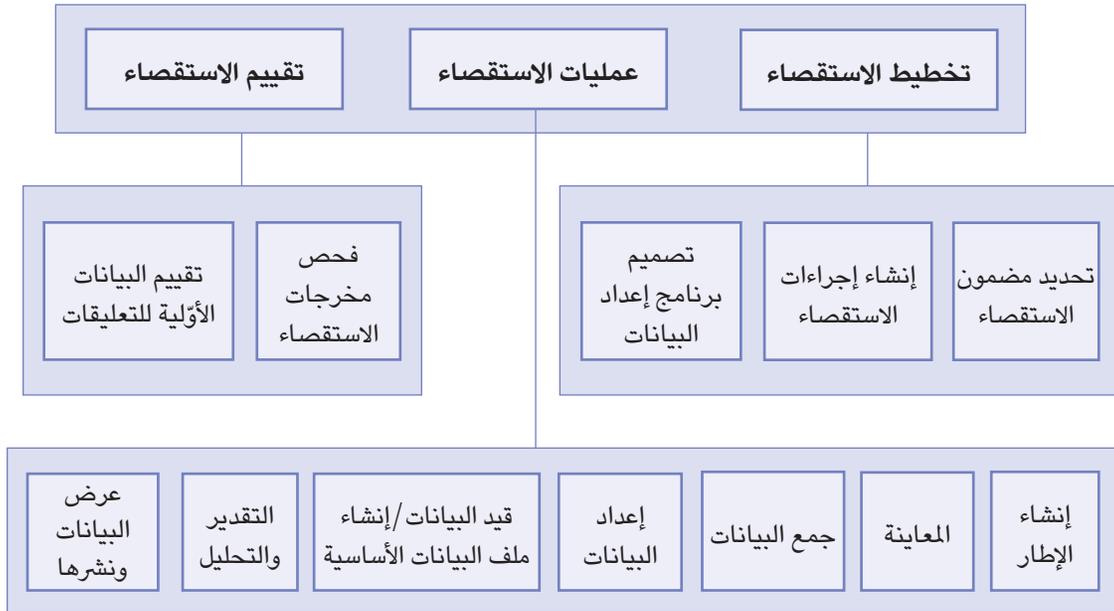
٥ - ظهر على مر السنين عدد من البرامج الحاسوبية لإعداد الاستقصاءات الإحصائية. تتناسب قوى كل واحد من هذه البرامج مع متطلبات كل واحدة من الخطوات المختلفة لإعداد البيانات. وقد يؤدي التذييل اللاحق بهذا الفصل دور مورد لاختيار برامج حاسوبية للخطوات المختلفة في إعداد بيانات الإحصاء؛ ويعطي أيضاً وصفاً لكل واحد من منتجات البرامج الحاسوبية الوارد وصفها في الفصل.

### ٩ - ٢ دورة استقصاء الأسر المعيشية

٦ - يصف الشكل ٩ - ١ الدورة النمطية لاستقصاء الأسر المعيشية. تمر جميع الاستقصاءات، من حيث المبدأ، بنفس نوع الدورات وتكون المراحل النمطية للدورة كالتالي:

الشكل ٩ - ١

## دورة استقصاء الأسر المعيشية



المصدر: سندغرين (١٩٩١).

- تخطيط الاستقصاء: يتخذ مصممو الاستقصاء قرارات بشأن المقاصد والاستعمالات الرئيسية لنتائج الاستقصاء، ومخرجاته الرئيسية ومدخلاته الرئيسية، والإجراءات المستخدمة للحصول على المدخلات (تصميم وإعداد الاستبيان والأدوات المتصلة بالاستقصاء) وتحويلها إلى مخرجات، وتصميم نظام إعداد البيانات والتوثيق.
- عمليات الاستقصاء: تتألف هذه العمليات من إنشاء إطار المعاينة، وتصميم العيّنات واختيارها، وجمع البيانات (القياس)، وإعداد البيانات (إدخال البيانات في الحاسوب، وتشفيرها، وتحريرها، وعزوها)، وملف إنشاء الملاحظة (البيانات الخام الأساسية)، والتقدير، بما في ذلك حساب معاملات الترجيح، وإنشاء المتغيرات المستمدة، وتحليلها، وعرضها، ونشر النتائج.
- تقييم الاستقصاء: يتألف التقييم من فحص وتقييم ما إذا كانت المنتجات النهائية المحددة قد أُنجِزَت، ونُشر الناتج وأُعلن عنه على الوجه الصحيح، ووثِّقت البيانات الأولية وحُزِّنت، إلخ.

٧ - قبل البدء في تصميم وتنفيذ نظام إعداد البيانات لاستقصاء معين، من الأهمية بمكان تصوّر النظام الإجمالي للاستقصاء المعين. وإن التتابع الصحيح للعمليات والإجراءات أيضاً ذو أهمية بالغة للنجاح في تنفيذ أيّ استقصاء للأسر المعيشية. وينبغي لأهداف الاستقصاء أن تحدد، في التتابع المستصوب، تصميم الناتج (كخطة الجدولة، مثلاً، وقواعد البيانات). فهذا بدوره سيملي الأنشطة اللاحقة من تصميم الاستقصاء، وجمع البيانات، وتحضير البيانات وإعدادها، وفي النهاية، تحليل النتائج ونشرها.

٨ - يمكن النظر إلى إعداد البيانات باعتباره عملية تُحوَّلُ بها الإجابات التي حُصِلَ عليها من الاستقصاء، أثناء جمع البيانات، إلى شكل مناسب للجدولة والتحليل. وهذا يستدعي مزيجاً من الأنشطة الآلية واليدوية. ويمكن أن يستغرق وقتاً طويلاً أيضاً ويستنفد موارد كثيرة، ويكون له أثر على جودة الناتج النهائي وتكلفته.

## ٩ - ٣ تخطيط الاستقصاء ونظام إعداد البيانات

### ٩ - ٣ - ١ أهداف الاستقصاء ومضمونه

٩ - ينبغي أن تكون الخطوة الأولى في تصميم أيّ استقصاء، كما نوقشت في الفصل الثاني، صياغة أهدافه الرئيسية وتوثيقها. استقصاءات الأسر المعيشية تعطي معلومات عن الأسر المعيشية في السكان. وهي تُنفَّذُ للإجابة على أسئلة يمكن أن تراود أصحاب المصلحة عن السكان المستهدفين. لذلك، كما تنعكس أهداف أيّ استقصاء بعينه في محاولة الإجابة على هذه الأسئلة، يجب أن يُعطي استبياناً الاستقصاء البيانات ذات الصلة.

١٠ - يمكن، في الأحوال النمطية، تصنيف الأسئلة التي ربما يحتاج أصحاب المصلحة إلى الإجابة عليها بواسطة استخدام بيانات استقصاء للأسر المعيشية، في عدد من الفئات (انظر: غليوي، ٢٠٠٣).

١١ - تسعى مجموعة من مجموعات الأسئلة إلى إثبات الخصائص الأساسية للسكان الذين هم قيد الدراسة (نسبة الفقراء من السكان، نسبة البطالة، إلخ).

١٢ - ثمة مجموعة أخرى من الأسئلة تسعى إلى تقييم أثر التدخلات أو التطورات العامة في خصائص الأسر المعيشية (كنسبة الأسر المعيشية المشاركة في برنامج معين، مثلاً، كيف تقارن خصائصها بخصائص الأسر غير المشاركة في البرنامج، وما إذا كانت أحوال معيشة الأسر آخذة في التحسن أو التدهور على مر الزمن، إلخ).

١٣ - وأخيراً توجد فئة الأسئلة المتعلقة بالعوامل التي تحدد ظروف الأسر المعيشية وخصائصها أو العلاقات فيما بينها (بمعنى أسئلة عما يحدث ولماذا يحدث، إلخ).

### ٩ - ٣ - ٢ إجراءات الاستقصاء وأدواته

#### ٩ - ٣ - ١ - ٢ خطط الجدولة والنواتج المتوقعة

١٤ - من التقنيات المفيدة لمساعدة مصمّم الاستقصاء على تحقيق الدقة في احتياجات المستخدمين إلى المعلومات إنتاج خطط جدولة وجداول صورية. الجداول الصورية هي مسودات جداول تشمل كل شيء ما عدا البيانات الفعلية. وينبغي أن تبيّن مسودة الجدول الخطوط العريضة للجدول، كحدّ أدنى، عناوين الجدول، وعناوين الأعمدة، وتبين المتغيرات المهنية المراد جدولتها، ومتغيرات الخلفية المراد استخدامها للتصنيف، والفئات السكانية (أهداف الاستقصاء، أو عناصره، أو وحداته) التي تنطبق عليها الجداول المختلفة (انظر الفصل الثاني). ومن المستصوب أيضاً بيان فئات التصنيف بأكبر قدر ممكن من التفاصيل، مع أن في الإمكان تعديل هذه الفئات في وقت لاحق، عندما تتوفر معرفة أفضل بتوزيع العيّينات على فئات المستجيبين.

١٥ - يمكن رؤية أهمية خطة الجدولة من عدد المنظورات. فعدد الجداول الصورية يبيّن إن كانت البيانات التي ستُجمَع ستؤدي إلى وضع جداول يمكن استخدامها. ولن تبين ما هو غير موجود فحسب، وإنما ستكشف

- أيضاً عما هو زائد عن الحاجة. علاوة على ذلك، سيُعوّض الوقت الذي يُصرّف في وضع الجداول الصورية بتوفير وقت أطول منه في مراحل جدولة البيانات، بتخفيض الوقت الذي يُصرّف في تصميم وإنتاج الجداول الفعلية.
- ١٦ - توجد أيضاً علاقة وثيقة بين خطة الجدولة وتصميم العيّنات المستخدم في الاستقصاء. فالتجزئة الجغرافية للجداول، مثلاً، لا يمكن إجراؤها إلا إذا كانت العيّنة مصمّمة بشكل يسمح بإجراء هذه التجزئة.
- ١٧ - تعطي الأمم المتحدة (١٩٨٢) وصفاً أكثر تفصيلاً لما تستدعيه خطة الجدولة ولمختلف فوائدها.
- ١٨ - القصد ممّا ذُكر أعلاه هو توكيد أهمية الدور الذي يمكن أن تؤديه خطة الجدولة فيما يتعلق بالتخطيط الفعّال للاستقصاء المحدد ونظام إعداد البيانات المتعلق به. غير أن من المهم توكيد أن خطة الجدولة لا تمثل إلا العمود الفقري لبعض النواتج التي يمكن توقُّع تحقيقها من الاستقصاء ذي الصلة. وتوجد في استقصاءات الأسر المعيشية إمكانية لتوليد ثروة من المعلومات. ويمكن رؤية مجموعة البيانات الدقيقة المنظّفة المستمدة من الاستقصاء بأنها الناتج الرئيسي والأساسي. وغالباً ما تحتاج مجموعة البيانات الدقيقة هذه إلى أن توضع في حزمة وتُتاح لأصحاب المصلحة في شكل مقبول لدى المستخدمين وبواسطة قنوات توزيع ملائمة.

### ٩-٣-٢-٢ تصميم الشكل وطباعته

- ١٩ - عندما تُحدّد أهداف الاستقصاء وخطة جدولته يمكن تطوير الاستبيان (أو الاستبيانات) ذات الصلة. يؤدي الاستبيان دوراً مركزياً في عملية الاستقصاء، التي تُنقل فيها المعلومات ممن توجد لديهم إلى من يحتاجون إليها (المستخدمين). فهي الوسيلة التي يُعرب فيها عن احتياجات المستخدمين بطرق عملية وهي كذلك الأساس الرئيسي للمُدخلات في نظام إعداد البيانات.
- ٢٠ - يستفاد من لندل (٢٠٠٣) أن لتصميم الاستبيان المراد استخدامه أثناء تعداد السكان آثاراً في التقاط البيانات والعكس صحيح أيضاً. فإذا وُجد أن تقنيات مسح الوثائق مفيدة لالتقاط البيانات، سيلزم استخدام تصاميم ذات شكل محدد؛ وسوف تكون هذه الأشكال مختلفة، وذلك يتوقف على مسألة ما إذا كان قد تقرّر شكل قيد البيانات يدوياً من لوحة المفاتيح إلى قرص الحاسوب.
- ٢١ - بغضّ النظر عن تقنية قيد البيانات يجب أن يكون كل استبيان قابلاً للتعريف بشكل منفرد. ويجب أن تُطبع بطاقة تعريف ذات شكل فريد على كل استمارة. ونظراً إلى أن إساءة تفسير بطاقة تعريف الاستمارة ربما تسفر عن تكرار القيود وعن مشاكل أخرى، ينبغي اتخاذ تدابير لتقليل المخاطر إلى الحد الأدنى. ومن البديهي أن تكون شفرات الخطوط المتوازية أفضل خيار لدى استخدام تقنيات المسح، وإذا أُخذ في الاعتبار قيد البيانات يدوياً يجب، مع ذلك، اتخاذ تدابير لكي تحتوي الاستمارة على فحص رقمي لمنع القيود غير الصحيحة. وينبغي أن تُعرّف شفرة الخطوط المتوازية كل استبيان تعريفاً فريداً ويجب دائماً أن تكون رقمية. وفي الأحوال النمطية تكون المعلومات المتعلقة بتحديد مُعاملات الترجيح في العيّنات أو توسيع نطاق العوامل (الطبقات، ووحدات المعاينة الأولية، وقطاعات المناطق، والتمييز بين المناطق الإدارية اللازمة للجدولة، إلخ).
- ٢٢ - توضع الاستمارات عادةً في كُتب (كتاب لكل منطقة أو حي تعداد، مثلاً، إلخ). ويجب أن يكون كل كتاب، شأنه في ذلك شأن الاستمارات، قابلاً للتعين بصورة فريدة، ويجب أن توجد علاقة محدّدة بوضوح بين كل كتاب والاستمارات الموجودة فيه، لضمان كون الاستمارة "س" تخص الكتاب "ص" دون غيره بشكل دائم. وستُستخدَم معلومات تعريف الكتب في جميع مرحلة إعداد البيانات، بدءاً ببيان وصول كتاب من الميدان إلى غرفة تصحيح الاستبيانات، وتستمر حتى نهاية استرداد استمارة ما عند اللزوم، مثلاً، لفحص شيء ما أثناء عملية

جدولة البيانات أو تحليلها. لذلك، يجب تقليل مخاطر إساءة تفسير معلومات تعريف كتاب معين، كإساءة تفسير معلومات تعريف استمارة معينة، إلى الحد الأدنى.

٢٣ - يجب تصميم كل ميدان تصميمياً يستوعب معه أكبر عدد ممكن من الرموز كمتغيرات؛ فيجب، مثلاً، أن يكون أكبر عدد ممكن من أفراد الأسر المعيشية واضحاً وضوحاً تاماً بغية وضع الحجم الصحيح للميدان.

٢٤ - من الأهمية بمكان أن نضمن عدم وجود أي عيب في تعريف وحدات الملاحظة، وأنماط التجاوز، والجوانب الأخرى للاستبيان. فكل استقصاء للأسر المعيشية يجمع معلومات عن وحدة إحصائية رئيسية (الهدف الرئيسي) — الأسرة المعيشية؛ وكذلك عن مجموعة مختلفة من الوحدات الثانوية (الأهداف المرتبطة) داخل الأسرة المعيشية — أشخاص، بنود ميزانية، وحدات الأراضي الزراعية، المحاصيل، إلخ. ويجب أن يكون الاستبيان واضحاً وصريحاً في بيان ماهية هذه الوحدات ويجب أيضاً أن يضمن أن توضع لكل وحدة ملاحظة بطاقة تعريف صحيحة كعريف فريد لها. ثمة طريقة متبعة في العادة لتعريف الأسر المعيشية، وهي تشكل جانباً هاماً أيضاً من جوانب نظام القيد اليدوي للبيانات، وهي استخدام رقم مسلسل بسيط يُكتب على صفحة غلاف الاستبيان، أو تطبعه المطبعة مسبقاً. ويمثل الرقم المسلسل في العادة تعريف الاستمارة.

٢٥ - نظراً إلى أن المسح الضوئي للصور وإعدادها كوسيلة سريعة لالتقاط البيانات يُستخدم الآن بصورة متزايدة، من المهم أيضاً تقديم بعض المناقشة للجوانب الخاصة المتصلة بتصميم الاستبيان على نحو يكون معه قابلاً للمسح الضوئي. ومن بين المسائل التي تظهر عند تصميم الاستبيان تصميمياً يسمح بإعداده بجهاز مسح للصور — بوسائل منها، على سبيل المثال: برامج حاسوبية لمعرفة الأرقام والحروف بصرياً، أو معرفة الأرقام والحروف ذهنياً، أو قراءة العلامات بصرياً، وترد مناقشة هذه المسائل أدناه.

٢٦ - يُستخدم في الاستبيان شفرتان اثنتان من شفرات الخطوط المتوازية. الأولى منهما هي الشفرة التي تعرّف كل صفحة من صفحات الاستبيان، وهذه مسألة هامة جداً، لا سيما حين تكون الصفحات متشابهة جداً في تصميمها وشكلها. وهذه هي الوسيلة الرئيسية التي يميز بها جهاز المسح الضوئي الحاسوبي بين الصفحات المختلفة في الاستبيان. وتوضع شفرة ثانية ذات خطوط متوازية، مع التفسير المرتبط بالصفحة عادة، على كل صفحة من صفحات الاستبيان، وتوضع هذه الشفرة نفسها بالضبط على كل الصفحات، لكنها تختلف من حيث الترتيب المسلسل عن نظيرتها في كل استبيان لاحق. هذه الشفرة تجمع صفحات الاستبيان معاً، لأن إعداد الاستبيان للمسح يفصل الصفحات بعضها عن بعض، وهي لذلك شفرة هامة جداً.

٢٧ - التصميم الدقيق لنطاقات الشكل المطبوع يمثل شكل قاموس البيانات للبيانات المراد جمعها من الاستقصاء. فإذا كان القصد وضع رمز لتعداد المقاطعة مؤلف من خمسة أرقام، مثلاً، يصمّم نطاق مطبوع محدد فيه مربع مقسم إلى خمس خانات في الاستبيان المطبوع، يوضع فيه رمز تعداد المقاطعة، ويجب وضعه في هذا المربع. وإذا كانت الشاشة مصمّمة لقيد البيانات يدوياً لا تكون هذه الدقة في طباعة الاستبيان ضرورية، وإنما يكون كل ما هو مطلوب وضع مربع يتسع لخمس أرقام. غير أن من المهم تهذيب تصميم نظام قيد البيانات بحيث يُبين النطاق برقم ويُملأ تلقائياً بغية تجنّب خروج الرقم المدخل عن مواعمة الترتيب أو إساءة تفسير البيانات.

٢٨ - يختلف تصميم الاستمارات لقيد البيانات اختلافاً كبيراً بحسب وسيلة القيد: أي بالمسح الضوئي أم باستخدام الآلة الكاتبة؟ وذلك لأن المسح الضوئي يعتمد للتعريف اعتماداً كلياً على مواقع نطاقات البيانات. ولا يحتاج قيد البيانات بالمسح، خلافاً للقيد اليدوي، إلى طباعة تعاريف النطاقات على الاستبيان، فيما عدا شيئاً من تعديل النطاقات على كل صفحة. ولإيجاد مستوى عال من الدقة في تفسير الصور بالمسح الضوئي ينبغي ألا

يكون النطاق صغيراً جداً ولا كبيراً جداً. وينبغي تشجيع المقابلين، في تدريبهم وكذلك أثناء إدارة العمل الميداني، على كتابة الأرقام والرموز بوضوح في منتصف نطاق البيانات.

٢٩ - يوجد أيضاً اختلاف أساسي فيما يتعلق بطريقة تصميم الاستبيان لإدخال رموز المهنة، مثلاً، بين عمليات الترميز اليدوي، المناسبة لقيد البيانات يدوياً بالآلة الكاتبة والتشفير على الشاشة باستخدام جدول تُستمدُّ منه المعلومات مطبوع بمساعدة الحاسوب، يحتوي على سجل رموز المهن على الشاشة. وعندما يصمّم الاستبيان لقيد البيانات بالآلة الكاتبة يُطبع رمز التصميم على الاستبيان لكي يستخدمه الشخص المسؤول عن تشفير الاستبيان بعد استعراض الإجابة المفتوحة على السؤال الذي نصه "ما هي مهنتك؟". وفي حالة الاستبيان المراد مسحه ضوئياً لا توجد ضرورة مطلقة لتخصيص فراغ على وجه الاستبيان المطبوع لهذا الغرض، لأن كتاب الرموز يشكّل جزءاً من القالب المطبوع المصمّم على الحاسوب وتُختصر أدوار واضح الرموز وكاتب قيد البيانات إلى دور مُحَقِّق واحد. في وقت قيد بيانات رمز المهنة، يُعرض المحقق مع الإجابة المفتوحة من الصورة المسوحة ضوئياً، وتُعرض قائمة تتدلى على الشاشة برموز المهن المفهرسة للاختيار منها بسرعة.

٣٠ - عندما يُنظر في استخدام تقنيات المسح الضوئي تصبح جودة طباعة الاستبيان أيضاً مسألة هامة. فجهاز المسح الضوئي أكثر حساسية للأخطاء التي تقع أثناء الطباعة من عين الإنسان. وقد تحدث مشاكل، مثلاً، بسبب استخدام ألوان معينة أو مجموعات بعض الألوان، وتفاوتات في تدرُّج اللون ووضوحه، ومطبوعات مائلة أو مفقودة، وأخطاء في التقييم التلقائي للصفحات، وأخطاء في التجليد.

٣١ - يعرض كاثرين (٢٠٠٣) ولندل (٢٠٠٣) بعض المسائل الرئيسية فيما يتعلق باستخدام تكنولوجيا المسح الضوئي لإعداد الاستقصاءات الإحصائية وعمليات تعداد السكان.

### ٩ - ٣ - ٣ تصميم نُظْم إعداد البيانات في استقصاءات الأسر المعيشية

#### ٩ - ٣ - ٣ النهج المعمّم لإعداد البيانات في استقصاءات الأسر المعيشية

٣٢ - تصميم النُظْم واحد من الأنشطة الرئيسية اللازمة في تخطيط استقصاء الأسر المعيشية. وما يلزم أساساً، في هذه الخطوة، هو تحديد بيانات الاستقصاء المراد جمعها ونظام إعداد البيانات بأسره وفقاً لمشروع ما مننظّم رسمياً.

٣٣ - يبحث جامبوا، وباريرنياوتا وروزين (١٩٨٩) الفوائد التي يعود بها اعتماد واستخدام مشروع مننظّم رسمياً لتصميم جميع هذه النُظْم وتطويرها وتوثيقها داخل الوكالة الإحصائية، لا سيما لاستقصاءات الأسر المعيشية وفقاً للأحكام التالية:

(أ) يمكن أن يكون المشروع هو البرنامج المشترك للتعاون اللازم بين الإحصائيين، وخبراء الموضوع ومحلي النُظْم/المبرمجين؛

(ب) ستوصف جميع عمليات الاستقصاء بوضوح وتوثق، ويمكن توفير مراجع لها في مرحلة لاحقة. وستكون الوثائق الناتجة (بمعنى مجموعة البيانات الأولية) ذات أهمية لتطوير وصيانة النُظْم ذات الصلة لإنتاج الإحصائيات؛

- (ج) تميل تكاليف تطوير النظم وصيانتها إلى أن تكون باهظة في الوكالة الإحصائية، نظراً إلى كثرة النظم المختلفة التي تكون لازمة في العادة. ويميل هيكل استقصاء الأسر المعيشية إلى أتباع نفس النمط ونفس المبادئ. فهي تميل، مثلاً، إلى الاشتراك في نفس هيكل الملفات والبيانات، ونظم الترميز، إلخ. لذلك، يمكن أن تستفيد الاستقصاءات اللاحقة من نظم إعداد البيانات التي وُضعت للاستقصاءات السابقة، ويتوقع أن يخفض هذا تكاليف تطويرها وصيانتها؛
- (د) سيكون اعتماد نهج منظم رسمياً مسألة هامة أيضاً لتكامل الاستقصاء إذا كانت ثمة رغبة، مثلاً، في إجراء بعض التحاليل المجتمعة لبيانات من استقصاءات مختلفة أو من جولات مختلفة من الاستقصاء نفسه.

### ٩-٣-٣-٢ السمات العامة لمخطط رسمي لتصميم النظم

٣٤ - يعطي هذا الجزء إشارة إلى بعض الجوانب العامة والأساسية للمخطط الرسمي المذكور أعلاه.

#### هيكل البيانات

٣٥ - إن القرارات المتعلقة بالمسألة الاجتماعية المراد تحليلها، والبيانات المراد استخدامها، والتقنية الإحصائية المراد تطبيقها كلها أساسية للتحليل الجيد. غير أن ثمة سؤالاً أكثر أساسية حتى من ذلك، وهو تعيين وتعريف مواضيع الاستقصاء أو وحدات التحليل. وقد تم توكيد هذه المسألة في الجزء المعني بتصميم الشكل وطباعته. وينبغي أثناء تصميم نظام إعداد البيانات للاستقصاء إجراء وصف أكثر رسمية وتفصيلاً لوحدة (موضوع) التحليل والمتغيرات.

٣٦ - ويستفاد من سندجرين (١٩٨٦) أنه يمكن تعيين موضوع أو وحدة التحليل بأنها أي كيان ملموس أو تجريدي (موضوع مادي، كائن حي، منظمة، حدث، إلخ). يود المستخدمون أن يحصلوا على معلومات عنه. وتعريف الموضوع في النهاية متصل بالمسألة الاجتماعية (أهداف الاستقصاء) التي تُجمع بيانات عنها وتُحلل. وكذلك لاستقصاءات الأسر المعيشية، أو مواضيعها، أو بنودها، أو عناصرها، أو وحداتها التي يود أصحاب المصلحة أن يحصلوا على معلومات عنها هي الأسرة المعيشية، مثلاً، والشخص، وقطع الأرض، إلخ. الموضوع الأساسي في معظم الحالات هو الأسرة المعيشية، ويوجد في العادة عدد من المواضيع المرتبطة بالموضوع الأساسي، وهذا يتوقف على الاستقصاء المعين.

٣٧ - يعرض الجدول ٩-١ المواضيع/الوحدات المعرفة في الاستقصاء الديمغرافي المجرى بين تعدادين للسكان في زمبابوي سنة ١٩٨٧. كانت الأسرة المعيشية هي الموضوع الأساسي والمواضيع المرتبطة بها هي "الشخص"، و"المرأة البالغة من العمر ١٢ سنة أو أكثر"، و"المتوفاة" (لاغرلوف، ١٩٨٨).

٣٨ - لكل موضوع توجد بضعة متغيرات هي موضع اهتمام. والمتغيرات هي صفات (توصيفات أو خصائص) للمواضيع. فالموضوع "الشخص"، مثلاً، يمكن أن تكون له متغيرات هي السن، والدخل، والمهنة، والوضع العائلي، إلخ. والمتغيرات ربما تكون نوعية أو كمية.

الجدول ٩ - ١

مثال على مواضيع/وحدات تحليل في استقصاء أُخِذَ من الاستقصاء الديمغرافي الذي أُجْرِيَ بين عمليتي تعداد للسكان في زمبابوي سنة ١٩٨٧

المواضيع ذات الصلة		تعريف الموضوع /			
المفتاح الخارجي	الموضوع	المتغيرات الهامة	الوحدة	المتغيرات المعرّفة	الموضوع/الوحدة
HID	الشخص	SOH = حجم الأسرة	البيت مجموعة من الأشخاص يعيشون ويأكلون معاً في العادة، ويُستثنى منه الزوار	HID = تعريف الأسرة المعيشية (المنطقة، القسم، القسم الفرعي)	الأسرة المعيشية
HID	المتوفى	المتغيرات الهامة		Eanr = رقم التعداد Ehnr = رب الأسرة المعيشية	
HID	الأسرة المعيشية	الجنس	الشخص فرد عادي في الأسرة المعيشية أو زائر في الليلة الماضية	HID	الشخص
HID, PID	المرأة البالغة من السن ١٢ سنة فما فوق	السن MARSTAT = الحالة الزوجية ETHNIC = مجموعة إثنية USMEM = فرد عادي في الأسرة المعيشية RELTH = العلاقة برب الأسرة		PID PID = تعريف الشخص	
HID	الأسرة المعيشية	SEXD = جنس المتوفى AGED = سن المتوفى	المتوفى الذي كان فرداً عادياً في الأسرة خلال الـ ١٢ شهراً الماضية	HID DID = تعريف المتوفى	المتوفون
HID, DID	الشخص	عدد الأطفال المولودين	كل امرأة بالغة من العمر ١٢ سنة أو أكثر وهي فرد عادي في الأسرة أو زائرة في الليلة الماضية	HID, DID	المرأة البالغة ١٢ سنة فما فوق

٣٩ - ينبغي أن يكون لكل موضوع تعريف فريد أيضاً. ويشير تعريف الموضوع المرتبط إلى الموضوع الأساسي الذي هو مرتبط به. فـ "الشخص"، مثلاً، يكون مرتبطاً بـ "أسرة"، ويُعرّف بتشكيلة من تعريف الأسرة (HID) وتعريف الشخص (PID)، أي الرقم المتسلسل داخل قائمة الأسر المعيشية (PID).

مدخلات نظام إعداد البيانات

٤٠ - تتألف المدخلات من القيم التي يحصل عليها المقابلون ويقيدونها وفقاً لاستبيان الاستقصاء.

## مخرجات نظام إعداد البيانات

٤١ - تتألف مخرجات النظام بصورة رئيسية من الجداول الإحصائية (المستندة إلى خطة الجدولة)، وقواعد بيانات تحتوي على بيانات جزئية وكلية، إلخ. وتتفاوت هذه بحسب نوع الموضوع، ونوع المتغير ونوع المقياس الإحصائي. والمتغيرات المُجدولة في العادة "أصلية"، لكنها يمكن أيضاً أن تكون مستمدة من متغيرات أصلية.

## تنظيم الملف

٤٢ - ينبغي في العادة أن يكون لدى المرء هياكل ملفات مختلفة في مرحلة المدخلات والمرحلة السابقة للجدولة. فملف طول المتغير، مثلاً (مقابل الملف المستوي) ربما يكون مفضلاً لقيد البيانات لاستقصاءات الأسر المعيشية، لأن الأسر المعيشية تختلف حجماً وتشكياً، ومن هنا جاءت الحاجة إلى سجلات متغيرة الطول أثناء قيد البيانات. تستخدم هذه الطريقة الفراغ بكفاءة لكنها غير مناسبة لعمليات الإعداد اللاحقة. غير أنه في نهاية المطاف توجد في الغالب إشارة إلى بيانات يُراد تنظيمها في ملفات مستوية بغية تيسير الجدولة والاستخدام الأمثل لأنواع مختلفة من البرامج الحاسوبية المعممة.

٤٣ - خريطة سير عمليات النظام: ينبغي وضع خريطة لسير عمليات استقصاء الأسر المعيشية تكون مفصلة إلى حد معقول. الخريطة هامة لأسباب كثيرة. منها، أولاً، أنها أداة لوضع جداول زمنية وتقدير الموارد اللازمة لإكمال عملية الاستقصاء. في الأحوال العادية، تشمل الأنشطة الرئيسية لإعداد البيانات لأي استقصاء ما يلي:

- (أ) فحص البيانات وتحريرها وترميزها؛
- (ب) قيد البيانات والتحقق منها وتصحيحها؛
- (ج) تحويل هيكل البيانات المستخدم في مرحلة المدخلات إلى هيكل بيانات مناسب لوضع الجداول؛
- (د) وضع الجداول.

٤٤ - ينبغي أن تشمل خرائط سير عمليات النظم أيضاً عمليات الملف الأساسي كالاختيار، والإسقاط، والبحث، ومقابلة الملفات، واستنباط متغيرات جديدة، والتجميع، والجدولة، والعرض في شكل بياني.

## نظام التوثيق

٤٥ - التوثيق الشامل الواضح (بمعنى مجموعة من البيانات الكلية) هامٌ لتطوير وصيانة نظم إعداد البيانات لاستقصاءات الأسر المعيشية. لذلك، من الأهمية بمكان توثيق الملفات والعمليات المختلفة، لكي يتسنى للأشخاص غير العاملين في تنفيذ النظام الأساسي أن يستخدموها هم أيضاً. لضمان كون الوثائق كافية، ينبغي استخدام قالب موحد وتخزينه إلكترونياً مع بياناته.

٤٦ - ينبغي، ما أمكن ذلك، أن تُستخدم نفس الأسماء، ونفس الرموز، ونفس أشكال البيانات للمتغيرات في نظم إعداد البيانات، في جميع الاستقصاءات المختلفة التي تجريها منظمة الاستقصاء، إذا كانت الرموز تحمل نفس المعاني. هذا على جانب كبير من الأهمية للمتغيرات التي تُستخدم في تعيين السجلات (المواضيع) في الملف، لأنه يمكن أن تستخدم هذه المتغيرات أيضاً عند تجميع (ضم) البيانات المأخوذة من النظم المختلفة.

٤٧ - لتصميم شاشات قيد البيانات أو المسح الضوئي للأشكال، تُبنى أدوات القالب كجزء من البرنامج الحاسوبي للمساعدة في التوثيق؛ ويلزم استخدامها استخداماً تاماً لتحقيق توثيق فعال. فأشكال (نظام إعداد عمليات تعداد السكان والاستقصاء)، مثلاً، أو قاموس البيانات للنظام المتكامل لإعداد البيانات على الحاسوب الصغير، تعرّف

موقع كل متغير في ملف البيانات — نقطة البداية ونقطة النهاية، سواء أكان المتغير رقماً أم حرفاً، وسواء أكان متكرراً أم لا، وإن كان متكرراً، كم مرة يتكرّر. ويضع القاموس أيضاً رقايع تعريف على القيم الموجودة داخل المتغير (سانت كاترين، ٢٠٠٣) (انظر التذييل الملحق بهذا الفصل للاطلاع على مزيد من المعلومات عن نظام إعداد عمليات تعداد السكان والاستقصاءات، والنظام المتكامل لإعداد البيانات على الحاسوب الصغير).

## ٩ - ٤ عمليات الاستقصاء وإعداد البيانات

### ٩ - ٤ - ١ إنشاء الأطر وتصميم العيّنات

٤٨ - كما نوقش في الفصلين الثالث والرابع، وحدات المعاينة في المرحلة الأولى لكثير من استقصاءات الأسر المعيشية هي مناطق تعداد السكان المعروفة في أحدث عمليات التعداد الوطني المتاحة. وإن إنشاء ملف حاسوبي يضم قائمة بجميع مناطق التعداد في البلد طريقة ملائمة وذات كفاءة لوضع إطار العيّنات في المرحلة الأولى، وأفضل طريقة لتحقيقه هي باستخدام برنامج يمكن من استخدام الجدول الإلكتروني (spreadsheet)، كبرنامج مايكروسوفت إكسل، مع وضع صف واحد لكل منطقة تعداد، ووضع أعمدة لجميع المعلومات التي يمكن أن تلزم (انظر التذييل الملحق بهذا الفصل للاطلاع على مزيد من المعلومات عن إكسل).

٤٩ - يجب أن يكون من السهل الوصول إلى الإطار واستخدامه في أعمال مختلفة، كالفرز، والترشيح، وإنتاج الإحصاءات الموجزة التي يمكن أن تساعد على تصميم العيّنات وتقديرها. مايكروسوفت إكسل برنامج سهل الاستخدام، وكثير من الناس يعرفون كيف يستخدمونه؛ وفيه وظائف الفرز والترشيح والتجميع، التي تلزم عند إعداد العيّنات من الإطار. ويمكن جلب الجداول الإلكترونية إلى معظم البرامج الحاسوبية الأخرى. ومن الأكثر ملاءمة بوجه عام إنشاء جداول إلكترونية مختلفة لكل طبقة من طبقات العيّنات.

٥٠ - ينبغي أن تكون محتويات سجلات وحدات الإطار على النحو التالي:

- يجب إدراج مُعرّف أولي ويجب أن يكون هذا المعرف رقمياً. وينبغي أن يكون له رمز يعرّف جميع الأقسام والأقسام الفرعية الإدارية التي توجد فيها وحدة الإطار تعريفاً فريداً. وستكون ثمة ميزة إذا كانت وحدات الإطار مرقّمة بترتيب جغرافي. وتكون لرموز مناطق التعداد في العادة هذه الخصائص مسبقاً
- ثمة معرف ثانوي وهو أيضاً هام، ذلك هو اسم القرية التي توجد فيها (أو التقسيم الفرعي الإداري الآخر الذي توجد فيه) وحدة الإطار. تُستخدم المعرفات الثانوية لتحديد موقع وحدة الإطار على الخرائط في الميدان
- ينبغي أن يشمل الملف أيضاً عدداً من خصائص وحدات المعاينة، كـمقاس الحجم (عدد السكان، الأسر المعيشية، كون الوحدة حضرية أو ريفية، كثافة السكان، إلخ). يمكن أن تُستخدم هذه الخصائص لتقسيم العينة إلى طبقات، أو إعطاء الوحدات مدى احتمال اختيارها، وكمتغيرات مساعدة في التقدير
- بيانات عملية كالمعلومات عن التغيرات في الوحدات، وينبغي شمول استخدام العيّنات

- ٥١ - ينبغي توثيق إجراءات الاختيار واحتمالات الاختيار في جميع وحدات العيّنات توثيقاً تاماً في كل مرحلة. وعندما تُستخدم عيّنات رئيسية يجب أن تكون ثمة سجلات تبين أيّ وحدات العيّنات الرئيسية هي التي استُخدمت في العيّنات المختارة لاستقصاءات معيّنة. ويجب استخدام نظام رقم تعريف موحد في وحدات المعاينة.
- ٥٢ - يمكن استخدام العيّنة الرئيسية في ناميبيا، بناءً على تعداد السكان والمساكن لسنة ١٩٩١، مثلاً لما نوقش أعلاه (مكتب الإحصاء المركزي في ناميبيا، ١٩٩٦).

### مثال

للتمكن من اختيار عيّنة عشوائية لمناطق جغرافية في ناميبيا، كان من الضروري وضع إطار عيّنات للمناطق الجغرافية. ولهذا الغرض، وُضع إطار للمناطق الجغرافية - وحدات المعاينة الأولية. كانت المنطقة منها تحتوي في المتوسط على ١٠٠ أسرة معيشية، وكانت معظم المناطق تضم ما يتراوح بين ٨٠ و ١٠٠ أسرة معيشية. وضعت قائمة المناطق من قائمة مناطق تعداد السكان والمساكن لسنة ١٩٩١. وقد ضُمَّت مناطق التعداد الصغيرة إلى مناطق مجاورة لتشكّل وحدات معاينة أولية ذات حجم كافٍ. كانت القاعدة التقديرية أن تضم وحدة المعاينة الأولية ما لا يقل عن ٨٠ أسرة معيشية. وكان ثمة ما مجموعه نحو ٦٨٥ وحدة معاينة أولية، مصنفة في طبقات من وحدات المعاينة الأولية الإقليمية، وريفية، وحضرية صغيرة، وحضرية.

٥٣ - كان التصنيف في طبقات بناءً على تصنيف مناطق التعداد أُجري أثناء الأعمال التحضيرية لتعداد سنة ١٩٩١. وأنشئ ما مجموعه ٣٢ طبقة، ووضعت فيها قوائم بوحدات المعاينة الأولية بترتيب جغرافي. في المناطق الحضرية والحضرية الصغيرة أُدرجت وحدات المعاينة الأولية في القوائم بحسب مستوى دخل المناطق. وكان أعلى ما في القوائم المناطق ذات الدخل المرتفع، تليها المناطق ذات الدخل المتوسط/المنخفض.

٥٤ - أعدّ مكتب الإحصاءات المركزي ملفات ببرنامج مايكروسوفت إكسل لإطار العيّنات الرئيسية لوحدات المعاينة الأولية. وضم الإطار ما يلي:

- الإقليم
- رقماً فريداً لوحدة المعاينة الأولية
- مستوى الدخل (للمناطق الحضرية فقط)
- المقاطعة
- رقم (أرقام) منطقة التعداد
- عدد الأسر المعيشية بناءً على تعداد سنة ١٩٩١
- مجموع عدد الأسر المعيشية في الطبقة الواحدة
- عدد السكان بحسب نوع الجنس، كما جاء في التعداد
- الحالة بالنسبة إلى العيّنة الرئيسية (ما إذا كانت وحدة المعاينة الأولية في العيّنة الرئيسية)
- رقم وحدة المعاينة في العيّنة الرئيسية (لوحدة المعاينة الأولية التي هي في العيّنة الرئيسية فقط)

- مُعامِلات الترجيح (عوامل رفع أو عوامل تضخم) لوحَدات المعاينة الأولى التي هي في العيِّنة الرئيسية فقط

٥٥ - كان هناك ملف مايكروسوفت إكسل واحد لكل إقليم، وصُنِّفت وحدات المعاينة الأولى في كل ملف مايكروسوفت إكسل في مجموعات بحسب كونها ريفية؛ أو حضرية صغيرة، أو حضرية — ذات دخل مرتفع؛ وحضرة صغيرة، وحضرية ذات دخل متوسط/منخفض.

٥٦ - بترسون (٢٠٠٣) يناقش مسائل مفصلة تتصل بالعيّنات الرئيسية. مونوز (٢٠٠٣) يبحث زيادة على ذلك كيف يمكن لإطار حاسوبي كالذي ورد وصفه أعلاه أن يثبت أهميته في تنفيذ إجراء المعاينة لاستقصاء الأسر المعيشية، بإرشاد الإجراء في جميع مراحلها: تنظيم إطار المرحلة الأولى، الذي يُبنى في العادة على أساس نتائج آخر تعداد للسكان والمسكن (مناطق التعداد)؛ اختيار وحدات المعاينة الأولى التي تتناسب احتماليتها مع حجمها (يقاس الحجم بعدد الأسر المعيشية أو الوحدات السكنية أو السكان)؛ وتحديث الجدول الإلكتروني على أساس وضع قوائم بأسر معيشية مختارة، وحساب احتمالات اختيارها، ومُعامِلات ترجيح المعاينة المقابلة. ويعطي الجزء ٩ - ٤ - ٣ - ٥ الوارد أدناه، والمتعلق بإجراءات تقدير النقاط وحساب مُعامِلات الترجيح، مناقشة مفصلة لحساب احتمالات الاختيار ومُعامِلات الترجيح المقابلة (انظر الفصل الخامس أيضاً). وسيتم الحصول على البيانات اللازمة لهذه الحسابات من جدول إلكتروني ورد وصفه أعلاه، ويمكن إجراء حساب مُعامِلات الترجيح باستخدام الجدول الإلكتروني الذي وضعه مونوز.

## ٩ - ٤ - ٢ جمع البيانات وإدارة البيانات

٥٧ - يمكن لاستقصاءات الأسر المعيشية أن تنتج كميات كبيرة من الاستبيانات المجاب عليها. وينبغي التفكير ملياً في إجراءات التناول المادي والإجراءات الحاسوبية لهذه الكميات الهائلة من الوثائق وضبطها في مرحلة مبكرة، إذا ما أُريدَ تجنب وقوع فوضى. وينبغي تخطيط الأعمال الروتينية للتناول اليدوي للاستبيانات (وضعها في الملفات واسترجاعها) بعناية، وينبغي أن تكون جاهزة للعمل قبل أن يبدأ وصول البيانات من الميدان بوقت طويل. وثمة جزء هام من هذا النظام يتطلب تقدير حجم البيانات المتوقعة لكي يتسنى الحصول على الملفات والصاديق، إلخ، وتخصيص أماكن لها على الرفوف أو في خزائن. والجزء الثاني من النظام عبارة عن سجلّ يمكن قيد المعلومات المتعلقة بالاستبيانات فيه فور وصولها، ويمكن متابعة تدفق البيانات عبر النظام فيها. هذه هي الجوانب الرئيسية لإدارة البيانات والشروط المسبقة الهامة لإدارة أي استراتيجية لإعداد البيانات للاستقصاء وتنفيذها بنجاح.

٥٨ - مسألة الأمن المادي للاستبيانات المجاب عليها مسألة هامة أيضاً في هذه المرحلة. ويرى هذا أنه مجال وُجد فيه أن المسح الضوئي للصور جذاب. وإذا مُسحت الاستبيانات بمجرد وصولها إلى المكتب تقلّ خطورة فقدانها نتيجة لحدوث شيء خطأ ممكن للبيانات التي في الاستبيانات. والمسح الضوئي للصور يوفّر مستوى إضافياً من مستويات الأمن، بالسماح بمسح الاستبيانات مسحاً ضوئياً في عين الموقع وبعد خروجها من الموقع لحفظها بعد مسحها (إدوين، ٢٠٠٣). غير أنه ينبغي ملاحظة أن النجاح في اعتماد هذه التكنولوجيا يعتمد إلى حد كبير على كيفية استخدامها من قبل المؤسسة المعنية وتنظيمها وإدارتها للعمليات المتصلة بها. ومن ثمّ، يمكن القول إنه بينما استُخدم المسح الضوئي بنجاح في بعض البلدان، لم ينجح في بلدان أخرى. ولكي ينجح العمل بالمسح الضوئي ينبغي إيلاء الاعتبار، في جملة عوامل أخرى، لطريقة تنظيم المكتب الإحصائي، من حيث كون إجراءاته مركزية أو لامركزية؛ والصورة الإجمالية للذين يُجرى لهم الاستقصاء، والضمانات المتعلقة بنوعية أدوات جمع البيانات للاستقصاء.

## ٩ - ٤ - ٣ إعداد البيانات

٥٩ - تحتاج البيانات التي تُجمع إلى وضعها في ملف بيانات. ويُسمى نقل البيانات من الاستبيانات إلى بيانات يقرأها الحاسوب قيد البيانات. وغالباً ما يكون من الضروري في هذا الصدد تقسيم القيم المتغيرة، التي تعطى كإجابات مفتوحة، إلى فئات؛ ويُشار إلى عملية التقسيم إلى فئات هذه بعبارة الترميز. ويمكن للمرء، بتحرير البيانات التي يُحصَل عليها، أن يعين البيانات الخاطئة. ويمكن بعد ذلك اتخاذ إجراءات مناسبة لفحص ما يُشتبه بأنها أخطاء، مثلاً، تجديد الاتصال بمصدر المعلومات. ويمكن أن يتبع هذه الفحوص تحديث (تصحيح). وتشمل خطوات إعداد البيانات: قيد البيانات، وترميزها، وتحريرها، وفحصها، وتحديثها/تصحيحها. ويشار إلى هذه الخطوات مجتمعة هنا بعبارة خطوة إعداد البيانات من عملية إعداد الاستقصاء.

## ٩ - ٤ - ٣ - ١ استراتيجيات إعداد البيانات

٦٠ - يتناول مونوز (٢٠٠٣) الجوانب المختلفة لإعداد البيانات وتشكيلاتها بالتفصيل. ويستتبع أكثر الترتيبات التنظيمية شيوعاً لاستقصاءات الأسر المعيشية القيام بإعداد البيانات في مواقع مركزية، بعد جمع البيانات في الميدان. وينطوي ترتيبٌ بديلٌ على دمج قيد البيانات مع العمليات الميدانية. وأحدث الابتكارات في هذا المضمار هو تقنية إجراء المقابلة بمساعدة الحاسوب.

## الإعداد المركزي للبيانات

٦١ - هذا هو الخيار الوحيد الذي كان موجوداً قبل مجيء الحاسوب الشخصي. وما زال في معظم الأحيان هو النهج الرئيسي المستخدم للاستقصاءات في البلدان النامية، مع شيء من التعديل نتيجة لإدخال الحواسيب الصغيرة. يُجرى قيد البيانات، بموجب هذا النهج، كعملية صناعية يُضطَلَعُ بها في عدد من المواقع بعد المقابلات. ويمكن إجراؤها في المقار الرئيسية لمكاتب الإحصاءات الوطنية أو في مكاتبها الإقليمية.

## إعداد البيانات في الميدان

٦٢ - نُظِرَ في الآونة الأخيرة إلى دمج أعمال مراقبة الجودة بمساعدة الحاسوب مع العمليات الميدانية على أنه واحد من العوامل الرئيسية لتحسين جودة استقصاءات الأسر المعيشية وتحسين توقيتها. وبموجب هذه الاستراتيجية، يُضطَلَعُ بأعمال مراقبة قيد البيانات واتساقها كجزء لا يتجزأ من العمليات الميدانية.

٦٣ - يستتبع واحد من أشكال هذه العملية جعل كاتب قيد البيانات يعمل باستخدام حاسوب مكتب في موقع ثابت (في المكتب الإقليمي لمكتب الإحصاءات الوطني، مثلاً) وتنظيم العمل الميداني بحيث يزور بقية أعضاء الفريق كل موقع من مواقع الاستقصاء (وحدة المعاينة الأولية بوجه عام) مرتين على الأقل، لإعطاء الكاتب وقتاً كافياً لقيد البيانات والتحقق من صحتها فيما بين الزيارات. في الزيارة الثانية والزيارات اللاحقة يطرح المقابِلون على الأسر المعيشية ذات الصلة مرة ثانية الأسئلة التي اكتشف مكتب قيد البيانات فيها أخطاءً أو إغفال معلومات أو عدم اتساق في المعلومات.

٦٤ - يوجد نهج آخر، وهو جعل كاتب قيد البيانات يعمل على حاسوب صغير محمول وينضم إلى الفريق في زيارته لمواقع الاستقصاء. ويبقى الفريق كله في الموقع حتى يتم قيد جميع البيانات ويصفها برنامج قيد البيانات بأنها كاملة وصحيحة.

٦٥ - تشمل الميزات النسبية المدركة لدمج جمع البيانات وإعداد البيانات معاً مجال الحصول على بيانات أعلى جودة لأنه يمكن تصحيح الأخطاء والمقابلون ما زالوا في الميدان، وإمكانية توليد قواعد بيانات، ووضع الجداول وإجراء التحاليل بعد انتهاء العمليات الميدانية بوقت قصير جداً، والمجال الأوسع لتوحيد جمع البيانات من قبل المقابلين.

٦٦ - في النهجين الوارد وصفهما أعلاه تكون الحاجة ماسة إلى توفّر الطاقة الكهربائية باستمرار في موقع العمليات. وفي البلدان التي تفتقر إلى الإمداد الجيد بالطاقة الكهربائية، كمعظم البلدان النامية، لا سيما في مناطقها الريفية، لا يكون من المجدي، ببساطة، استخدام هذه الخيارات. ومن الجدير بالذكر أنه توجد تحديات تنظيمية وسوقية مرتبطة باستخدام المعدات المتحركة لجمع البيانات وإعدادها. ومن بين الشروط المسبقة للنجاح في استخدام هذه الاستراتيجية وجود نظام إداري فعّال؛ وأمن فعّال للمعدات والبيانات؛ وتوفّر أماكن تخزين كافية لحفظ نُسخ من البيانات؛ وتوفر إمدادات كافية من المواد المستهلكة كالبطاريات الاحتياطية لاستخدامها في الميدان.

#### إجراء المقابلات بمساعدة الحاسوب

٦٧ - إجراء المقابلات الشخصية بمساعدة الحاسوب شكّل من أشكال المقابلات الشخصية، يُحضر المقابل معه حاسوباً محمولاً أو يدوياً لكي يقيّد فيه البيانات مباشرة في قاعدة البيانات، بدلاً من ملء استبيان مطبوع على الورق. هذه الطريقة توفّر وقتاً في إعداد البيانات، وتُريح المقابل من عبء حمل مئات الاستبيانات في جولاته. غير أنه مع كون التكنولوجيا متوفرة من سنين عديدة، لم يُفعل سوى شيء ضئيل جداً بصورة جدية لتطبيق هذه الاستراتيجية في الاستقصاءات المعقدة في البلدان النامية. يمكن أن يكون اتخاذ ترتيبات للعمل بهذه الطريقة لجمع المعلومات باهظ التكلفة، ويتطلب أن تتوفر لدى المقابلين مهارات استخدام الحاسوب والطباعة. ويتطلب إجراء المقابلات بمساعدة الحاسوب أيضاً مقابلات مرتبة ترتيباً جيداً لها بداية ولها نهاية. غير أن معظم الاستقصاءات في البلدان النامية تتطلب زيارات متعددة لكل أسرة معيشية، وإجراء مقابلات منفصلة مع أفراد الأسرة، إلخ. في عملية ليست مرتبة ترتيباً دقيقاً وإنما تعتمد على المهارة الشخصية للمقابل نفسه.

#### ٩ - ٤ - ٣ - ٢ ترميز وتحضير بيانات الاستقصاء

٦٨ - ربما يمثل فحص البيانات وتحريها وترميزها أصعب مرحلة من مراحل إعداد البيانات. ففي تنظيم إدارة البيانات وإعدادها غالباً ما يواجه فنيو الاستقصاءات المدربون حديثاً على أكبر الصعوبات. وإن أفضل طريقة لإجراء عمليات فحص البيانات وتحريها وترميزها هي إجراء ذلك بطريقة آلية، إن أمكن ذلك. غير أنه في حالة الترميز يجب بحكم الطبيعة أن يُنظر في الحالات التي يتعذر فيها تعيين الرموز آلياً وفي هذه الحالة يجب بحكم الضرورة تعيين الرموز يدوياً.

#### الترميز

٦٩ - يتمثل الهدف في إعداد البيانات بشكل مناسب لقيدها في الحاسوب. وتنطوي عملية الترميز بصورة رئيسية على تعيين رموز رقمية لإجابات مسجلة بالكلمات (كما في حالة المواقع الجغرافية، مثلاً، والمهنة، والصناعة، إلخ). وربما تقتضي أيضاً نسخ كلمات مسجلة صوتياً تحوّل بموجبه الرموز الرقمية التي تكون قد عُيّنَت وسُجّلت من قبل أثناء المقابلة إلى صحائف ترميز.

٧٠ - ينبغي إعداد كتاب يدوي لإعطاء إرشاد واضح للعاملين في الترميز. وينبغي أن يحتوي هذا الدليل على مجموعة من الفئات غير المترابطة، تغطي جميع الإجابات المقبولة على الأسئلة التي هي قيد النظر. ومن المستصوب في الاستقصاءات الكبيرة للأسر المعيشية توسيع نطاق الترميز المسبق للأسئلة وإقفالها.

#### تحرير البيانات وفحصها

٧١ - القصد من فحص الاستبيانات و/أو تحريرها هو (أ) لتحقيق الاتساق داخل البيانات والاتساق داخل الجداول وفيما بينها، (ب) للكشف عن الأخطاء الخارجية والتحقق منها وتصحيحها أو إزالتها، لأن القيم المتطرفة تساهم مساهمة كبيرة في قابلية العيّنات للتغير في تقديرات الاستقصاء.

٧٢ - ينطوي التحرير على مراجعة القيود في الاستبيان أو تصحيحها. وربما يُنظر إليه على أنه إجراء تصحيح، تُكتشف به الاختلافات والمستحيلات في البيانات وتُصحّح، أو إجراء إحصائي تُجرى فيه الفحوص بناءً على تحليل إحصائي للبيانات. ويتجه الميل الآن إلى قيام الحاسوب بدور متزايد في التحرير، إما عند قيد البيانات وإما في جولات تحرير خاصة للبيانات. وإن جولات التحرير الخاصة هذه يمكن أن تتفاعل بعضها مع بعض وربما لا تتفاعل. ومعنى ذلك أن الكاتب ربما يصحح الأخطاء فوراً وربما لا يصححها غير أن تصويب الأخطاء الأكثر تعقيداً يحتاج إلى وقت أطول وتحليل أعمق قبل أن يمكن إيجاد التصويب الصحيح، ولهذا تكون جولات التحرير غير المتفاعل أنسب. وتورد المواد المرجعية التي كتبها أولسون (١٩٩٠) مناقشة مفصلة لمختلف جوانب فحص وتحرير بيانات الاستقصاء.

٧٣ - الفحص والتحرير اليدوي: المهمة الرئيسية للفحص أو التحرير اليدوي هي الكشف عن كلمات أُغفلت أو عدم اتساق وغير ذلك من الأخطاء البديهية في الاستبيانات قبل مراحل الإعداد اللاحقة. وينبغي أن يبدأ التحرير اليدوي بسرعة وفي أقرب مكان ممكن إلى مصدر البيانات، أي في المكتب الإقليمي مثلاً، أو في مكتب المقاطعة أو مستوى أدنى منهما. والشئ المثالي هو أن تُكتشف أغلبية الأخطاء وتُصحّح في الميدان قبل إرسال الاستمارات إلى مركز الإعداد. لذلك، يأمر التدريب ودليل التعليمات عادةً المقابل والمشرّف بفحص الاستبيانات وتصحيح أيّ أخطاء توجد فيها في الميدان قبل إرسال البيانات منه. هذه مهمة هامة وصعبة، يصبح أداؤها دالةً على جودة العمل الميداني، وفعالية الإشراف، وإدارة الاستقصاء، إلخ.

٧٤ - التحرير بمساعدة الحاسوب: يمكن التحرير بواسطة الحاسوب: (أ) بصورة غير تفاعلية في مرحلة قيد البيانات، أو (ب) باستخدام إعداد البيانات على دفعات بعد قيدها، أو (ج) باستخدام توليفة من البندين (أ) و(ب). التحرير التفاعلي يميل إلى أن يكون أكثر فائدة في حالة الأخطاء البسيطة (كالأخطاء المطبعية، مثلاً): ومن شأنه أن يؤخر عملية التقاط البيانات في حالة الأخطاء التي تستدعي التشاور مع المشرفين. وينبغي أن تُترك معالجة هذه الأخطاء، بما في ذلك الأخطاء الناتجة عن عدم الاستجابة، لعملية تحرير منفصلة بواسطة الحاسوب.

٧٥ - تُصمّمُ برامج التحرير بمساعدة الحاسوب في الغالب باستخدام برامج قواعد بيانات مثل النظام المتكامل للإعداد بواسطة الحاسوب (IMPS)، والنظام المتكامل لتحليل الاستقصاء (ISSA)، ونظام إعداد تعداد السكان والاستقصاء (CSPro)، وبرنامج فيجيوال بايسك (Visual Basic)، ومايكروسوفت أكسس (Microsoft Access) (يوجد مزيد من التفاصيل عن البرامج المذكورة أعلاه في التذييل الملحق بهذا الفصل). أبسط البرامج تسمح للبيانات مساحاً ضوئياً، سَجِلاً فَسَجِلاً، وتلاحظ الاختلافات استناداً إلى قواعد تحرير مكتوبة في البرنامج. وفي برامج التحرير الأكثر تقدماً، يمكن مقارنة المتغيرات (تعيين المتغيرات، مثلاً) بين ملف وآخر ويمكن ملاحظة الاختلافات.

ويتألف ناتج النُظم من قوائم بالأخطاء، وهذه تُفحص في الغالب يدوياً، مقارنةً بالبيانات الخام. وتُصحّح الأخطاء في نسخة من ملف البيانات الخام.

#### أنواع الفحوصات

٧٦ - تحتاج البيانات المتعلقة بالاستبيانات إلى إخضاعها لأنواع مختلفة من المراجعات، وتشمل المراجعات النموذجية مراجعات النطاق، ومراجعات مقارنةً ببيانات مرجعية، ومراجعات التخطّي، ومراجعات الاتساق، والمراجعات الطباعية (مونوز، ٢٠٠٣).

٧٧ - مراجعات النطاق: يقصد بمراجعات النطاق أن تضمن أن كل متغير في الاستقصاء يحتوي فقط على بيانات تقع ضمن نطاق محدود من القيم الصحيحة. فالمتغيرات الفئوية لا يوجد فيها إلا واحدة من القيم المحددة لها مسبقاً في الاستبيان ( فنوع الجنس، مثلاً لا يمكن ترميزه إلا بـ "١" للذكر و"٢" للأنثى). ومتغيرات التسلسل الزمني يجب أن تحتوي على تواريخ صحيحة والمتغيرات الرقمية يجب أن تقع بين قيم الحد الأدنى وقيم الحد الأقصى (كأن تقع بين صفر و٩٥ سنة في حالة السن). وتحدث حالة خاصة في مراجعات النطاق عندما يمكن فحص البيانات الآتية من ميدانين أو أكثر توجد بينها صلة وثيقة مقارنةً بجداول مرجعية خارجية.

٧٨ - مراجعات التخطّي: هذه المراجعات تتحقّق مما إذا كانت أنماط التخطّي قد اتّبعَت على نحو سليم. فالمراجعة البسيطة، مثلاً، تتحقّق من أن الأسئلة التي لا توجّه للأطفال المدارس لا تُسجّل لطفل أجاب بالنفي على سؤال أوّليّ عما إذا كان ملتحقاً بالمدرسة. أما المراجعة الأكثر تعقيداً فهي تتحقّق من أن الوحدات الأولى الصحيحة للاستبيان قد مُلئت لكل مُجيب. ويفترض أن يجيب كل فرد في الأسرة المعيشية على (أو يتخطّى) أجزاء معينة من الاستبيان، ويتوقف ذلك على سن وجنس الشخص المجيب. فيمكن، مثلاً، إدراج النساء في الفئة العمرية ١٥ - ٤٩ في الجزء المتعلق بالخصوبة، أما الرجال فلا يمكن إدراجهم فيه.

٧٩ - مراجعات الاتساق: تتحقّق مراجعات الاتساق من توافق القيم الآتية من سؤال واحد مع القيم الآتية من سؤال آخر. وتكون المراجعة بسيطة عندما تأتي القيمتان من وحدة إحصائية واحدة مثل تاريخ ميلاد فرد معين وعمره. وتشمل مراجعات الاتساق الأكثر تعقيداً من وحدتين مختلفتين أو أكثر من وحدات المشاهدة. مثال ذلك أنه ينبغي أن يكون الوالدان أكبر من أطفالهما بـ ١٥ سنة على الأقل.

٨٠ - مراجعات الطباعة: من أخطاء الطباعة العادية عكس موقّعي رقمين (كأن يُكتب "١٤" بدلاً من "٤١") في القيود الرقمية. هذا الخطأ فيما يتعلق بالسن يمكن كشفه في مراجعات الاتساق في الحالات الزوجية أو العلاقات الأسرية. فإذا كان شخص متزوج أو أرمل يبلغ من العمر ٤١ سنة، وقُدِّد عمره خطأً ١٤ سنة، يظهر البرنامج إشارة لوجود الخطأ في مراجعة الحالة الزوجية. غير أن نفس الخطأ يمر بسهولة إذا وقع في قيد النفقات الشهرية على اللحم، ولا يلاحظه البرنامج لأن كلا المبلغين ١٤ دولاراً و٤١ دولاراً جائز. ومن الإجراءات العادية للكشف عن هذا الخطأ أن تقيد البيانات من كل استبيان مرتين من قبل كاتبين مختلفين.

#### معالجة البيانات المفقودة

٨١ - عندما يصل الاستقصاء إلى مرحلة الإعداد، من المؤكد أن يكون ثمة قدر كبير من البيانات المفقودة. فقد تكون بعض الأسر المعيشية قد انتقلت من مسكنها أو رفضت الإجابة على الاستبيان. أو ربما تكون بعض الأسئلة في الاستبيان قد تُركت دون جواب؛ أو ربما تكون بعض البيانات مزيفة أو غير متسقة مع معلومات أخرى في الاستبيان. والنتيجة، أيّاً كان السبب، هي عدم وجود قيد ما أو وجود خانة فارغة أو فارغة جزئياً في السجل.

٨٢ - من المهم التمييز بين البيانات المفقودة — بمعنى أن بيانات كان يجب أن تكون موجودة لكن قيمتها الصحيحة غير معروفة — والبيانات التي قيمتها صفر. فيمكن، مثلاً، أن يكون استبيان ما فارغاً لأن الأسرة رفضت المشاركة، بينما يمكن أن يكون جزء من استبيان آخر فارغاً لأن الأسرة لم تزرع أي محصول في حقولها، مثلاً. في الحالة الثانية، يجب أن يكون في خانة المتغير "المساحة المزروعة" صفر. ينبغي إبقاء هذه القيود في الملف من أجل التحليل والجدولة.

٨٣ - يتوقف النهج الذي يجب اتخاذه بشأن البيانات المفقودة فعلاً على نوع البيانات المفقودة. فيمكن أن يكون عنصر مختار من عناصر العينة مفقوداً كله بسبب رفض الأسرة المشاركة في الاستقصاء أو عدم قدرة الأسرة المستجيبة على الإجابة على مجموعة الأسئلة الواردة في الاستبيان بأسرها. وفي هذه الحالات قيل إنه حدث تدوين "عدم استجابة الوحدة".

٨٤ - إذا استطاع المستجيب الإجابة على بعض الأسئلة فقط دون غيرها، يُقال إنه حدث "عدم استجابة على بند / عدم استجابة جزئي"، لأنه حصل على بعض البيانات بشأن الأسرة المعيشية لا كلها.

٨٥ - تثير البيانات المفقودة من أي النوعين مسألة تقديرات متحيزة في الاستقصاء، كما تكرر تأكيد ذلك في هذا الدليل. للاطلاع على مناقشة مفصلة بشأن مُعاملة ملائمة لعدم الاستجابة، بما في ذلك طرق تعديلها، انظر الفصل السادس.

٨٦ - في حالة عدم الاستجابة الجزئي ربما يكون من الضروري، لتحقيق الاتساق في المجاميع، الاستعاضة عن القيم المفقودة بتقديرات معقولة. وهذا يسمى العزو [الاستعاضة عن القيمة المفقودة بقيمة مناظرة من استقصاء مشابه]. كما لوحظ في الفصل السادس. توجد نُهجٌ مختلفة يمكن استخدامها لعزو القيم البديلة. ومن هذه النهج ما يلي:

- عزو متوسط القيمة: استخدام متوسط القيمة (في وحدة المعاينة الأولية أو مجموعة البيانات كلها) للتعويض عن القيمة المفقودة؛
- العزو من "السطح الساخن": استعارة القيم المفقودة من سجل (مانح) شبيه بالسجل الناقص. وينبغي أن يكون السجل المانح قد اجتاز جميع فحوص التحرير؛
- العزو الإحصائي: استخدام علاقة (تراجع، نسبة) مع متغير آخر مستمد من بيانات كاملة للاستعاضة بها عن القيمة المفقودة.

٨٧ - ما ذكر أعلاه ما هو إلا بعض الطرق المتاحة للعزو: توجد عدة طرق أخرى لهذا الغرض. تتوقف كفاءة العزو، طبعاً، على مدى نجاح نموذج العزو (المستعاض به) في التكافؤ مع عدم الاستجابة. فيما يتعلق باختيار المعلومات المساعدة المتاحة، من المهم أن تكون ثمة علاقة بين المتغير والمتغير المراد الاستعاضة عنه (انظر أولسون (١٩٩٠) للاطلاع على مزيد من المعلومات عن هذه المسألة).

## ٩ - ٤ - ٣ - ٣ قيد البيانات

٨٨ - الهدف من قيد البيانات هو تحويل المعلومات المكتوبة على الورق في الاستبيان إلى مُنتج وسيط (ملفات قابلة للقراءة بالحاسوب)، ينبغي تهذيبه زيادة على ما هو عليه بواسطة برامج تحرير وعمليات كتابية بغية الحصول على ما يسمى بقواعد بيانات نظيفة كمنتج نهائي. في المرحلة الأولية من مراحل قيد البيانات تكون الأولوية للسرعة وضمان كون المعلومات التي في الملفات تتفق اتفاقاً تاماً مع المعلومات التي جُمعت في الاستبيانات.

٨٩ - يجب أن يُتخذ القرار بشأن الطريقة المستخدمة في قيد البيانات المنقولة من الاستبيانات إلى الحاسوب في مرحلة مبكرة، لأن ذلك يؤثرُ كثيراً في سير العمل الأساسي، وتقنية تخزين البيانات، وتصميم الشكل، وكذلك في تشكيل هيئة الموظفين العاملين في ذلك.

قيد البيانات من لوحة المفاتيح إلى قرص الحاسوب

٩٠ - يشتمل قيد البيانات من لوحة المفاتيح إلى قرص الحاسوب على طباعة بيانات مُشفّرة على قرص صلب، مثلاً، أو قرص صغير، أو قرص مدمج. وقد اكتسبت منظمات استقصاء كثيرة في البلدان النامية خبرة كبيرة في استخدام هذه الطريقة لقيد البيانات. وهذا هو النهج الرئيسي المستخدم وقد تعزز استخدامه بمجىء الحواسيب الشخصية والبرامج الحاسوبية ذات الصلة.

٩١ - تطبيق قيد البيانات: يتألف تطبيق قيد البيانات عادة من ثلاث وحدات. الوحدة الأولى هي عندما تقيّد كل المعلومات. والوحدة الثانية، للتحقق من البيانات المقيدة، هي عندما يُشهد بجودة البيانات المقيدة ومتابعة أداء الكتبة الذين قيدوا البيانات. والوحدة الثالثة هي لتصويب المعلومات المقيدة، إذ ربما تدعو الحاجة إلى تصحيح أخطاء في القيم لم تُكتشف أثناء قيد البيانات أو أثناء عملية التصحيح.

٩٢ - توجد لتطبيق قيد البيانات في العادة قائمة رئيسية، يختار منها الشخص المسؤول عن قيد البيانات الوظيفة التي يحتاج إليها: قيد البيانات أو التحقق منها أو تصويبها. ويجب على المستخدم، قبل أن يبدأ العمل بالقائمة، أن يشهد، بعد وضع اسم المستخدم وكلمة المرور، أنه مأذون له بدخول التطبيق. وإذا تعدّر عليه الدخول (أي إذا كان ثمة خطأ في اسم المستخدم أو كلمة المرور)، يُغلق التطبيق فوراً. جميع أسماء المستخدمين وكلمات المرور مخزونة في جدول في آخر الملف الذي تشقّر فيه كلمات المرور. وعندما يدخل المستخدم إلى النظام بكلمة مرور صحيحة تُحدّث الجداول التي في نهاية الملف.

٩٣ - وحدة قيد البيانات: وحدة قيد البيانات هي الصلة بين الاستبيان وملف البيانات أو قاعدة البيانات. يجب أن يكون نظام الإدخال هذا بسيطاً يسهل على كاتب قيد البيانات استخدامه. وتوجد بعض المتطلبات الهامة التي تشترط ما يلي:

- يجب أن تبدو شاشة قيد البيانات أشبه ما تكون بالصفحات المقابلة لها من الاستبيان. فبهذا يسهل على كاتب قيد البيانات من نظره إلى الاستبيان تحديد الخانة المقابلة على الشاشة.
- السرعة في قيد البيانات أمر بالغ الأهمية. ولا يريد الكاتب أن ينتظر حتى يُقيّم له نظام الحاسوب كل قيمة يقيدّها. لذلك يجب أن تكون عملية التقييم سريعة جداً؛ وهذا يعني ضمناً أن النظام لا يستطيع البقاء على اتصال بالجهاز الخادم أكثر من اللازم، ويعني ذلك في هذه الحالة أن القيم لن تُحفظ في قاعدة البيانات حتى يتم قيد جميع القيم المتعلقة بالأسرة المعيشية. العيب هنا هو أن البيانات المقيدة حالياً عن الأسرة المعيشية تُفقد إذا أُغلق التطبيق لسبب ما. غير أن فائدة السرعة العالية نسبياً أهم من هذا.
- ينبغي أن يكون لكل قيمة في الاستبيان رمز رقمي للتمكين من استخدام لوحة مفاتيح رقمية، وهي أساس السرعة العالية.
- ينبغي أن يكون لوحدة قيد البيانات جهاز متغير لمراقبة الصحة يتلقّى منه كاتب البيانات إشارة إلى خطأ حينما يقيد قيمة غير صحيحة. ويجب أن يُعنى جهاز مراقبة الصحة أيضاً بالقيم ذات الصلة؛ فإذا كانت لنوع الجنس القيمة "١" (للذكور)، يجب تعطيل المعلومات المتعلقة بالخصوبة.

• ينبغي طبعاً أن يطلق برنامج قيد البيانات إشارة خطأ إلى أي وضع يبدو أنه مستحيل منطقياً أو طبيعياً ( كأن تكون ابنة، مثلاً، أكبر سناً من أمها)، أو أنه مستبعد جداً (كأن تكون ابنة أصغر من أمها بفترة أقل من ١٥ سنة).

• من الأهمية بمكان متابعة عدد ضربات المفاتيح ووقت قيد البيانات لاستخدامهما إحصائياً فيما بعد، للتنبؤ بمجموع الوقت الذي استغرقه قيد البيانات.

٩٤ - وحدة التحقق من البيانات: الغرض من نظام التحقق هو تقديم معلومات عن جودة البيانات المقيدة ونسبة فشل كل كاتب من كتبة قيد البيانات. ويكون تصميم شاشة هذه الوحدة نفس تصميم شاشة وحدة قيد البيانات، دون أي فرق منظور بينهما. والفرق الرئيسي، بدلاً من ذلك، لا يقف عند حد أن عدد ضربات المفاتيح موجز، وإنما يوجد عدد الأخطاء أيضاً. وتشمل خيارات نوع التحقق، التحقق الكلي، حيث يتم تحقيق كل مناطق التعداد والاستبيانات في منطقة التعداد المحققة، أو التحقق بالعيّنة، حيث يتم التحقق من بعض مناطق التعداد وبعض الاستبيانات فقط.

٩٥ - وحدة تصويب البيانات: تُستخدم وحدة تصويب البيانات بصورة رئيسية لتصحيح البيانات التي لم يكن في الإمكان، لسبب من الأسباب، إكمالها في وحدة قيد البيانات. ويمكن في هذه الوحدة إضافة، أو حذف، أو تحديث معلومات تتراوح من معلومات أسرة معيشية كاملة إلى معلومات قيمة واحدة.

٩٦ - تطبيق إدارة المشرف: تطبيق الإدارة هو الأداة التي يحقق بها المشرفون التغييرات في قاعدة البيانات. وتُستخدم الأداة بصورة رئيسية لتصحيح ملف المجموعات الرئيسي، وتلقي التقارير عن أداء المستخدمين. ومن المهم أن:

- يسيطر المشرفون سيطرة كاملة على ملف الدفعات الرئيسي من التطبيق. ويتسنى لهم أن يضيفوا أو يحذفوا أو يحدّثوا المعلومات التي في ملف الدفعات الرئيسي
- يمكن أن يُضاف مستخدمون وأن تُحذف أسماءهم، وأن يُحصل على قائمة كاملة بكل المستخدمين. ويمكن مراجعة وضع جميع المستخدمين، أو مراجعة مستخدم واحد فقط
- يمكن مراجعة إحصاءات ضربات المفاتيح وطبعتها. ويمكن اختيار فترات زمنية مختلفة
- يمكن مراجعة نسبة فشل مستخدم واحد، ومتوسط نسب فشل كل المستخدمين وطباعتها
- يمكن إعادة ترتيب وضع منطقة تعداد لقيد البيانات أو للتحقق من البيانات
- يمكن الحصول من هذا التطبيق على جميع المعلومات التي يحتاج إليها المشرفون لإدارة عملهم

يتناول سفينسون (١٩٩٦) بالتفصيل مختلف جوانب نظم قيد البيانات من لوحة المفاتيح إلى قرص الحاسوب.

٩٧ - مناهج نظم قيد البيانات من لوحة المفاتيح إلى قرص الحاسوب: توجد مناهج عديدة لتطوير برامج قيد البيانات وتحريرها متاحة في الأسواق. فقد أثبت نظام إعداد تعداد السكان والاستقصاء، وسلفه النظام المتكامل للإعداد بالحاسوب الصغير، قدرتهما على دعم تطوير برامج فعّالة لقيد البيانات وتحريرها للاستقصاءات الوطنية المعقدة في كثير من البلدان النامية. وأثبت البرنامج أيضاً أنه يسهل الحصول عليهما واستخدامهما (مونوز، ٢٠٠٣).

#### المسح الضوئي

٩٨ - استخدام المسح الضوئي في إعداد بيانات تعداد السكان والاستقصاءات آخذ بالازدياد بسرعة. فقبل بضعة سنوات فقط كان قيد البيانات في البرامج الرئيسية مرادفاً للنظام العامل بلوحة المفاتيح. ولم تكن برامج منافسة

كثيرة متاحة في السوق. أما اليوم فقد تغير المشهد، وأصبحت أكثر نُظُم قيد البيانات مبيعاً كُلُّها تستند إلى تقنيات المسح الضوئي. توجد بضعة أقسام فرعية لهذه التقنيات، ولكل منها ميزاته ونقاط ضعفه. ومن بين أكثر التقنيات شيوعاً ما يلي: التعرف على الحروف ضوئياً، وهذا يعني التعرف على الحروف المطبوعة بالحاسوب؛ والتعرف الذكي على الحروف، وهذا يعني التعرف على الحروف المكتوبة بخط اليد؛ والتعرف الضوئي على العلامات، وهذا يعني التعرف على العلامات المرسومة بقلم الحبر أو قلم الرصاص في مواقع مقررّة سلفاً، ويكون ذلك عادة في مربعات العلامات؛ والتعرف على شفرة الخطوط المتوازية، وهذا يعني التعرف على البيانات المشفرة في شفرات الخطوط المتوازية المطبوعة.

٩٩ - يُستفاد ممّا كتبه لندل (٢٠٠٣)، فيما يتعلق باستخدام تكنولوجيا المسح الضوئي للاستقصاءات الإحصائية وعمليات تعداد السكان، أن الخيار هو بصورة رئيسية بين التعرف الذكي على الحروف والتعرف الضوئي على العلامات. فالبلد الذي فيه عدد كبير من السكان يفضل التعرف الضوئي على العلامات، بينما الاستبيان المعقد يفضل التعرف الذكي على الحروف. التعرف الضوئي على العلامات يقيد تصميم الاستمارة لكنه يوفر الإعداد السريع ويتطلب من الموظفين مهارة أقل نسبياً. أما التعرف الذكي على الحروف فيسمح بالحرية في تصميم الاستمارات لكن الإعداد به يتطلب سرعة أكبر في الحاسوب ومهارة أكثر من الموظفين. شفرات الخطوط المتوازية تُستخدم في العادة فقط لاستعادة معلومات التعريف، كأرقام الاستمارات، مثلاً، لأن شفرة الخطوط المتوازية تستخدم معلومات فحص الأرقام لتقليل الأخطاء إلى الحد الأدنى.

١٠٠ - أثناء عملية المسح الضوئي تُمسح الاستبيانات بسرعة تتراوح بين ٤٠ و ٩٠ صحيفة ذات صفحتين في الدقيقة. والسرعة هي أهم عامل في تقرير اختيار المسح الضوئي على الأشكال التقليدية لقيد البيانات التي تستخدم لوحة المفاتيح لطباعة البيانات. ثم يُستخدم برنامج المسح الضوئي لتعيين صفحات الاستبيان وتقييم مضمونها باستخدام التعرف الذكي على الحروف والتعرف الضوئي على العلامات. وتُرسل البنود التي هي موضع سؤال أو التي هي بحاجة إلى تشفير إلى المحقق، الذي يستعرض البنود المكتوبة بخط سيئ ويشفر الاستبيانات المفتوحة من الجداول الإلكترونية التي تتدلى على الشاشة، والتي تشكّل جزءاً أساسياً من القلب. ويوجد قدر كبير من المرونة فيما يتعلق بكيفية أداء مراجعات التحقق هذه، وذلك يتوقف على كيفية تنظيم قالب المسح الضوئي. ويمكن مراجعة المتغيرات الحاسمة كلياً أو جزئياً لرفع مستوى دقة الإجابة المسجلة في ملف البيانات.

١٠١ - أظهر أن استخدام عملية المسح الضوئي للصور يمكن أن تزيد كفاءة تسجيل البيانات بنسبة ٧٠ في المائة (إدوين، ٢٠٠٣). ويمكن إبطال كثير من المشاكل المرتبطة بالمسح الضوئي بواسطة التنظيم التقني الصحيح للعملية. ويمكن معالجة مشكلة الصفحات المفقودة أو غير المتوافقة مع غيرها باستخدام شفرات متوازية الخطوط مطبوعة مسبقاً في الاستبيان. وإذا وجدت صيانة صحيحة وإشراف صحيح، واستُخدمت باعتبارهما مَرَكَبَةً للوصل بين مختلف صفحات الاستبيان. وإذا كانت ثمة صيانة صحيحة وإشراف صحيح على المعدات والبرامج الحاسوبية فإن تكلفة عملية المسح الضوئي (بما في ذلك شراء المعدات والبرامج الحاسوبية) ربما يظهر أنها أقل كثيراً من تكلفة عملية طباعة البيانات باستخدام لوحة المفاتيح.

١٠٢ - كانت الخبرة في استعمال المسح الضوئي لاستقصاءات الأسر المعيشية محدودة جداً بوجه عام، لا سيما في الإقليم الواقع جنوبي الصحراء الكبرى. غير أن استخدامه في جولة سنة ٢٠٠٠ من عمليات تعداد السكان والمساكن كان كبيراً جداً، وربما يمثل نقطة تحوّل فيما يتعلق باعتماده بوجه عام. فقد استُخدم المسح الضوئي في كينيا، مثلاً، وجمهورية تنزانيا المتحدة، وجنوب أفريقيا، وناميبيا، وزامبيا في أحدث عمليات التعداد التي أجريت فيها.

وفي الآونة الأخيرة استُخدم أيضاً في جميع استبيانات المؤشرات الرئيسية للرفاه الاجتماعي، مدفوعاً بالبنك الدولي. واعتمده بلدان مثل ناميبيا وجنوب أفريقيا أيضاً في برامجها لاستقصاء الأسر المعيشية.

#### ٩ - ٣ - ٤ هيكَل الملفات وتنظيم مجموعات البيانات

##### تخزين البيانات

١٠٣ - في استقصاءات الأسر المعيشية، التي تحتوي في الأحوال النموذجية على معلومات على مستوى الأسر المعيشية وعلى مستوى الفرد كليهما، يمكن أن يستدعي استخدام حيز التخزين بكفاءة قالباً تبعياً أو متغير الطول من قوالب الملفات، لأن الأسر المعيشية المختلفة تحتوي على عدد مختلف من الأفراد الملحقين بالأسرة. وإن الملف المستوي، الذي يشغل نسبة كبيرة بلا لزوم من الحيز الفارغ، لا يكون ملائماً إلا إذا أشارت جميع الأسئلة إلى الأسرة كوحدة إحصائية، لكن الأمر ليس كذلك، كما أشير أعلاه. فبعض الأسئلة تشير إلى وحدات إحصائية ثانوية تبدو في الأعداد المتغيرة داخل الأسرة المعيشية، كالأشخاص والمحاصيل والبنود الاستهلاكية وما أشبه ذلك. وإن تخزين سن وجنس كل فرد من أفراد الأسرة المعيشية كمتغيرات مختلفة على مستوى الأسرة المعيشية سيكون فيه تبديد لأن عدد المتغيرات اللازمة سيُعرَّف بحجم أكبر الأسر المعيشية بدلاً من متوسط حجم الأسرة المعيشية.

١٠٤ - سوف يُستخدَم الملف المتغير الطول في الأحوال العادية لقيد بيانات استقصاءات الأسر المعيشية. ولأن الأسر المعيشية تختلف في حجمها وتكوينها، ستلزم سجلات متغيرة الحجم أثناء قيد البيانات. ومع أن كل نوع من أنواع السجلات سيكون ثابتاً من حيث الطول والقالب، ستوجد أنواع مختلفة من السجلات في الملف الواحد. وسيكون كل ملف أساساً صورة حاسوبية للاستبيانات كما مُلئت. وسوف يشكّل كل خط أو مجموعة في الاستبيان سجلاً. وسيبدأ كل سجل بخيط من المحددات التي تصل السجل بالأسرة المعيشية ووحدة الملاحظة، وهلم جراً. هذه الطريقة تستخدم الحيز بكفاءة لكنها غير ملائمة للإعداد في وقت لاحق، حيث يصبح الرجوع إلى البيانات من ملفات مختلفة مسألة حرجة.

١٠٥ - يستخدم نظام إعداد تعداد السكان والاستقصاء (CSPro)، مثلاً، هيكَل ملفات يتناول التعقيدات التي تظهر في التعامل مع وحدات إحصائية مختلفة كثيرة تناولاً جيداً، بينما يقلل متطلبات التخزين إلى الحد الأدنى ويتقابل جيداً مع البرامج الحاسوبية الإحصائية في المرحلة التحليلية.

١٠٦ - يحافظ هيكَل البيانات على تطابق واحد مع واحد بين كل وحدة إحصائية ملاحظة والسجلات الموجودة في الملفات الحاسوبية، مستخدماً نوعاً مختلفاً من السجلات لكل نوع من الوحدات الإحصائية. فلإدارة البيانات الواردة في قائمة الأسر المعيشية، مثلاً، يُعيّن نوع من السجلات للمتغيرات التي في القائمة، وتخزن البيانات المقابلة لكل فرد في سجل مستقل من ذلك النوع. وكذلك في وحدة استهلاك الأغذية، يقابل نوع من السجلات البنود الغذائية، وتُخزن البيانات المقابلة لكل بند فردي في سجلات مختلفة من هذا النوع.

١٠٧ - يُسمَح بتفاوت عدد السجلات التي في كل نوع من أنواع السجلات. فهذا يوفر حيز التخزين اللازم، لأن الملفات لا تحتاج إلى السماح لكل حالة بأن تكون أكبر ما يمكن.

١٠٨ - عقب إدراج المحددات، تُسجل البيانات الفعلية، التي يسجلها الاستقصاء لكل وحدة بعينها، في ميادين ثابتة الطول بنفس ترتيب الأسئلة في الاستبيان. وتُخزن جميع البيانات الفعلية بالقالب القياسي للشفرة المعيارية الأمريكية لتبادل المعلومات [المسماة أسكي: ASCII].

١٠٩ - يوردُ مونوز (٢٠٠٣) والبنك الدولي (١٩٩١) مناقشات أكثر تفصيلاً بشأن إدارة ملفات استقصاءات الأسر المعيشية.

ترتيب هياكل مجموعات البيانات لاستخدامها في عمليات أخرى

١١٠ - لتيسير التحليل الملائم يجب أن تضم قاعدة البيانات المرتبطة جميع المعلومات عن إجراء المعاينة؛ ورقاع تعريف لطبقات تصميم العينة، و وحدات المعاينة الأولية، ووحدات المعاينة الثانوية، إلخ؛ ومُعاملات ترجيح العينة لكل وحدة معاينة. هذه المعلومات تلزم لتقدير الإحصاءات اللازمة، ولتقدير أخطاء المعاينة في هذه التقديرات أيضاً.

١١١ - غالباً ما يكون ضرورياً، بعد قيد البيانات، أن يُعاد تشكيل هيكل مجموعة البيانات وتوليد ملفات جديدة، وتسجيل بعض ميادين البيانات القائمة من أجل تعريف متغيرات جديدة أنسب للجدولة والتحليل. ربما يكون هذا ضرورياً للسماح بأداء بعض العمليات في البيانات، بما في ذلك عملية التقدير.

١١٢ - ربما يحتوي أول ملف بيانات كامل للاستقصاء في الواقع معلومات عن وحدات معاينة من مختلف المجموعات السكانية (روزن، ١٩٩١). فلإجراء استقصاء لميزانية الأسر المعيشية، مثلاً، يمكن أن توجد البيانات عن الأسر المعيشية المعاينة والأشخاص المعيّنين في نفس الملف الأولي. ولتقدير الخصائص الإحصائية للأسر المعيشية وللأشخاص، يحتاج المرء إلى ملف فيه سجل واحد لكل أسرة معيشية معاينة وسجل واحد لكل شخص معاين، على حدة. وتُستخدم مجموعات البيانات أو الملفات، المستندة إلى الأسر المعيشية كوحدات (مواضيع)، لإنتاج إحصائيات (جداول) عن أشخاص من أسر معيشية خاصة. وتُستخدم مجموعات البيانات أو الملفات، المستندة إلى الأفراد كوحدات (مواضيع)، لإنتاج إحصائيات (جداول) عن أفراد من أسر معيشية.

١١٣ - كما يتجلى ممّا ذكر أعلاه، يوجد في الأحوال النمطية نوعان رئيسيان من الملفات، من استقصاءات الأسر المعيشية: ملفات الأسر المعيشية، وملفات الأفراد (الأشخاص كلاً على حدة). الملفات، في معظم الحالات، ملفات أسر معيشية بمعنى أنها تحمل قيماً لمتغيرات الأسر المعيشية (متغيرات تتصل بوحدة ملاحظة أو موضوع "أسرة معيشية") بعض هذه الملفات ملفات فردية (ملفات أشخاص)، بمعنى أنها تحمل قيماً لمتغيرات بشأن أفراد (متغيرات تتصل بوحدة ملاحظة أو موضوع "شخص"). وستحتوي ملفات البيانات الكاملة والنهائية (مجموعات البيانات) على معلومات عن جميع الأسر المستجيبة والأفراد المستجيبين من كل وحدة من وحدات المعاينة الأولية المشمولة بالاستقصاء.

١١٤ - يوضّح الجدول ٩ - ٢ الوارد أدناه كيف أُعيد تنظيم الملف الكبير للاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي في سنة ١٩٨٧، لتيسير مزيد من إعداد البيانات. والمثال الثاني، في الجدول ٩ - ٣ يعرض ملفات نمطية لاستقصاء ميزانيات الأسر المعيشية. وتستند هذه الأمثلة إلى موادّ عمل عليها لاغرلوف (١٩٨٨) وروزن (١٩٩١).

الجدول ٩ - ٢  
ملفات الأسر المعيشية والأفراد المستخدمة في الاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في  
زمبابوي في سنة ١٩٨٧

الملف	النوع	المضمون
أسرة معيشية	ملف الأسرة المعيشية	تعريف الأسرة (الإقليم، المنطقة، المقاطعة، إلخ.) إجابات على كل الأسئلة المتصلة بالأسرة المعيشية متغيرات، مثل حجم الأسرة المعيشية مستمدة (من ملفات أفراد الأسرة)، إلخ.
شخص	ملف فردي	تعريف الأسرة المعيشية زائداً تعريف الشخص الخصائص الديمغرافية: السن، نوع الجنس، الحالة الزوجية، فرد عادي في الأسرة المعيشية، العلاقة برب الأسرة المعيشية
شخص متوفى	ملف فردي	تعريف الأسرة المعيشية زائداً تعريف الشخص المتوفى تفاصيل عن الشخص المتوفى الذي كان فرداً عادياً في الأسرة: نوع الجنس، السن (السن عند الوفاة)
امراة البالغة ١٢ سنة فما فوق	ملف فردي HID, PID	تفاصيل عن كل امرأة، في الأسرة المعيشية، لا يقل عمرها عن ١٢ سنة

الجدول ٩ - ٣  
الملفات النمطية لاستقصاء ميزانيات الأسر المعيشية

الملف	النوع	المضمون
الأسرة المعيشية	ملف الأسرة المعيشية	تعريف الأسرة المعيشية (الإقليم، المنطقة، المقاطعة، إلخ.) إجابات على كل الأسئلة المتصلة بالأسرة المعيشية متغيرات، مثل حجم الأسرة المعيشية مستمدة (من ملفات أفراد الأسرة)، إلخ.
أفراد الأسرة	ملف فردي	تعريف الأسرة المعيشية زائداً تعريف الشخص الخصائص الديمغرافية: السن، نوع الجنس، الحالة الزوجية، المستوى التعليمي، إلخ.، لأفراد الأسرة معلومات عن الأنشطة الرئيسية: وضع العمالة، المهنة، إلخ.
الدخل	ملف فردي	تعريف الأسرة المعيشية زائداً تعريف الشخص، زائداً تعريف مصادر الدخل:
الغذاء	ملف الأسرة المعيشية	تعريف الأسرة المعيشية زائداً تعريف بنود الغذاء النفقات على الأغذية:
سلع أخرى غير معمرة	ملف الأسرة المعيشية	تعريف الأسرة المعيشية زائداً تعريف بنود السلع النفقات على السلع:
سلع معمرة	ملف الأسرة المعيشية	تعريف الأسرة المعيشية زائداً تعريف بنود السلع المعمرة النفقات على السلع المعمرة:
الزراعة	ملف الأسرة المعيشية	تعريف الأسرة المعيشية زائداً تعريف البنود الزراعية النفقات على الزراعة:
رأس المال الزراعي	ملف الأسرة المعيشية	تعريف الأسرة المعيشية زائداً تعريف بنود رأس المال الزراعي النفقات على رأس المال الزراعي:

١١٥ - لأغراض الجدولة من الضروري وضع ملفّ مستوٍ لمعظم البرامج الحاسوبية الإحصائية. ويحتاج كثير من البرامج العامة المتاحة إلى بيانات في قالب المستوي. وفي الملفّ المستوي يكون لجميع السجلات نفس مجموعة المتغيرات أو الميادين، وتكون كلها ذات طول واحد. يوصف الملفّ بأنه مستوٍ عندما تكون جميع ميادين البيانات موجودة لكل مستجيب. وتُرتَّب ميادين البيانات ترتيباً متماثلاً في كل سجلّ، ويوجد عدد ثابت من السجلات وتكون متماثلة في التصميم. يعرض الجدول ٩ - ٤ قالب الملفّ المستوي لملفّ الأسرة المعيشية المستخدم في الاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي في سنة ١٩٨٧.

الجدول ٩ - ٤

قالب الملفّ المستوي كما استُخدم في ملفّ الأسرة المعيشية للاستقصاء الديمغرافي الذي أُجري بين عمليتي تعداد للسكان في زمبابوي في سنة ١٩٨٧

المتغير الترجيحي	القيم المتغيرة		مَعْلَمَات تصميم المعاينة								تعريف			
	z	y	x	$m_{hi}$	$M_{hi}$	$S_{hi}$	$b_{hr}$	$R_{hi}$	$a_{hi}$	$S_{hi}$	منطقة	أُسرة	قسم	طبقة
W											معيشية <td>تعداد <td></td> <td></td> </td>	تعداد <td></td> <td></td>		
$W_{hrj}$	$Z_{hrj}$	$Y_{hrj}$	$X_{hrj}$								h	i	r	h

١١٦ - يحتوي ملفّ الأسرة المعيشية على سجلّ واحد لكل أسرة معيشية ملاحظة، ويحتوي كل سجلّ على معلومات عن:

- تعريف الأسرة المعيشية
- مَعْلَمَات تصميم المعاينة
- القيم الملاحظة لمتغيرات (الأسرة المعيشية)
- مَعْلَمَات الترجيح

تعريف الأسرة المعيشية: تقول التوليفة  $hrj$  إن الأسرة المعيشية  $j$  تخص منطقة التعداد  $i$  في القسم الفرعي  $r$  من الطبقة  $h$ .

مَعْلَمَات تصميم المعاينة: هذه المَعْلَمَات في هذا المثال بالذات هي كما يلي:

- $S_b$  = عدد الأسر المعيشية في الطبقة المعاينة في سنة ١٩٨٢
- $a_b$  = حجم عيّنة مناطق التعداد في طبقة المعاينة
- $R_b$  = عدد الأقسام الفرعية الممتلئة في العيّنة من طبقة المعاينة
- $b_{br}$  = عدد مناطق التعداد من القسم الفرعي
- $S_{bi}$  = عدد الأسر المعيشية في منطقة التعداد في سنة ١٩٨٢
- $M_{bi}$  = عدد الأسر المعيشية في منطقة التعداد في سنة ١٩٨٧

•  $m_{hi}$  = حجم عينة الأسر المعيشية من منطقة التعداد

القيم المتغيرة الملاحظة:  $x, y, z$  تشير إلى متغيرات الأسر المعيشية.

القيم المتغيرة لمعاملات الترجيح:  $w$  تشير إلى متغير مُعاملات الترجيح للأسرة المعيشية.

١١٧ - يشبه تنظيم ملف الشخص تنظيم ملف الأسرة المعيشية المقدم أعلاه. الفرق الطفيف هو أن التعريف سيكون تعريف الشخص، وأن الرقم القياسي ( $k$ ) سيكون للشخص الفرد، بينما تشير "المتغيرات" إلى متغيرات الأفراد.

١١٨ - يتعين تنظيم مجموعات بيانات الاستقصاء كملفات مستوية منفصلة فقط (ملف واحد لكل نوع من أنواع السجلات)، من أجل نشرها، لأن قالب المجال الثابت الطول في الهيكل المحلي يلائم أيضاً نقل البيانات إلى النظم المعيارية لإدارة قواعد البيانات من أجل مواصلة معالجتها، أو إلى برامج إحصائية قياسية من أجل جدولتها وتحليلها. ونقل البيانات إلى نظم إدارة قواعد البيانات سهل للغاية لأن الهيكل المحلي يتحول بصورة مباشرة تقريباً إلى القالب المعياري لقواعد البيانات الذي يقبله الجميع كمدخل للجدول الفردية (في هذه الحالة تعمل محددات التسجيل بمثابة وصلات ترابط طبيعي بين الجداول) (مونوز، ٢٠٠٣).

#### ٩ - ٤ - ٣ - ٥ إجراءات التقدير وحساب مُعاملات الترجيح

١١٩ - يعطي الفصل السادس الأساس المنطقي للترجيح وطريقة لحساب مُعاملات الترجيح في استقصاء الأسر المعيشية (انظر الإشارات المرجعية لروزن في نهاية هذا الفصل). ويشار إلى خوارزمية حساب، نابعة من قيم ملاحظة لتقديرات خصائص إحصائية، لإجراء تقدير النقاط. في الخطوة الأولى لتقدير النقاط يُحسب مُعامل ترجيح لكل موضوع مجاب عليه، ثم تُحسب "مجاميع" التقديرات بجمع قيم الملاحظة المرجحة (بضرب القيمة الملاحظة بمُعامل ترجيحها).

١٢٠ - يقدم مونوز (٢٠٠٣) وصفاً جيداً لكيفية استخدام نظام حاسوبي هو الجدول الإلكتروني لمايكروسوفت إكسل في تنفيذ إجراء المعاينة لاستقصاء الأسر المعيشية بإرشاده في مراحل الرئيسية: تنظيم إطار المرحلة الأولى؛ واختيار وحدات المعاينة الأولية مع تناسب احتمال الاختيار مع الحجم؛ وحساب احتمالات الاختيار ومُعاملات ترجيح المعاينة المقابلة لها.

١٢١ - التشكيل الفعلي لأدوات التقييم المرجحة تشكيل مباشر. يُبدأ بمجموعة بيانات العينة الأصلية وإنشاء مجموعة بيانات جديدة بضرب كل ملاحظة برقم مُعامل الترجيح المحدد لها، ثم استخدام المعادلات المعيارية لحساب مَعْلَمَة الحدود باستخدام مجموعة البيانات المرجحة.

١٢٢ - غير أنه ينبغي ملاحظة أن مُعاملات الترجيح الدقيقة يجب أن تشمل ثلاثة عناصر (يانسانة، ٢٠٠٣) من بينها التعديلات المختلفة اللازمة (انظر الفصل السادس أيضاً). فمُعاملات الترجيح الأساسية هي سبب التفاوت في تفاوت احتمالات الاختيار من بين مجموعات مختلفة من الأسر المعيشية، كما يشترط التصميم الأولي للاستقصاء. والتعديل الثاني يتعلق بالتفاوت في عدم الاستجابة من المجالات أو المجموعات الفرعية. وأخيراً يمكن أن يوجد في بعض الحالات تعديلات لاحقة للتقسيم إلى طبقات تكون لازمة لجعل بيانات الاستقصاء تتفق مع توزيعات من مصدر مستقل كأحدث تعداد للسكان، مثلاً.

١٢٣ - ينشأ تعقيد آخر في عملية التقدير من الطلب المتزايد على الإحصاءات على مستوى المجال. والمجال، كما نُوقش في الفصل الثالث، مجموعة فرعية يُستصوب إجراء تقديرات مستقلة لها. يُمكن، في الأحوال العادية تحديدها في مرحلة تصميم العيّنة، لكنه يمكن أيضاً حسابها من البيانات المستمدة. ويمكن أيضاً أن يكون المجال طبقة، أو مجموعة طبقات، أو مناطق إدارية (إقليم، مقاطعة، مستوى ريفي/ حضري، إلخ). ويمكن أيضاً تعريفه من حيث الخصائص الديمغرافية أو الاجتماعية - الاقتصادية (مثل ذلك السن أو نوع الجنس أو المجموعة الإثنية أو الفقراء، إلخ). وما يلي هو محاولة لوصف كيفية تشكيل مجموعات البيانات لتسهيل تقدير المجالات.

١٢٤ - نبدأ بتصوّر ملف بيانات (ملاحظة) (ملف الأسرة المعيشية، مثلاً) كما ورد أعلاه بشأن الاستقصاء الديمغرافي المُجرى بين عمليّتي تعداد للسكان في زبابوي. هذا الملف فيه سجل واحد لكل أسرة معيشية مشمولة بالعيّنة. وفي نهاية عملية الاستقصاء يجب أن يتضمن الملف المعلومات التالية عن كل أسرة معيشية:

(أ) تعريف الأسرة المعيشية؛

(ب) مَعْلَمَات المعايينة؛

(ج) قيم متغيرات الدراسة  $x, y$  and  $z$ ؛

(د) قيمة مُعاملٍ ترجيحٍ لتقدير الأسرة المعيشية؛

(هـ) مسألة ما إذا كانت الأسرة المعيشية تنتمي إلى الفئة ج (c)؛

(و) مسألة ما إذا كانت الأسرة المعيشية تنتمي إلى المجال ز (g).

١٢٥ - فيما يلي ملاحظات تعريفية لهذه المعلومات (باستثناء مَعْلَمَات المعايينة):

- $HID =$  علامة تعريف للأسر المعيشية المدرجة في العيّنة. ولأجل البساطة نستخدم الأرقام المسلسلة،  $1, 2, n, \dots$ . ومن ثم ترمز  $n$  إلى مجموع حجم العيّنة.
- $x, y, \text{ and } z =$  هي القيم الملاحظة للمتغيرات  $X, Y, \text{ and } Z$  للأسرة المعيشية.
- $c = 1$  إذا كانت الأسرة المعيشية تنتمي إلى الفئة  $c$ ، وإلا فهي تساوي صفرًا.
- $g = 1$  إذا كانت الأسرة المعيشية تنتمي إلى الفئة  $g$ ، وإلا فهي تساوي صفرًا.
- $w =$  مُعاملٍ ترجيحٍ التقدير للأسرة المعيشية.

١٢٦ - قيم متغيرات المؤشرين  $c$  و  $g$  مستمدة في العادة من متغيرات أخرى وليست ملاحظة بصورة مباشرة. فيمكن أن ترمز الفئة  $c$ ، مثلاً، إلى عبارة "تحت خط مستوى الفقر". ولا تُسأل الأسرة المعيشية إن كانت تنتمي إلى تلك الفئة أم لا. وإنما التصنيف مستمد، مثلاً، من بيانات دخل الأسرة المعيشية وهي خط مشروط لمستوى الفقر. وكذلك يُقتضى في الغالب أن تلزم متغيرات أخرى لتقرير ما إذا كانت الأسرة المعيشية تنتمي إلى مجال دراسة محدد أم لا (فالمجال  $g$ ، مثلاً، يمكن أن يتألف من الأسر المعيشية التي لديها ثلاثة أطفال فأكثر). وينبغي، في مرحلة التقدير، أن تكون هذه المؤشرات متاحة في ملف الملاحظات.

١٢٧ - عندما تكون جميع البيانات موجودة في ملف الملاحظات ستبدو كما هي في الجدول ٩ - ٥ أدناه، ما عدا مَعْلَمَات المعايينة، فهي غير مشمولة.

الجدول ٩ - ٥

## ملف الملاحظات وفيه البيانات النهائية لمنحدرات استقصاء الأسر المعيشية

W	G	C	Z	Y	X	HID
W <sub>1</sub>	g <sub>1</sub>	C <sub>1</sub>	Z <sub>1</sub>	y <sub>1</sub>	x <sub>1</sub>	1
W <sub>2</sub>	g <sub>2</sub>	C <sub>2</sub>	Z <sub>2</sub>	y <sub>2</sub>	x <sub>2</sub>	2
W <sub>3</sub>	g <sub>3</sub>	C <sub>3</sub>	Z <sub>3</sub>	y <sub>3</sub>	x <sub>3</sub>	3
.	.	.	.	.	.	.
.	.	.	.	.	.	W.
W <sub>n</sub>	g <sub>n</sub>	C <sub>n</sub>	.	y <sub>n</sub>	x <sub>n</sub>	N

١٢٨ - اقتصر المناقشة الواردة أعلاه على تقدير الخصائص الإحصائية للأسر المعيشية. وتُجرى تقديرات الخصائص الإحصائية للأشخاص على غرار ذلك. ومُعالمٍ ترجيح التقديرات المتعلقة بالأشخاص، هو بوجه عام نفس مُعاملٍ الترجيح المستخدم للأسرة المعيشية التي ينتمي إليها الشخص المعني. ولما كان جميع أفراد الأسرة المعيشية النمطية مدرجين في الاستبيان، لا يُدرج الشخص المعين في عينة الأشخاص إلا، ونؤكد كلمة "إلا"، إذا كانت أسرته مدرجة في عينة الأسر المعيشية. ومن ثم يكون احتمال إدراج الشخص هو نفس احتمال إدراج الأسرة المعيشية التي ينتمي إليها الشخص المذكور. غير أنه تجدر ملاحظة أن هذا الافتراض لا يصح عندما تُختار عينات فرعية داخل الأسر المعيشية. ففي بعض تصاميم العينات، مثلاً، ربما يدعو الإجراء إلى اختيار شخص بالغ واحد فقط من كل أسرة معيشية، أو اختيار ذكر واحد وأنثى واحدة من كل أسرة؛ ففي هذه الحالات يُحسب مُعاملٍ الترجيح للشخص المختار (أو الأشخاص المختارين) حساباً مستقلاً ولا يساوي مُعاملٍ الترجيح المحدد للأسرة.

١٢٩ - من أجل الاكتمال، يجب أن يحتوي جزء من إجراء التقدير على تقديرات لأخطاء المعاينة (أو الأخطاء المعيارية) في الاستقصاء، لا سيما في أهم الإحصائيات التي تُؤكَّد وتُنشر على الملأ. وقد خُصَّص الفصل السابع من الدليل بأسره لهذا الموضوع.

## ٩ - ٤ - ٣ - ٦ الجدولة، ومجموعات البيانات للجدولة، وقواعد البيانات

١٣٠ - توجد ثلاثة مخرجات أساسية رئيسية من الاستقصاء الإحصائي (سندغرين، ١٩٩٥):

- البيانات الكلية: "الإحصائيات" التي تمثل تقديرات لخصائص إحصائية معينة؛ إنتاج هذه الإحصائيات هو الغرض الأساسي من الاستقصاء الجاري تنفيذه.
- البيانات الجزئية: "ملاحظات المواضيع الفردية" التي تقوم عليها البيانات الكلية التي ينتجها الاستقصاء؛ هذه البيانات أساسية لاستخدامها في المستقبل ولتفسير نتائج الاستقصاء.
- البيانات الأولية: "هي البيانات التي تصف المعنى والدقة والإتاحة وغيرها من الميزات الهامة التي تقوم عليها البيانات الكلية والجزئية"؛ وهذه أساسية لتعريف البيانات الإحصائية ذات الصلة لمشكلة محددة تعريفاً صحيحاً واستعادة هذه البيانات، وكذلك لتفسير البيانات الإحصائية تفسيراً صحيحاً وإعادة استخدامها.

وسيكون من المفيد أيضاً النظر في تصميم جداول البيانات المتعددة الأبعاد (المكعبات)، من أجل توفير إمكانيات أكثر مرونة للوصول إلى نتائج الاستقصاء، وكذلك زيادة القدرة على الرجوع إليها بواسطة المواقع الإلكترونية، مثلاً.

١٣١ - من شأن برنامج استقصاءات الأسر المعيشية أن ينتج في نهاية المطاف وضعاً يستند فيه تحصيل البيانات إلى توليفة من البيانات على المستويين الجزئي والكلي. ولتحقيق ذلك، يجب أن يكون ثمة وصف مفصل لهيكل المعلومات المجموعة بواسطة استقصاءات متعددة.

١٣٢ - ينبغي النظر في تخزين البيانات في ثلاث مراحل (لونديل، ٢٠٠٣):

- التخزين: ينبغي أثناء قيد البيانات أن تُخزن البيانات بطريقة تعمل بواسطتها جيداً مع قيد البيانات والطرائق المستخدمة في تنظيف البيانات، كما سبقت مناقشته.
- وضعها في المستودعات: عندما تقيّد البيانات وتنظّف، يجب إضافتها إلى مستودع يُعدّل هيكله على نحو يناسب أدوات وطرائق تحليل البيانات ونشرها.
- وضعها في المحفوظات (الأرشيف): ينبغي وضع بيانات المشروع في المحفوظات (الأرشيف) بطريقة تتفق مع المعايير الطويلة الأجل لكي تُضمّن استعادة البيانات في المستقبل من دون تعقيد.

١٣٣ - يمكن إنشاء مستودع بيانات يحتوي على بيانات نظيفة بعدة طرق، باستخدام واحدة من الطرائق التالية (انظر التذييل الملحق بهذا الفصل للاطلاع على معلومات عن البرامج الحاسوبية):

- الملفات المستوية
- قاعدة البيانات المنطقية (جهاز خادم مايكروسوفت للاستفسارات اللغوية المنظمة (SQL))
- البرامج الحاسوبية (نظام التحليل الإحصائي (SAS)، مثلاً، أو البرامج الإحصائية للعلوم الاجتماعية (SPSS))

١٣٤ - لحفظ البيانات النهائية حفظاً طويلاً في المحفوظات (الأرشيف) يوجد خيار رئيسي واحد. يجب حفظ البيانات بقالب أسكي (ASCII) بسيط للملفات المستوية وإرفاق توصيفات للسجل بها. ويمكن لمعظم نُظُم قواعد البيانات والبرامج الحاسوبية الإحصائية إرسال البيانات إلى هذه الملفات دون كبير عناء، ويمكن أيضاً أخذ البيانات منها بسهولة.

## ٩ - ٥ تذييل

## خيارات البرامج الحاسوبية للخطوات المختلفة في إعداد بيانات الاستقصاء

خيارات البرامج الحاسوبية	نوع العملية
Microsoft Structured Query Language (SQL) Server 2000, Standard Edition Microsoft Access Statistical Analysis System (SAS)	نظام إدارة قاعدة البيانات
Visual Basic Microsoft Access Integrated Microcomputer Processing System (IMPS) Census and Survey Processing System (CSPRO)	قيد البيانات وتحريرها
Statistical Analysis System (SAS) Statistical Package for the Social Sciences (SPSS) Microsoft Access Microsoft Excel	استرجاع البيانات
Microsoft Word Microsoft Excel Statistical Analysis System (SAS) Statistical Package for Social Sciences (SPSS)	الجدولة والتحليل والعرض
CENVAR: variance calculation component of (IMPS) Computation and Listing of Useful STatistics on Errors of Sampling (CLUSTERS) Integrated System for Survey Analysis (ISSA) Survey Data Analysis (SUDAAN) Statistical Analysis System (SAS) Statistical Package for the Social Sciences (SPSS) Cluster Analysis and Regression Package (PC-CARP)	تقدير التفاوت

## ٩ - ٥ - ١ مكتب مايكروسوفت (The Microsoft Office)

مكتب مايكروسوفت الذي طورته شركة مايكروسوفت مجموعة برامج شاملة تضم بضعة برامج حاسوبية من بينها:

- Microsoft Office Access، هو برنامج إدارة قاعدة البيانات من برامج المكتب، يوفر سهولة محسنة في استخدام، وقدرة موسعة على استيراد، المعلومات من الملفات وتصديرها إليها، والعمل مع ملفات البيانات بـ (Extensible Markup Language (XML).
- Microsoft Office Excel، هو برنامج الجدول الإلكتروني من برامج المكتب، الذي يشمل دعماً لبرنامج XML، وملاحج جديدة تُسهّل تحليل المعلومات وإطلاع الغير عليها.
- Microsoft Office Word، هو برنامج تحضير الكلمات من برامج المكتب.

- Microsoft SQL Server 2000، هو الجهاز الخادم لقاعدة البيانات من برامج المكتب، للقدرات على إدارة المشروع التجاري والموارد إدارة تامة.
  - Microsoft Office Outlook، هو برنامج إدارة المعلومات الشخصية والاتصالات من برامج المكتب، الذي يقدم مكاناً موحّداً لإدارة البريد الإلكتروني، والتقويم السنوي، إلخ.
- الموقع على الإنترنت: <http://www.microsoft.com/office/system/overview.msp#EDAA>.

### ٩ - ٥ - ٢ فيجيوال بيسك (Visual Basic)

أطلقت شركة مايكروسوفت برنامج فيجيوال بيسك في سنة ١٩٨٧. هذا البرنامج ليس مجرد لغة برمجة حاسوبية وإنما هو أيضاً بيئة تطوير غرافيكية كاملة. تسمح هذه البيئة للمستخدمين الذين لا توجد لديهم خبرة تذكر في البرمجة بتطوير تطبيقات نوافذ مايكروسوف (Microsoft Windows) مفيدة بسرعة، لديها القدرة على استخدام مواضيع برنامج ربط المواضيع وتركيبها (Object Linking and Embedding (OLE)، مثل الجدول الإلكتروني المنتج ببرنامج إكسل (Excel Spreadsheet). توجد لدى برنامج فيجيوال بيسك أيضاً القدرة على تطوير برامج يمكن استخدامها كتطبيقات ابتدائية لنظام قواعد بيانات، تعمل بمثابة سطح بيني للمستخدم يجمع مدخلاته ويعرض المخرجات مُشكّلة في قالب أكثر جاذبية وفائدة مما تستطيع أشكال كثيرة من برنامج SQL إنتاجه.

النقطة التي هي أكثر ما تكون مساعدة على بيع فيجيوال بيسك هي السهولة التي يمكن بها المستخدم من وضع برامج غرافيكية جميلة المنظر دون تشفير يذكر من قبل المبرمج. الموضوع الرئيسي في فيجيوال بيسك يُسمى القالب، وهذا يسهّل تطوير شاشات قيد البيانات.

الموقع على الإنترنت: <http://www.engin.umd.umich.edu/CIS/course.des/cis400/vbasic/vbasic.html>.

### ٩ - ٥ - ٣ سنفار (CENVAR)

سنفار هو عنصر حساب التفاوت في نظام الإعداد المتكامل بالحاسوب الصغير (IMPS)، وهو عبارة عن سلسلة من البرامج الحاسوبية لقيد بيانات تعداد السكان والاستقصاء وتحريرها وجدولتها وتقديرها وتحليلها ونشرها. ومكتب تعداد السكان في الولايات المتحدة هو الذي طوّر (IMPS).

الموقع على الإنترنت: <http://www.census.gov/ipc/www/imps/>.

### ٩ - ٥ - ٤ بي سي كارب (PC CARP)

يستند سنفار إلى برنامج تحليل المجموعات وتراجُعها لبرامج الحواسيب الشخصية (Cluster Analysis and Regression Package for Personal Computers (PC CARP) الذي طوره أصلاً جامعة ولاية أيوا. ويستخدم بي سي كارب إجراء الضبط الخطي لحساب التفاوتات.

الموقع على الإنترنت: <http://www.census.gov/ipc/www/imps/>.

### ٩ - ٥ - ٥ نظام إعداد تعداد السكان والاستقصاء (CSPPro) (Census and Survey Processing System)

نظام CSPPro برنامج حاسوبي لمجال عام يُستخدم لقياد بيانات تعداد السكان والاستقصاء. صُمم هذا البرنامج ونُفذ في مجهود مشترك بين مُطوّري IMPS و ISSA : وهم مكتب تعداد السكان في الولايات المتحدة، وماكرو إنترناشيونال (Macro International)، وسيربرو أس آيه (Serpro, S A). وقَدّم تمويل تطوير البرنامج مكتب السكان في وكالة الولايات المتحدة للتنمية الدولية. وقد صُمم البرنامج ليحل في النهاية محل البرنامجين IMPS و ISSA.

الموقع على الإنترنت: <http://www.census.gov/ipc/www/imps/>.

### ٩ - ٥ - ٦ حساب إحصاءات مفيدة لأخطاء المعاينة ووضع قائمة بها (CLUSTERS) (Computation and Listing of Useful Statistics on Errors of Sampling)

حساب الإحصاءات المفيدة لأخطاء المعاينة ووضع قائمة بها (CLUSTERS) برنامج طُوّر أصلاً لحساب أخطاء المعاينة في برنامج استقصاء الخصوبة العالمية (WFS). وهو يستخدم أسلوب تايلور للضبط الخطي لحساب أخطاء المعاينة. واستُخدم أيضاً لحساب أخطاء المعاينة لاستقصاءات مختلفة للأسر المعيشية، لا سيما تلك التي أُجريت في إطار برامج الاستقصاءات الديمغرافية والصحية في كثير من البلدان النامية (انظر فيرما، ١٩٨٢).

### ٩ - ٥ - ٧ النظام المتكامل لتحليل الاستقصاءات (ISSA) (Integrated System for Survey Analysis)

طورت شركة ماكرو إنترناشيونال إنك. (Macro International Inc.) النظام المتكامل لتحليل الاستقصاءات خصيصاً لبرنامج الاستقصاءات الديمغرافية والصحية. واستُخدم لجميع جوانب إعداد البيانات، وقياد البيانات، وتحريها، وجدولتها. وفيه أيضاً وحدة لتناول أخطاء المعاينة كنسب الخصوبة والوفيات، باستخدام طريقة المطواة (انظر ماكرو إنترناشيونال إنك.، ١٩٦٦ (1966) (Macro International Inc.)).

### ٩ - ٥ - ٨ نظام التحليل الإحصائي (SAS)

نظام التحليل الإحصائي (SAS)، الذي طورته شركة نظام التحليل الإحصائي (SAS Inc.) في سنة ١٩٦٦، برنامج حاسوبي لتحليل البيانات، وإدارة الملفات، وحساب أخطاء المعاينة (انظر آن وواتس للاطلاع على عرض لبعض من أحدث الملامح في نظام التحليل الإحصائي).

### ٩ - ٥ - ٩ البرنامج الحاسوبي الإحصائي للعلوم الاجتماعية (SPSS)

البرنامج الحاسوبي الإحصائي للعلوم الاجتماعية (Statistical Package Social Sciences)، الذي طورته شركة البرنامج الحاسوبي الإحصائي للعلوم الاجتماعية، إنك. (SPSS, Inc.)، برنامج حاسوبي لتحليل البيانات ومعالجة الملفات، إلخ. (انظر: SPSS, Inc. (1988) للاطلاع على عرض لبعض من أحدث الملامح).

## ٩ - ٥ - ١٠ تحليل بيانات الاستقصاء (SUDAAN) (SURvey DATA ANALysis)

تحليل بيانات الاستقصاء (SUDAAN)، الذي طوره معهد مثلث البحوث (Research Triangle Institute) في حديقة مثلث البحوث، كارولاينا الشمالية (Research Triangle Park, North Carolina)، برنامج حاسوبي للاستقصاء الشامل للعيّنات (والبيانات المرتبطة به)، وفيه قوى للتحليل الوصفي والنموذجي. (يمكن الاطلاع على مزيد من التفاصيل في شاه، وبارنويل، وبيلبير (١٩٩٦)).

### المراجع ومواد أخرى للقراءة

- An, A., and D. Watts (2001)., *New SAS Procedures for Analysis of Sample Survey Data*. SUGI paper, No. 23, Cary, North Carolina: SAS Institute, Inc.
- Arnic, and others (2003). "Metadata production systems within Europe: the case of the statistical system of Slovenia, Paper presented at the Metadata Production Workshop, Luxembourg. Eurostat Document 3331.
- Australian Bureau of Statistics (2005). *Labour statistics: concepts, sources and methods*. Canberra: Statistical Concepts Library.
- Available from [www.abs.gov.au/AUSSTATS/abs@nsf/DirClassManually Catalogue/59D849DC7BOIFCC ECA257/10FOOI F6E5B](http://www.abs.gov.au/AUSSTATS/abs@nsf/DirClassManually%20Catalogue/59D849DC7BOIFCC%20ECA257/10FOOI%20F6E5B). Open Document Catalogue No. 6102.0.55.001.
- Backlund, S. (1996). *Future directions on IT issues*. Mission report to National Statistical Centre, Lao People's Democratic Republic, Vientiane.
- Brogan, D. (2003). *Comparison of Data Analysis Software Suitable for Surveys in Developing Countries*, United Nations Statistics Division, New York.
- Central Statistics Office, Namibia (1996). *The 1993/1994 National Household Income and Expenditure Survey (NHIES)*. Administrative and technical report, Windhoek: National Planning Commission.
- Chromy, J., and S. Abeysasekara (2003). *Analytical Uses of Survey Data*, United Nations Statistics Division, New York.
- Chronholm, P. and Edsfeldt (1996). *Course and seminar on systems design*, Mission report to Central Statistics (CSS), Pretoria.
- Giles, M. (1996). *Turning Data into Information: A Manual for Social Analysis*. Canberra: Australian Bureau of Statistics.
- Glewwe, Paul (2005). *An overview of questionnaire design for household surveys in developing countries*. In *Household Sample Surveys in Transition and Developing Countries*. Studies in Methods, No. 96. Sales No. E.05.XVII.6.
- Graubard, B., and E. Korn (2002). *The Use of Sampling Weights in the Analysis of Survey Data*, United Nations Statistics Division, New York.
- International Labour Office (1990). *Survey of Economically Active Population, Employment, Unemployment and Underemployment: ILO Manual on Concepts and Methods*. Geneva: International Labour Office.

- Jambwa, M. and L. Olsson (1987). Application of database technology in the African context. Invited paper, 46<sup>th</sup> session of International Statistical Institute, Tokyo.
- Jambwa, M., C. Parirenyatwa, and B. Rosen, (1989). Data processing at the Central Statistical Office: Lessons from recent history. Central Statistics Office, Harare.
- Lagerlöf, Birgitta. (1988), Development of systems design for national household surveys. SCB R&D report, No. 4. Stockholm: Statistics Sweden.
- Lehtonen, R., and E. Pahkinen (1995). *Practical Methods for Design and Analysis of Complex Surveys*. New York: Wiley & Sons.
- Lundell, L. (1996). Information systems strategy for CSS. Report to Central Statistical Service (CSS), Pretoria.
- \_\_\_\_\_ (2003). Census data processing experiences. Report to Central Bureau of Statistics (CBS), Windhoek.
- Macro International, Inc. (1996). *Sampling Manual*, DHS-III Basic Document No. 6. Calverton, Maryland: Macro International, Inc.
- Muñoz, Juan. (2003). A Guide for data management of household surveys. *In Household Sample Surveys in Developing and Transition Countries*. Studies in Methods, No. 96. Sales No. E.05.XVII.6.
- Olofsson, P. (1985). Proposals for Survey Design, Kingdom of Lesotho, Report on short-term mission on a labour-force survey to Bureau of Statistics, Maseru.
- Olsson, Ulf. (1990)a. Approaches to agricultural statistics in developing countries: an appraisal of ICO's experiences, No. 12. Stockholm: SCB (Statistics Sweden, International Consulting Office). 20 July.
- \_\_\_\_\_ (1990)b. Applied statistics lecture notes: special reports. TAN 1990:1. Stockholm Statistics Sweden International Consulting Office.
- Pettersson, Hans. (2005). The design of master sampling frames and master samples for household surveys in developing countries. *In Household Sample Surveys in Developing and Transition Countries*. Studies in Methods, No. 96. Sales No. E.05.XVII.6.
- Puide, Annika, (1995). Report on a mission to Takwimu, Dar es Salaam, 21 November – 21 December 1994. TANSTAT 1994: 20 (20 January 1995). Stockholm: Statistics Sweden, International Consulting Office.
- Rauch, L. (2001). Best Practices in Designing Websites for Dissemination of Statistics. Conference of European Statisticians Methodological Material. Geneva: United Nations Statistical Commission and Economic Commission for Europe.
- Rosen, B. and B. Sundgren. (1991). Documentation for re-use of microdata from surveys carried out by Statistics Sweden, Working paper for Research and Development Unit, Statistics Sweden, Stockholm.
- (2002)a. Mission on sampling: framework for the master sample, Kingdom of Lesotho. Report from a mission to the Bureau of Statistics, Maseru, Lesotho, 1-15 June 2002. LESSSTAT 2002:7. Stockholm: Statistics Sweden, International Consulting Office.
- (2002)b. Report on the short-term mission on estimation procedure for master sample surveys. Maseru: Bureau of Statistics, Kingdom of Lesotho.

- Rosen, Beugt. (1991). Estimation in the income, consumption and expenditure survey, ZIMSTAT 1991: 8:1.
- Shah, B., B. Barnwell, and G. Bieler (1996). *SUDAAN User Manual: Release 7.0*, Research Triangle Park, North Carolina. Research Triangle Institute.
- Silva, P. Pedro Luis do Nascimento. (2005). Reporting and compensating for nonsampling errors for surveys in Brazil: current practice and future challenges. *In Household Sample Surveys in Developing and Transition Countries*. Studies in Methods, No. 96. Sales No. E.05.XVII.6.
- SPSS, Inc. (1988), *SPSS/PC+V2.0 Base Manual*. Chicago, Illinois: SPSS.
- St. Catherine, Edwin (2003). Review of data processing, analysis and dissemination for *Designing Household Survey Samples: Practical Guidelines*. United Nations Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys, New York, 3-5 December 2003.
- Sundgren, B. (1984). *Conceptual Design of Databases and Information Systems*, P/ADB Report E19. Stockholm: Statistics Sweden.
- \_\_\_\_\_ (1986). *User-Oriented Systems Development at Statistics Sweden*. U/ADB Report E24, Stockholm: Statistics Sweden.
- \_\_\_\_\_ (1991). *Information Systems Architecture for National and International Statistics Offices: Guidelines and Recommendations*. Geneva: United Nations Statistical Commission and Economic Commission for Europe.
- \_\_\_\_\_ (1995). *Guidelines: Modelling Data and Metadata*. Geneva: United Nations Statistical Commission and Economic Commission for Europe.
- Svensson, R. (1996). The Census Data Entry Application. Report from a mission to Central Statistical Service (CSS), Pretoria.
- Thiel, Lisa Olson. (2001). Designing and developing a web site. Report from a mission to Bureau of Statistics, Maseru, Lesotho, 12-23 November 2001. LESSTAT: 2001:17. 28 December. Stockholm: SCB Statistics Sweden, International Consultancy Office.
- United Nations (1982). National Household Survey Capability Programme: survey data processing: a review of issues and procedures. DP/UN/INT-81-041/1. New York: United Nations Department of Technical Co-operation for Development Statistical Office.
- \_\_\_\_\_ (1985). National Household Survey Capability Programme: household income expenditure surveys: a technical study. DP/UN/INT.88-X01/6E. New York: United Nations Department of Technical Co-operation for Development and Statistical Office.
- Verma, Vijay. (1982). The estimation and presentation of sampling errors, World Fertility Survey, Technical Bulletins No. 11 (December). The Hague: International Statistical Institute. Voorburg, Netherlands.
- Wallgren, Anders, and others (1996). *Graphing Statistics and Data: Creating Better Charts*. Thousand Oaks, California: Sage Publications, Inc.
- World Bank (1991). The SDA survey instrument: an instrument to capture social dimensions of adjustment. Washington, D.C. Poverty and Social Policy Division, Technical Department, Africa Division.
- Yansaneh, I. (2005). Overview of sample design issues for household surveys in developing and transition countries. United Nations Statistics Division, New York. *Household Sample Surveys in Developing and Transition Countries*. Studies in Methods, No. 96. Sales No. E.05.XVII.6.

# المرفق الأول

## أساسيات تصميم عيّنات الاستقصاء

### ميم - ١ مقدمة

١ - المعاينة (أخذ العيّنات) تقنية يُختار بها جزء من السكان وتُعمّم النتائج المتولدة من هذا الجزء على جميع السكان الذين اختير منهم هذا الجزء أو العيّنة. يوجد بوجه عام نوعان من العيّنات هما عيّنات احتمالية وعيّنات غير احتمالية. وينصبُّ تركيزنا في هذا الدليل على العيّنات الاحتمالية. وستغطي النظرة العامة وحدات الاستقصاء، وتصميم العيّنات، والاستراتيجيات الأساسية للمعاينة، مع الأمثلة.

### ميم - ٢ وحدات الاستقصاء ومفاهيمه

٢ - نبدأ بتعريف وحدات الاستقصاء ومفاهيمه المستخدمة بصورة مشتركة في أخذ عيّنات للاستقصاء. العناصر: عناصر (وحدات) السكان هي وحدات يُلتصق الحصول على معلومات بشأنها. ربما تكون هي الوحدات الابتدائية التي يتألف منها السكان، الذين يُراد التوصل إلى استدلالات بشأنهم. ففي استقصاء خصوبة، مثلاً، للأسر المعيشية، تكون النساء اللاتي هن في مرحلة الحمل والولادة هن العناصر النسائية في نهاية المطاف. ولتيسير جمع البيانات في الاستقصاء من الأساسي جداً تعريف العناصر تعريفاً جيداً، ومن السهل مادياً تحديد هوياتها.

٣ - السكان: السكان هم مجموع السكان المعرفين أعلاه. فالعناصر إذن هي الوحدات الأساسية التي يتكون منها السكان وهي التي تعرّف السكان. ومن الأساسي تعريف السكان على النحو التالي:

- المضمون، الذي يدعو إلى تعريف نوع وخصائص العناصر التي يتكوّن منها السكان
- المدى، الذي يشير إلى الحدود الجغرافية بقدر ما تتصل بالتغطية
- الوقت، الذي يشير إلى الفترة الزمنية التي يكون فيها السكان موجودين.

٤ - وحدات الرصد: هذه وحدات يُحصَل منها على الملاحظات. وفي الاستقصاءات القائمة على المقابلات تُسمّى بالمستجيبين. وحدات الإبلاغ هي العناصر التي تقدم المعلومات المطلوبة في الاستقصاء. لاحظ أنه في بعض الحالات ربما تكون وحدات الرصد ووحدات الإبلاغ مختلفة بعضها عن بعض. ففي استقصاء للأطفال الذين هم دون سن الخامسة، الآباء والأمهات هم الذين يعطون في العادة المعلومات المتعلقة بأطفالهم نيابة عن هؤلاء الأطفال. في هذه الحالات يكون الأطفال المختارون هم وحدات الرصد بينما الآباء والأمهات هم وحدات الإبلاغ.

٥ - وحدات المعاينة: تُستخدم وحدات المعاينة لاختيار عناصر لإدخالها في العيّنة. وفي معاينة العناصر تحتوي وحدة المعاينة على عنصر واحد، بينما في معاينة المجموعات، مثلاً، تتألف وحدة المعاينة من مجموعة من العناصر تسمى "المجموعة". فمنطقة تعداد السكان، مثلاً، تحتوي، بوصفها وحدة معاينة في المرحلة الأولى، على مجموعة من الأسر المعيشية. ويمكن لنفس الاستقصاء أن يستخدم وحدات معاينة مختلفة. ومن الأمثلة الجيدة على ذلك المعاينة المتعددة المراحل، التي تستخدم سلسلة من وحدات المعاينة في ترتيب هرمي (انظر الفصل الثالث).

٦ - وحدات العيّنة: يمكن تسمية وحدات المعاينة المختارة "وحدات العيّنة"، وتُعرف قيم الخصائص التي هي قيد الدراسة لوحدة العيّنة باسم "ملاحظات العيّنة": وحدة التحليل: هذه وحدة تستخدم في مرحلة الجدولة والتحليل. وربما تكون هذه الوحدة وحدة ابتدائية أو مجموعات وحدات ابتدائية. ومن الجدير بالملاحظة، كما ذُكر من قبل، أنه لا داعي بالضرورة لأن تكون وحدة التحليل ووحدة الإبلاغ نفس الوحدة.

٧ - إطار المعاينة: يستخدم إطار المعاينة لتعيين وحدات معاينة واختيارها للدخول في العيّنة، ويستخدم أيضاً كأساس لجعل التقديرات تستند إلى بيانات العيّنة. ويعني هذا ضمناً أن السكان الذين ستختار منهم العيّنة يجب أن يكونوا ممثلين بشكل طبيعي. وينبغي في الأحوال المثالية أن يحتوي الإطار على جميع وحدات المعاينة التي تنتمي إلى المجموعة السكانية التي هي قيد الدراسة مع تعيين خصائصها على الوجه الصحيح. ويجب أن تكون الأطر جامعة مانعة، ويفضّل أن يُجَبَّ بعضها بعضاً (للاطلاع على مزيد من التفاصيل راجع إلى الفصل الرابع). الأطر المستخدمة بصورة شائعة في الاستقصاءات أطر قائمة، وأطر منطقة وأطر متعددة.

٨ - إطار القائمة: يحتوي إطار القائمة على قائمة وحدات معاينة يمكن أن تُختار عيّنة منها مباشرة. ويفضل أن يكون الإطار على صلة بكل وحدة معاينة وفيه معلومات دقيقة عن هذه الوحدات، كالحجم والخصائص الأخرى. وتساعد المعلومات الإضافية على تصميم و/أو اختيار عيّنات ذات كفاءة.

٩ - أطر المنطقة: أطر المنطقة أطر متعددة المراحل وهي تُستخدم استخداماً شائعاً بوجه عام في استقصاءات الأسر المعيشية. وفي هذا الصدد يتألف الإطار من مرحلة أو أكثر من مراحل وحدات المنطقة. ففي تصميم عيّنة ذات مرحلتين، مثلاً، سيتألف الإطار من مجموعات يمكن أن تُسمّى وحدات معاينة أولية؛ وفي وحدات المعاينة الأولية المختارة تصبح قائمة الأسر المعيشية إطار المرحلة الثانية. والأطر تلزم بوجه عام في كل مرحلة من مراحل الاختيار. وتنقص مدة بقاء الإطار كلما انخفضت رتبة الوحدة في التسلسل الهرمي.

١٠ - وحدات المنطقة: تغطي وحدات المنطقة مناطق برية محددة لها حدود معروفة بوضوح، يمكن أن تكون ملامح تضاريس طبيعية كالطرق، أو الشوارع، أو الأنهار، أو خطوط السكة الحديدية، أو خطوطاً وهمية تمثل الحدود الرسمية بين الأقسام الإدارية. وتُنشأ مناطق تعداد السكان في العادة داخل الوحدات الإدارية الأصغر مساحة التي توجد في البلد. وهذا يسهّل الإبلاغ عن الأعداد المتعلقة بالوحدات الإدارية باعتبارها مجالات.

١١ - الإطار أو الأطر المستخدمة لاستقصاء الأسر المعيشية يجب أن تكون قادرة على توفير إمكانية الوصول إلى جميع وحدات المعاينة في المجموعة السكانية المشمولة بالاستقصاء، لكي يكون لكل وحدة احتمال معروف للاختيار لدخول العيّنة وليس صفرًا. يمكن تحقيق الوصول بالاختيار للعيّنة من الأطر ويكون ذلك عادة في مرحلتين أو أكثر من مراحل الاختيار. ويجب أن يشمل إطار المرحلة الأولى من مراحل المعاينة جميع وحدات المعاينة المسماة. وفي المراحل اللاحقة من مراحل الاختيار لا تلزم الأطر إلا لوحدة العيّنة التي اختيرت في المرحلة السابقة. ويمكن تخزين إطار المعاينة في قرص صلب و/أو واسطة إلكترونية.

### ميم - ٣ - تصميم العيّنة

١٢ - يشير تصميم العيّنة بوجه عام إلى اختيار العيّنات وتقديرها. من ثمّ يُعنى الموضوع بكيفية اختيار جزء من السكان لشموله بالاستقصاء. وينطوي تصميم العيّنات، من زاوية عملية، على تحديد حجم العيّنة والهيكل ويأخذ في الحسبان تكاليف الاستقصاء. وأفضل تصميم للعيّنات هو التصميم الذي يسفر عن أعلى مستويات الدقة بتكلفة معيّنة للاستقصاء، أو عن أقل تكلفة ممكنة بمستوى محدد من الدقة.

١٣ - غير أنه ينبغي منذ البداية تأكيد أنه لا يمكن فصل تصميم العيّنات عن الجوانب الأخرى لتصميم الاستقصاء وتنفيذه. تُعنى نظرية المعاينة، بوجه عام، بكيفية اتصال تقديرات مجموعة معيّنة من السكان، المستمدة من أخطاء الاستقصاء وأخطاء المعاينة المرتبطة بها، بحجم العيّنة وهيكلها.

### ميم - ٣ - ١ المتطلبات الأساسية لتصميم عيّنة احتمالية

- يجب تعريف السكان المستهدفين بوضوح
- يجب أن يكون ثمة إطار أو أطر في حالة العيّنات متعددة المراحل
- يجب تحديد أهداف الاستقصاء بوضوح لا غموض فيه من حيث مضمون الاستقصاء، والمتغيرات التحليلية ومستويات التقسيم (هل تريد، مثلاً، تقديرات أو بيانات على المستوى الوطني، أو الريفي / الحضري، أو الإقليمي، أو على مستوى المقاطعة؟)
- ينبغي أن تؤخذ في الحسبان قيود الميزانية والقيود الميدانية
- يجب إعلان متطلبات الدقة بكل وضوح بغية تحديد حجم العيّنة

### ميم - ٣ - ٢ أهمية المعاينة الاحتمالية في الاستقصاءات الكبيرة للأسر المعيشية

- تسمح بتغطية جميع السكان المستهدفين في اختيار العيّنات
- تخفض تحييز المعاينة
- تسمح بالتعميم في الحديث عن نتائج العيّنة للسكان الذين اختيرت منهم العيّنة
- تسمح بحساب أخطاء المعاينة، التي هي مقاييس لاحتمالية
- قيل إنها تسمح لمنقذ الاستقصاء أن يقدم نتائج دون الاضطرار إلى الاعتذار عن استخدام أساليب غير علمية

### ميم - ٣ - ٣ إجراءات الاختيار والتنفيذ والتقدير

- ينبغي تمثيل كل عنصر من عناصر السكان في الإطار الذي ستُختار منه العيّنة
- ينبغي أن يستند اختيار العيّنة إلى عملية عشوائية تعطي كل وحدة احتمال اختيارٍ محدداً
- ينبغي تعداد سكان جميع الوحدات المختارة، والمختارة فقط
- يجب، في تقدير مَعْلَمَات السكان، ترجيح كل وحدة/ عنصر وفقاً لاحتمال الاختيار

١٤ - يخفض الاختيار العشوائي للوحدات فرصة الحصول على عيّنة غير تمثيلية. فالاختيار العشوائي طريقة آمنة للتغلب على آثار عوامل التحيز غير المنظورة. ويتوقف الأسلوب المستخدم لاختيار العيّنات على خطة المعاينة المستخدمة. فكلما ازداد تصميم العيّنة تعقيداً أصبحت إجراءات الاختيار أصعب.

## ميم - ٤ أساسيات استراتيجيات المعاينة الاحتمالية

١٥ - يوجد عدد من استراتيجيات المعاينة الاحتمالية وُضعت بقصد إنتاج تصميم للعيّنات، من بينها العشوائية البسيطة، والمنهجية، ومعاينة الطبقات والمجموعات. ستناقش هذه التقنيات فيما يلي أدناه مع بعض الأمثلة.

### ميم - ٤ - ١ المعاينة العشوائية البسيطة

١٦ - المعاينة العشوائية البسيطة (SRS) أسلوب لاختيار عيّنة بناءً على الاحتمال، حيث تتاح لكل عنصر من عناصر السكان فرصة/احتمال للاختيار على قدم المساواة. ويمكن أن يكون اختيار العيّنة مع البديل أو بلا بديل. هذا الأسلوب نادراً ما يستخدم في استقصاءات الأسر المعيشية الواسعة النطاق لأنها باهظة التكلفة من حيث إعداد القوائم والسفر. يمكن اعتبار هذا الأسلوب القلب الأساسي للمعاينة الاحتمالية التي تنطبق على أوضاع لا توجد فيها معلومات سابقة متاحة عن هيكل السكان. المعاينة العشوائية البسيطة جذابة بفضل بساطتها من حيث إجراءات الاختيار والتقدير (أخطاء المعاينة، مثلاً).

١٧ - وإن كانت المعاينة العشوائية البسيطة لا تُستخدَم كثيراً، فهي أساسية لنظرية المعاينة، ويعود معظم السبب في ذلك إلى بساطة خصائصها الرياضية. لذلك، تفترض معظم النظريات والتقنيات الإحصائية الاختيار العشوائي البسيط للعناصر. بل إنه يمكن أن يُرى في جميع الاحتمالات الأخرى لاختيار العيّنات الاحتمالي قيود على المعاينة العشوائية البسيطة تمنع اختيار بعض توليفات العناصر السكانية. تقوم المعاينة العشوائية البسيطة بوظيفتين اثنتين:

- تضع خط أساس لمقارنة الكفاءة النسبية لتقنيات المعاينة الأخرى
- يمكن استخدامها كأسلوب أخير لاختيار الوحدات الابتدائية في سياق التصاميم الأكثر تعقيداً، كتصاميم معاينة المجموعات والطبقات

توضح الأمثلة الواردة أدناه حساب احتمال الاختيار بموجب المعاينة العشوائية البسيطة:

١ - أولاً ننظر في مجموعة محدودة من السكان قوامها ١٠٠ أسرة معيشية ...  $H_1, H_2, \dots, H_i, \dots, H_{100}$  وقيم دخلها هي  $X_1, X_2, \dots, X_i, \dots, X_{100}$ .

يكون احتمال اختيار أي وحدة بعينها في هذا المثال  $\frac{1}{100}$

٢ - ونلاحظ، كمثال ثانٍ، أنه يمكن للحصول على عيّنة من الأسر المعيشية إعطاء الأسر المعيشية أرقاماً متسلسلة في إطار/قائمة. باستخدام أرقام عشوائية، يمكن اختيار عيّنة ليكُن حجمها، مثلاً، ٢٥. ولأن أسلوب الاختيار بتساوي الاحتمالات (EPSEM) هو كسر المعاينة الإجمالي للعناصر.

$$f = \frac{n}{N} \text{ لذلك فإن}$$

وإذا كان  $n = 25$  هو حجم العيّنة، و  $N = 100$  هو مجموع عدد الأسر المعيشية، فإن كسر العيّنة، الذي يساوي احتمال الاختيار هو

$$\frac{25}{100} = \frac{1}{4}$$

### ميم - ٤ - ١ - ١ أنواع اختيار العيّنات بموجب المعاينة العشوائية البسيطة

١٨ - توجد طريقتان شائعتان بموجب المعاينة العشوائية البسيطة، وهما:

(أ) المعاينة العشوائية البسيطة مع البديل (SRSWR):

(ب) المعاينة العشوائية البسيطة بلا بديل (SRWOR).

#### المعاينة العشوائية البسيطة مع البديل

١٩ - المعاينة العشوائية البسيطة مع البديل تستند إلى اختيار عشوائي من مجموعة سكانية يُجرى بتبديل العنصر المختار بعد كل سحب. ويظل احتمال اختيار عنصر ما دون تغيير بعد كل سحب، وإن كل العيّنات المستقلة التي تُختار تبقى مستقلة بعضها عن بعض. هذه الخاصية تفسر سبب استخدام العيّنة العشوائية البسيطة باعتبارها تقنية المعاينة التلقائية في كثير من الدراسات الإحصائية النظرية. يُضاف إلى ذلك أنه نظراً إلى أن افتراض المعاينة العشوائية البسيطة يسهّل معادلات التقديرات كثيراً، مثل تقديرات التباين، فهي تستخدم بمثابة مرجع. نُعطي في الفقرة ٢٠ أدناه معادلات لتقدير المتوسط (ميم - ١) والتباين (ميم - ٢) لمتوسط العيّنة، بموجب المعاينة العشوائية البسيطة مع البديل. وتوضّح المعادلات بأمثلة رقمية.

٢٠ - فإذا أخذنا عيّنة فيها العدد  $n$  من الوحدات، التي اختيرت باستخدام المعاينة العشوائية البسيطة مع البديل، وجمعت لها مَعْلَمَات عن المتغير  $x$ ، يعطى المتوسط والتباين بما يلي:

#### ١ - المتوسط

(ميم - ١)

$$\bar{x} = \frac{1}{n} \sum_i^n x_i = \frac{1}{n} [x_1 + x_2 + \dots + x_n]$$

عندما:  $x_1 = 24, x_2 = 30, x_3 = 27, x_4 = 36, x_5 = 31, x_6 = 38, x_7 = 23, x_8 = 40, x_9 = 25, x_{10} = 32$

$$\bar{x} = \frac{24 + 30 + 27 + \dots + 25 + 32}{10} = 30.6 \quad \text{ثم:}$$

#### ٢ - التباين

(ميم - ٢)

$$V(\bar{x}) = \frac{s^2}{n}$$

(ميم - ٣)

$$s^2 = \frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left[ \sum_i^n x_i^2 - \frac{x^2}{n} \right] = \frac{1}{n-1} \left( \sum x_i^2 - n\bar{x}^2 \right) \quad \text{حيث:}$$

$$x^2 = (\sum x_i)^2 = 93,636$$

عندما نحسب بهذه القيم،

$$s^2 = \frac{(9,684 - 9,364)}{9} = 35.56$$

$$V(\bar{x}) = \frac{35.56}{10} = 3.56$$

$$Se(\bar{x}) = \sqrt{3.56}$$

المعاينة العشوائية البسيطة بلا بديل

٢١ - من الأفضل، بحكم السليقة، أخذ العيّنات بلا بديل، لأن المرء يحصل على معلومات أكثر، لأنه لا توجد إمكانية لتكرار وحدات العيّنة. لذلك، فإن استراتيجية المعاينة العشوائية البسيطة بلا بديل أكثر ما تستخدم إجراء المعاينة العشوائية البسيطة. عملية الاختيار، في هذا الإجراء، متواصلة حتى تُختار وحدات معاينة عددها  $n$  ويتم تجاهل جميع حالات التكرار. هذا يعادل الاحتفاظ بالوحدة أو الوحدات التي اختيرت، وتُختار وحدة أخرى مع تساوي الاحتمالات من بين الوحدات الباقية في المجموعة السكانية.

فيما يلي بعض خواصّ المعاينة العشوائية البسيطة بلا بديل:

- تعطينا حجماً ثابتاً للعيّنة
- تسفر عن الاختيار مع تساوي الاحتمالات لكل عنصر/وحدة (EPSEM)
- متوسط العيّنة والتباين، كما هو الحال في المعاينة العشوائية البسيطة مع البديل، تقديرات غير متحيزة لمعلّمات السكان.

٢٢ - في الفقرة ٢٣ أدناه، نقدم معادلات تُستخدم في تقدير المتوسط والتباين بموجب المعاينة العشوائية البسيطة بلا بديل (ميم - ٤ وميم - ٥). نعطي، بالإضافة إلى ذلك، أمثلة رقمية لكيفية حساب المتوسط والتباين في العيّنة.

٢٣ - افترض أن مجموع عدد المدارس الابتدائية في إقليم ما يساوي ٢٧٥ مدرسة. اختيرت عيّنة قوامها ٥٥ بلا بديل. الأرقام الواردة أدناه هي عدد الموظفين ( $y_i$ ) في كل مدرسة من المدارس المختارة.

٢	٨	٦	٣٢	١٠	٥
٦	٥٠	٧	٣٥	١٦	١٥
٦	٢٠	٢٠	٤٧	٦	٢
٢	١٦	٦	٣٥	٦	٧
٢	١٥	٤	٤٨	٢	٢١
	٧	٦	٤٦	٥	٧
	٦	٢	٨	٤	٤
	٢	٨	٧	٢	٧
	٢	٦	١٠	١٢	٥
	١٩	٧	٧	٤٠	٢

$$\sum y_i = 688, \text{ مجموع عدد الموظفين}$$

$$\sum y_i^2 = 18,182$$

١ - متوسط العينة هو

(ميم - ٤)

$$\bar{y} = \frac{\sum y_i}{n}$$

حيث  $n$  هو حجم العينة.

عندما نحسب بهذه القيم،

$$\bar{y} = \frac{688}{55} = 12.5$$

٢ - تباين متوسط العينة هو

(ميم - ٥)

$$V(\bar{y}) = 1 - f \frac{s_y^2}{n}$$

حيث  $1-f$  هو عامل تصحيح المجموعة السكانية، و

(ميم - ٦)

$$s_y^2 = \frac{1}{n-1} [\sum y_i^2 - n\bar{y}^2] = \frac{1}{54} [18,182 - 8,594] = 177.56$$

عندئذٍ

$$Se(\bar{y}) = \sqrt{2.58} \text{ و } V(\bar{y}) = \left(1 - \frac{55}{275}\right) 177.56/55 = 2.58$$

## ميم - ٤ - ٢ المعاينة المنهجية

٢٤ - المعاينة المنهجية طريقة اختيار للعينة بالاحتمال، تُختار فيه العينة باختيار كل عنصر رقمه  $k^{\text{th}}$  من السكان، حيث تساوي  $k$  عدداً صحيحاً أكثر من ١. ويجب أن يُختار العدد الأول من العينة عشوائياً من بين أول عناصر تحمل الرقم  $k$ . ويجري الاختيار من قائمة منظمة. هذا أسلوب مرغوب للاختيار، لا سيما عندما تكون الوحدات عديدة وتكون مرقمة من ١ إلى  $N$ . افترض أن  $N$ ، الذي هو مجموع عدد الوحدات، عدد صحيح وهو من مضاعفات حجم العينة المطلوب  $n$  وأن  $k$  عدد صحيح بحيث يكون  $N$  مساوياً لـ  $(N = nk)$ . ثم يُختار عدد عشوائي بين ١ و  $k$ . لنفترض أن ٢ هو البداية العشوائية، عندئذٍ يكون حجم العينة  $n$  وتكون الوحدات مرقمة ترقيمياً تسلسلياً كالتالي:

$$2, 2 + k, 2 + 2k, \dots \dots 2 + (n-1)K$$

ويلاحظ أن العينة تشكل أول وحدة تُختار عشوائياً، وتُختار كل وحدة رقمها  $k^{\text{th}}$ ، حتى تبلغ العينة المختارة الحجم المطلوب. الفاصل  $k$  يقسم السكان إلى مجموعات. ونحن، بهذا الإجراء، نختار مجموعة واحدة من الوحدات يكون

احتمال اختيارها  $1/k$ . وحيث إن الرقم الأول سُحب عشوائياً من مجموعة أرقام تتراوح من ١ إلى  $k$ ، فإن كل وحدة في المجموعات التي يفترض أنها متساوية يكون لها نفس احتمال الاختيار،  $1/k$ .

### ميم - ٤ - ٢ - ١ المعاينة المنهجية الخطية

٢٥ - إذا كان  $N$ ، وهو مجموع عدد الوحدات، من مضاعفات حجم العينة المطلوب، بعبارة أخرى إذا كان  $N = nk$ ، وكان  $n$  هو حجم العينة المطلوب و  $k$  هو الفاصل — فإن عدد الوحدات في كل عينة من العينات المنهجية الممكنة هو  $n$ . في هذه الحالة يكون النظام بمثابة تقسيم الوحدات ذات الرقم  $N$  إلى فئة عينات رقمها  $k$  وكل عينة منها مؤلفة من وحدات عددها  $n$ ، وتُختار مجموعة يكون احتمال اختيارها  $1/k$ . وعندما يكون  $N$  مساوياً لـ  $nk$ ، يكون  $\bar{y}$  التقدير غير المتحيز لمتوسط عدد السكان  $\bar{Y}$ . ومن جهة أخرى، عندما لا يكون  $N$  من مضاعفات  $n$ ، فإن عدد الوحدات المختارة باستخدام التقنية المنهجية، مع كون فاصل المعاينة  $k$  مساوياً للعدد الصحيح الذي هو أقرب ما يكون إلى  $N/n$  لا يكون بالضرورة مساوياً لـ  $n$ . لذلك، عندما لا يكون  $N$  مساوياً لـ  $nk$ ، تختلف أحجام العينات ويكون متوسط حجم العينة تقديراً متحيزاً لمتوسط عدد السكان. يوضح الشكل ميم - ١ أدناه اختيار العينة بموجب المعاينة المنهجية الخطية.

يوضح المثال الوارد أدناه اختيار عينة قوامها ٤ طلاب من صف فيه ٢٠ طالباً. البداية العشوائية هي ٣،  $N = 20$ ، و  $n = 4$ ، و  $k = 5$ . تضم العينة الناتجة وحدات تحمل الأرقام ٣، ٨، و ١٣، و ١٨.

الشكل ميم - ١

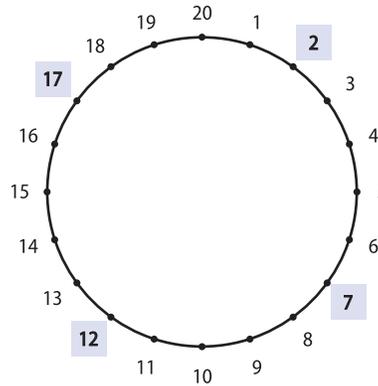
### المعاينة المنهجية الخطية (اختيار العينات)

١٠	٩	٨	٧	٦	٥	٤	٣	٢	١
٢٠	١٩	١٨	١٧	١٦	١٥	١٤	١٣	١٢	١١

### ميم - ٤ - ٢ - ٢ المعاينة المنهجية الدائرية

٢٦ - لاحظنا أنه في المعاينة المنهجية الخطية يمكن أن يكون حجم العينة الفعلي مختلفاً عن الحجم المطلوب، وأن متوسط الحجم تقديراً متحيزاً لمتوسط السكان عندما لا يكون  $N$  من مضاعفات  $n$ . غير أن تقنية المعاينة المنهجية الدائرية تتغلب على القصور المذكور أعلاه. ففي الاختيار المنهجي الدائري تُرتب القوائم في دائرة بحيث تلي الوحدة الأولى الوحدة الأخيرة. تُختار بداية عشوائية بين الرقم ١ والرقم  $N$ ، بدلاً من أن تكون بين الرقم ١ والرقم  $k$ . ثم تضاف الوحدة ذات الرقم  $k^{\text{th}}$  حتى تُختار العناصر التي عددها  $n$  بالضبط. وعندما نصل إلى نهاية القائمة نواصل من البداية. ويعطي الشكل ميم - ٢ توضيحاً لاختيار عينة بموجب المعاينة المنهجية الدائرية، حيث  $N = 20$ ، و  $n = 4$ ، و  $k = 5$ ، وتكون البداية العشوائية ٧. لذلك تكون الوحدات المختارة ٧ و ١٢ و ١٧ و ٢٠.

الشكل ميم - ٢  
اختيار العيّنات بالعيّنة المنهجية الدائرية



ميم - ٤ - ٢ - ٣ التقدير بموجب العيّنة المنهجية

٢٧ - تعطى المعادلات تقدير المجموع (ميم - ٧)، ومتوسط العيّنة (ميم - ٨)، والتباين (ميم - ٩)، وتقدّم أمثلة رقمية تبيّن العدد التقديري للسكان، ومتوسط العيّنة، والتباين.

١ - لتقدير المجموع يُضرب مجموع العيّنات بفاصل العيّنة، لذلك

(ميم - ٧)

$$\hat{Y} = k \sum y_i$$

ويكون تقدير متوسط السكان

(ميم - ٨)

$$\bar{y} = k \frac{\sum y_i}{N}$$

٢ - تقدير التباين معقد لأنه لا يمكن إعطاء تقدير دقيق جداً من عيّنة منهجية واحدة. والمخرَج هو افتراض أن ترقيم الوحدات عشوائي؛ وفي هذه الحالة يمكن معاملة العيّنة المنهجية كأنها عيّنة عشوائية. لذلك يُعطى تقدير التباين للمتوسط بالمعادلة التالية:

(ميم - ٩)

$$V(\bar{y}) = \frac{1}{n} \left(1 - \frac{n}{N}\right) \sum s^2$$

حيث:  $y = \sum y_i$  و  $s^2 = \frac{1}{n-1} \sum (y_i - y)^2$

٢٨ - يمكن حساب تقدير دقيق جداً لتباين غير متحيّز من عيّنة منهجية باختيار أكثر من عيّنة منهجية واحدة من مجموعة سكانية معيّنة.

الأمثلة الرقمية

٢٩ - افترض أنه يوجد ١٨٠ مزرعة تجارية في منطقة تضم الواحدة منها ٣٠ رأساً من الماشية أو أكثر. تؤخذ عيّنة فيها ٣٠ مزرعة باستخدام العيّنة المنهجية ويكون  $k$  مساوياً لـ ٦ ( $k = 6$ ).

يرد أدناه عدد رؤوس المواشي ( $y_i$ ) في الـ ٣٠ مزرعة المختارة.

	١٢٠	٣٠	٤١	٣٥	٧٩	٤٠	٥٠	٤٥	٢٠٠	٦٠
$\sum y_i = 2,542$ و	٤٠	٣٢	٦٧	٥١	٤٢	٢٠٠	١٢٠	١١١	٦٥	٣٠٠
	٥٠	٣١	٨٢	٤٧	٩٠	٦٣	١٠٠	٢٥٠	٥٥	٤٦

١ - العدد المقدر للمواشي هو

$$\hat{Y} = k \sum y_i = 6 \times 2,542 = 15,252$$

٢ - المتوسط المقدر لعدد المواشي في كل مزرعة هو

$$\bar{y} = k \frac{(\sum y_i)}{N} = 6 \times 2,542 / 180 = 84.7 \approx 85$$

٣ - تباين متوسط العيّنة، المحسوب على أساس افتراض أن ترقيم المزارع عشوائي، هو

$$V(\bar{y}) = 1 - f \frac{s_y^2}{n}$$

$$\text{حيث: } s_y^2 = \frac{1}{n-1} \left\{ \sum y_i^2 - \frac{(\sum y_i)^2}{n} \right\} = \frac{1}{29} (348,700 - 215,392.13) = 4,596.80 \quad (\text{ميم - ١٠})$$

$$\text{لذلك: } V(\bar{y}) = (0.833)(153.227) = 127.64$$

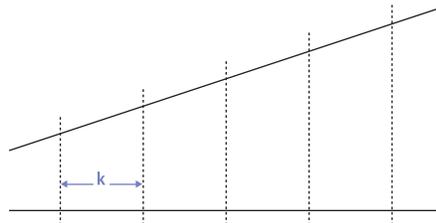
$$\text{و } Se(\bar{y}) = \sqrt{127.64} = 11.30$$

٣٠ - يوجد عدد من الميزات والعيوب المرتبطة باستخدام المعاينة المنهجية.

(أ) الميزات:

- اختيار الوحدة الأولى يحدد العيّنة بأكملها. ويعمل هذا في صالح العمليات الميدانية لأنه يمكن اختيار وحدات المعاينة النهائية في الميدان من قبل العدّادين عندما يضعون قوائم بالوحدات.
- تُنشر العيّنة بالتساوي بين السكان عندما تكون الوحدات المشمولة بالإطار مرقمة ترقيماً ملائماً. غير أن تقدير العيّنة سيكون أكثر دقة إذا كان ثمة ميلٌ من نوع ما في السكان.
- المعاينة المنهجية تعطي تقسيماً طبقياً ضمناً. يوضّح الشكل ميم - ٣، الوارد أدناه، التقسيم الطبقي الضمني بواسطة اتجاه خطي رتيب.

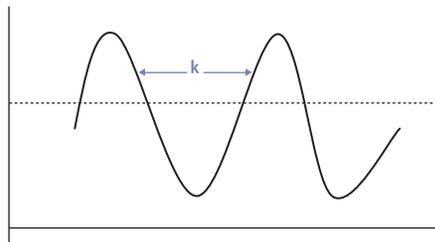
الشكل ميم - ٣  
الاتجاه الخطي الرتيب



(ب) أوجه القصور:

- إذا كان ثمة تغير دوري في السكان، يمكن أن تسفر المعاينة المنهجية عن نتائج تكون تقديراتها أقل من الواقع أو أكثر منه. وفي هذه الحالة يكون فاصل المعاينة متنسقاً مع البيانات. فإذا كنت مثلاً تدرس تدفق حركة السير لمدة ٢٤ ساعة في شارع يغصّ بحركة السير من شوارع المدينة، وحل فاصلك في ساعات ازدحام حركة السير، ستحصل دائماً على أرقام مرتفعة. لذلك، ستسفر الدراسة عن تشغيل العاملين ساعات إضافية. يوضح الشكل ميم - ٤ الفاصل الدوري الذي يمكن أن يساهم في التقديرات غير الموثوقة بموجب المعاينة المنهجية
- إذا تحرّينا الدقة، لا نستطيع أن نحصل على تقدير دقيق جداً من عيّنة منهجية واحدة
- أسلوب الاختيار عرضة للإساءة من قبل بعض العدّادين/الموظفين الميدانيين

الشكل ميم - ٤  
التقلبات الدورية



### ميم - ٤ - ٣ المعاينة الطباقية

٣١ - في أسلوب المعاينة الطباقية تقسم وحدات المعاينة في السكان إلى مجموعات تُسمّى طبقات. ويتم التقسيم إلى طبقات لكي يتسنى تقسيم السكان تقسيماً فرعياً إلى مجموعات غير متجانسة تكون متجانسة داخلياً. وبوجه عام، عندما تكون وحدات المعاينة متجانسة فيما يتعلق بالمتغيرات المساعدة، المسماة متغيرات طباقية، ينخفض تباين تقديرات الطبقات في العادة. ومن الجدير بالملاحظة أيضاً أنه توجد مرونة كبيرة في التقسيم إلى طبقات، بمعنى أن إجراءات المعاينة والتقدير يمكن أن تكون مختلفة من طبقة إلى أخرى.

٣٢ - لذلك، في المعاينة الطباقية، نجمّع الوحدات/العناصر المتشابهة إلى حد ما، لكي يكون التباين  $\delta_h^2$  داخل كل طبقة صغيراً. ومن الأساسي في الوقت نفسه أن تكون متوسطات ( $\bar{x}_h$ ) الطبقات المختلفة مختلفة قدر الإمكان. ويُحصل على تقدير ملائم للسكان ككل بالجمع بصورة مناسبة بين التقديرات الطباقية للخاصية التي هي قيد النظر.

### ميم - ٤ - ٣ - ١ ميزات المعاينة الطباقية

٣٣ - الميزة الرئيسية للمعاينة الطباقية هي الزيادة الممكنة في دقة التقديرات وإمكانية استخدام إجراءات المعاينة المختلفة في الطبقات المختلفة. بالإضافة إلى ذلك، وُجد أن التقسيم إلى طبقات مفيد:

- في حالات المجموعات السكانية المتخالفة يمكن استخدام كسور كبيرة من كسور المعاينة في اختيار أمثلة كثيرة من وحدات قليلة أكبر حجماً. وهذا يعطي وزناً أكبر للوحدات الكبيرة جداً، وفي النهاية ينخفض التباين في المعاينة الطباقية

- عندما يكون لمنظمة استقصاءات عدة مكاتب ميدانية في مناطق مختلفة، قُسم إليها البلد لأغراض إدارية، في هذه الحالة يمكن أن يكون من المفيد معالجة المناطق كطبقات لتسهيل العمل الميداني
- عندما تلزم تقديرات في حدود هامش خطأ محدد، لا للسكان ككل فقط، وإنما لمجموعات فرعية محددة أيضاً، كالأقاليم، والمناطق الريفية أو الحضرية، ونوع الجنس، إلخ. يمكن، بواسطة التقسيم إلى طبقات، تقديم هذه التقديرات بصورة مؤاتية
- إذا كان إطار المعاينة متاحاً في شكل أُطر فرعية، تخص المناطق أو فئات وحدات محددة، وفي هذه الحالة يكون من المؤاتي عملياً والاقتصادي معاملة الأُطر الفرعية كطبقات لاختيار العيّنات

### ميم - ٤ - ٣ - ٢ موجز الخطوات المتخذة في المعاينة الطباقية

- تقسيم سكان وحدات المعاينة ككل إلى مجموعات سكانية فرعية متجانسة داخلياً وغير متجانسة خارجياً
- في كل طبقة تُختار عيّنة منفصلة من جميع وحدات المعاينة في الطبقة
- في كل عيّنة تُختار من كل طبقة، يُحسب متوسط طبقي مستقل (أو أية إحصائية أخرى). ثم توضع معاملات ترجيح، مثلاً، على الوجه الصحيح لمتوسطات الطبقات لتشكّل تقديراً مجتمعاً لمتوسط السكان
- تُستخدم في العادة معاينة متناسبة داخل الطبقات عندما تكون التقديرات الإجمالية — الوطنية، مثلاً، هي هدف الاستقصاء ويكون الاستقصاء متعدد الأغراض
- تُستخدم المعاينة غير المتناسبة عندما تكون الأولوية لمجالات المجموعات الفرعية، مثلاً: في الحالات التي يلزم فيها وضع تقديرات للمناطق دون الوطنية مع التساوي في الموثوقية

### ميم - ٤ - ٣ - ٣ ملاحظات

- ٣٤ - الرموز والحروف السفلية ترتبط بالمعاينة الطباقية. لذلك نبدأ بتعريف بعض الملاحظات والرموز الشائعة المستخدمة في استراتيجية المعاينة هذه.

القيم السكانية

للطبقات  $H$ ، سيشار إلى مجموع عدد العناصر في كل طبقة بالرموز التالية:

$$.N_1, N_2, \dots \dots N_b, \dots \dots N_H$$

وتكون هذه المعلومات غير معروفة في العادة. ويكون مجموع القيمة السكانية:

(ميم - ١١)

$$\sum_b^H N_b = N$$

ويكون متوسط الطبقة كما يلي:

(ميم - ١٢)

$$\bar{X}_{hi} = \frac{1}{N} \sum_i^{N_b} X_{hi} = \frac{X_b}{N}$$

حيث  $X_{hi}$  هي قيمة العنصر  $h^{th}$  في  $h^{th}$  الطبقة، و  $X_b$  هو مجموع قيم الطبقة  $h^{th}$ .

ميم - ٤ - ٣ - ٤ مُعَامِلَات التَرْجِيح

٣٥ - تمثل معامِلَات التَرْجِيح بوجه عام نسب العناصر السكانية في الطبقات و

(ميم - ١٣)

$$W_b = \frac{N_b}{N}$$

$$\sum W_b = 1 \text{ ومن ثمَّ:}$$

(ميم - ١٤)

$$S_b^2 = \frac{1}{N-1} \sum_{i=1}^{N_b} (X_{hi} - \bar{X})^2$$

ميم - ٤ - ٣ - ٥ قيم العيّنات

٣٦ - قيمة العيّنة تقدير محسوب من عناصر  $nh$  مختارة في الطبقة. ونَصِفُ في هذا الجزء الرموز الشائعة المستخدمة في مسألة المعاينة الطباقية.

(أ) للطبقات  $H$  دعنا نرمز إلى أحجام العيّنات في كل طبقة بالرموز  $n_1, n_2, \dots, n_n$  حيث  $\sum n_b = n$  هو مجموع حجم العيّنة.

(ب) ليكن  $x_{hi}$  عنصر العيّنة  $i$  في الطبقة  $h$ .

(ج) ثم ليكن

(ميم - ١٥)

$$\bar{x}_b = \frac{1}{n_b} \sum_{i=1}^{n_b} x_{hi} \text{ متوسط العيّنات في الطبقة } h.$$

(د) ثم ليكن

(ميم - ١٦)

$$\bar{x}_{st} = \sum W_b \bar{x}_b \text{ متوسط إجمالي العيّنات.}$$

(هـ) ثم ليكن

(ميم - ١٧)

$$f_b = \frac{n_b}{N_b} \text{ كسر المعاينة في الطبقة } h.$$

يُعطى تباين العنصر  $n_h^{th}$  في الطبقة  $h^{th}$  بما يلي:

(ميم - ١٨)

$$v(\bar{x}_h) = \sum \left[ 1 - \frac{n_h}{N_h} \right] \frac{s_h^2}{n_h}$$

حيث  $s_h^2$  هو تباين العنصر في الطبقة  $h^{th}$ ، وهو مُعطى بما يلي:

(ميم - ١٩)

$$s_h^2 = \frac{\sum (x_{hi} - \bar{x}_h)^2}{(n_h - 1)}$$

(ميم - ٢٠)

$$v(\bar{x}_{st}) = \sum W_h^2 (1 - f_h) \frac{s_h^2}{n_h} \text{ : ويُعطى تباين متوسط العيّنات بما يلي:}$$

يُبحثُ فيما يلي أدناه نوعان من استراتيجيات المعاينة الطباقية، هما المعاينة المتناسبة والمعاينة غير المتناسبة.

#### ميم - ٤ - ٣ - ٦ التقسيم المتناسب إلى طبقات

٣٧ - يقتضي التخصيص المتناسب في المعاينة الطباقية استخدام كسر معاينة موحد في جميع الطبقات. وهذا ينطوي على اختيار نفس النسبة من الوحدات من كل طبقة. فإذا قررنا، مثلاً، أن نختار عيّنة يبلغ مجموعها ١٠ في المائة، فهذا يعني أننا سنختار ١٠ في المائة من وحدات كل طبقة. ولما كانت نسب المعاينة في جميع الطبقات واحدة، فإن عناصر العيّنة المختارة في العيّنة ستتفاوت من طبقة إلى أخرى. وفي كل طبقة، سيكون حجم العيّنة متناسباً مع عدد العناصر الموجودة في الطبقة.

في هذه الحالة يُعطى كسر المعاينة بـ  $f_h = \frac{n_h}{N_h} = \frac{n}{N}$  وهذا ينطوي على تصميم معاينة مع تساوي الاحتمالات (EPSEM).

(ميم - ٢١)

$$\bar{x}_{st} = \sum W_h \bar{x}_h \text{ : ومتوسط العيّنات هو:}$$

(ميم - ٢٢)

$$v(\bar{x}_{st}) = \frac{(1-f)}{n} \sum W_h s_h^2 \text{ : ويكون تباين المتوسط الإجمالي:}$$

#### ميم - ٤ - ٣ - ٧ التقسيم غير المتناسب إلى طبقات

٣٨ - يقتضي أسلوب المعاينة غير المتناسبة استخدام نسب معاينة مختلفة في طبقات مختلفة. والقصد هو تعيين نسب معاينة للطبقات بطريقة يمكن معها حصول أقل تباين ممكن في إجمالي المتوسط لتكلفة كل وحدة.

٣٩ - عند استخدام هذا الأسلوب تكون نسبة المعاينة في طبقة معيّنة متناسبة مع الانحراف القياسي لتلك الطبقة. وهذا يعني أن عدد وحدات المعاينة التي ستُختار من أية طبقة لا يتوقف على مجموع عدد العناصر فحسب، وإنما أيضاً على الانحراف القياسي للمتغير المساعد.

في التخصيص غير المتناسب، تُقدّم أيضاً فكرة وظيفة التكلفة. فمثلاً:

(ميم - ٢٣)

$$C = C_o + \sum c_h n_h$$

حيث  $C_o$  هي التكلفة الثابتة و  $c_b$  هي تكلفة تغطية العيّنة في طبقة معيّنة. يمكننا، في أوضاع كثيرة، أن نفترض أن  $c_b$  عدد ثابت في كل الطبقات. ومن بين المعادلات المستخدمة في تخصيص العيّنات غير المتناسب للطبقات معادلة تخصيص نيبمان. حيث  $c_b$  عدد ثابت و  $\sum n_h$ ، الذي هو حجم العيّنة الإجمالي، عدد ثابت. يُعطى عدد الوحدات المراد اختيارها من طبقة ما بما يلي:

(ميم - ٢٤)

$$n_b = \frac{W_b s_b n}{\sum W_b s_b} \quad \text{أو} \quad n_b = \frac{N_b s_b \cdot n}{\sum N_b s_b}$$

ويُعطى التباين بما يلي:

(ميم - ٢٥)

$$v(\bar{x}_{st}) = \frac{(\sum W_b s_b)^2}{n} - \frac{1}{N} \sum W_b s_b^2$$

الرقم الوارد إلى يسار علامة الناقص (-) عامل تصويب لمجموعة سكانية محدودة، يمكن إسقاطه إذا كنا نختار عيّنات من مجموعة سكانية كبيرة جداً، أي إذا كان كسر المعاينة صغيراً.

## ميم - ٤ - ٣ - ٨ ملاحظات عامة

- القيم السكانية  $S_b$  و  $C_b$  بوجه عام غير معروفة؛ لذلك، يمكن أن توضع التقديرات من استقصاءات معيّنة سابقة أو تجريبية
- التخصيص غير المتناسب ليس ذا كفاءة عالية لاختيار النسب
- ربما توجد أوجه تنازع فيما يتعلق بالمتغيرات المراد وضعها في موضع مثالي، في حالة الاستقصاءات متعددة الأغراض
- يسفر التخصيص غير المتناسب، بوجه عام، عن أقل تباين ممكن

٤٠ - توضّح الأمثلة الواردة أدناه حساب أحجام العيّنات والتباينات بموجب التقسيم المتناسب وغير المتناسب إلى طبقات. وفي هذا المثال الافتراضي تقسم المدارس إلى طبقات بناءً على أساس عدد الموظفين. ويبلغ مجموع عدد المدارس الابتدائية في الإقليم ٢٧٥ مدرسة. وتُختار منها عيّنة تضم ٥٥ مدرسة وتقسّم إلى طبقات بناءً على أساس عدد الموظفين.

## ميم - ٤ - ٣ - ٩ تقرير أحجام العيّنات داخل الطبقة

٤١ - ينبغي الرجوع إلى الجدول ميم - ١ أدناه.

التخصيص المتناسب

للتخصيص المتناسب يُستخدم كسر المعاينة المشترك.

لذلك  $f = \frac{n}{N}$  هو الكسر الإجمالي للمعاينة المطبق على مجموع عدد الوحدات في الطبقة.

وفي المثال المذكور أعلاه  $f = \frac{55}{275} = 0.2$  أو  $f = 20\%$  في المائة.

ويرد توزيع أحجام العيّنات في العمود ٤ من الجدول ميم - ١ أدناه؛ فحجم العيّنة في الطبقة ١، مثلاً هو  $n_b = 0.2 \times 80 = 16$ .

التخصيص غير المتناسب

تُعطى معادلة الحصول على أحجام العيّنات في الطبقات المختلفة بما يلي:

(ميم - ٢٦)

$$n_b = \frac{W_b s_b}{\sum W_b s_b} (n)$$

مثال ذلك، المعادلة  $n_b = \frac{0.3750}{2.4474} \times 55 = 8$  للطبقة ١.

يرد ما تبقى من النتائج في العمود ٥ من الجدول.

الجدول ميم - ١

عدد المدارس بحسب عدد الموظفين

الطبقة	عدد الموظفين في كل مدرسة مختارة ( $y_{hi}$ )	مجموع عدد المدارس في كل طبقة ( $N_b$ )	عدد المدارس المختارة بحسب الطبقة		$W_b$	$s_b^2$	$S_b$	$W_b S_b$	$W_b s_b^2$
			التخصيص غير المتناسب ( $n_b$ )	التخصيص المتناسب ( $n_b$ )					
١	٤،٢،٢،٤،٢،٤،٢،٢،٤،٢،٢،٤،٢،٢،٥،٥،٢،٢،٢،٢	٨٠	١٦	٨	٠،٢٩٠٩	١،٦٣٣	١،٢٨٩	٠،٣٧٥٠	٠،٤٨
٢	٨،٦،٧،٧،٧،٧،٦،٧،٦،٧،٦،٨،٧،٦،٨،٦،٦،٦،٦،٧،٦	١٠٠	٢٠	٦	٠،٣٦٣٦	٠،٥٣٧	٠،٧٣٣	٠،٢٦٦٥	٠،١٩
٣	١٥،١٠،١٢،١٠،٢٠،١٦،٢١،١٥،١٩،١٦،٢٠	٥٥	١١	١٨	٠،٢٠٠٠	١٥،٥٦٤	٣،٩٤٥	٠،٧٨٩٠	٣،١١
٤	٣٥،٣٥،٢٣،٤٧،٤٦،٤٨،٤٠،٥٠	٤٠	٨	٢٣	٠،١٤٥٥	٤٨،٨٣٦	٦،٩٨٩	١،٠١٦٩	٧،١٠
المجموع		٢٧٥	٥٥	٥٥	١،٠٠٠			٢،٤٤٧٤	١٠،٩٠

ملاحظة:  $N$  = مجموع عدد المدارس الابتدائية

$n$  = مجموع عدد المدارس الابتدائية في العيّنة كلها

$N_b$  = حجم الطبقة رقم  $b$

$n_b$  = حجم العيّنة من الطبقة رقم  $b$

## ميم - ٤ - ٣ - ١٠ حساب التباينات

٤٢ - يُوضّح حساب التباينات بموجب التقسيم المتناسب وغير المتناسب إلى طبقات بتطبيق المعادلتين ميم - ٢٧ و ميم - ٢٨، على التوالي.

التقسيم المتناسب إلى طبقات

(ميم - ٢٧)

$$V(\bar{y}_{prop}) = \frac{1-f}{n} \sum w_h s_h^2 = \frac{(1-0.2)}{55} (10.9) = 0.16$$

التقسيم غير المتناسب إلى طبقات

(ميم - ٢٨)

$$V(\bar{y}_{opt}) = \frac{(\sum w_h s_h)^2}{n} - \frac{1}{N} \sum w_h s_h^2 = \frac{(2.4474)^2}{55} - \frac{10.9}{275} = 0.07$$

## ميم - ٤ - ٣ - ١١ بوجه العموم

(ميم - ٢٩)

$$v(\bar{x}_{st})_{OP} \leq v(\bar{x}_{st})_{PROP} \leq v(\bar{x}_{st}) \leq (\bar{x}_{st})_{SRS}$$

## ميم - ٤ - ٤ معاينة المجموعات

٤٣ - المناقشات التي وردت في الفروع السابقة كانت حول أساليب معاينة اعتُبرت فيها وحدات المعاينة الابتدائية مرتبةً في قائمة من الإطار؛ وكان الترتيب أنه يمكن اختيار العيّنات الفردية مباشرةً من الإطار. وفي معاينة المجموعات، تحتوي وحدات الاختيار الأرفع مستوىً كمناطق تعداد السكان، مثلاً، (انظر الفصل الثالث)، على أكثر من وحدة ابتدائية واحدة. ووحدة المعاينة، في هذه الحالة، هي المجموعة. وثمة أسلوب بسيط لاختيار عيّنة عشوائية من الأسر المعيشية في مدينة، مثلاً، يستتبع وجود قائمة بجميع الأسر المعيشية في المدينة. وربما يكون هذا غير ممكن لأنه ربما لا يوجد، عملياً، إطار كامل بكل الأسر المعيشية في المدينة. ويمكن، لتجاوز هذه المشكلة، تشكيل مجموعات على شكل أحياء. ثم يمكن اختيار عيّنة من الأحياء، وفيما بعد يمكن وضع قائمة بالأسر المعيشية في الأحياء المختارة. ويمكن، إذا دعت الحاجة، تشكيل عيّنة من الأسر المعيشية من كل حي، تضم ١٠ في المائة من الأسر، مثلاً.

## ميم - ٤ - ٤ - ١ أسباب استخدام معاينة المجموعات

٤٤ - فيما يلي بعض الأسباب التي قُدّمت لتأييداً لاستخدام معاينة المجموعات، لا سيما في تصاميم العيّنات متعددة المراحل.

- يخفض تشكيل المجموعات تكاليف السفر وتكاليف أخرى متصلة بجمع البيانات
- يمكن أن يحسّن الإشراف والرقابة ومتابعة التغطية وجوانب أخرى لها تأثير في جودة البيانات التي يجري جمعها

- يصبح بناء الإطار أقل تكلفة لأنه يتم على مراحل. ففي المعاينة متعددة المراحل، مثلاً كما نوقشت في الفصل الثالث، لا يلزم إطار يغطي جميع السكان إلا لاختيار وحدات المعاينة الأولى، أي المجموعات على مستوى المرحلة الأولى. وفي أية مرحلة أدنى منها لا يلزم الإطار إلا في الوحدات التي اختيرت في المرحلة السابقة
- بالإضافة إلى ذلك، تميل أطُر الوحدات الأكبر حجماً والأعلى مرحلة إلى أن تكون أكثر دواماً ولذلك أكثر قابلية للاستخدام على مدى فترات أطول. وتميل قوائم الوحدات الصغيرة كالأُسَر المعيشية والأفراد، بوجه خاص، إلى أن تصبح قديمةً في غضون فترة زمنية قصيرة
- توجد فوائد إدارية لتنفيذ الاستقصاء

٤٥ - نلاحظ بوجه عام أننا، بمقارنة عيّنة مجموعات بعينة عناصر من نفس الحجم، نجد أن تكلفة العنصر في عيّنة المجموعات أقل، وذلك يعود إلى قلة تكاليف وضع قوائم بالعناصر و/أو التعرف على مواقعها. من جهة أخرى، يكون تباين العناصر أكثر تكلفة نتيجة لعدم انتظام تجانس العناصر (الترابط داخل الطبقة الواحدة) في المجموعات. ونوضّح المعاينة الأساسية للمجموعات بالنظر في تصميم ذي مرحلة واحدة (قُدِّمَت تصاميم متعددة المراحل ونوقشت بالتفصيل في الفصل الثالث).

#### ميم - ٤ - ٤ - ٢ معاينة المجموعات في مرحلة واحدة

٤٦ - ربما لا يكون من الممكن عملياً الحصول على قائمة بالأسر المعيشية، في مقاطعة بعينها، ثم اختيار عيّنة من هذه القائمة. غير أنه ربما يمكن أن نجد قائمةً بالقرى أعدت أثناء عملية تعداد سابقة واحتفظ بها لأغراض إدارية. في هذه الحالة، نحصل على عيّنة من القرى، ثم نحصل منها على معلومات عن الأسر المعيشية في القرى المختارة. هذا يمثل تصميم معاينة مجموعة ذات مرحلة واحدة، لأنه بعد اختيار عيّنة من القرى يمكن الوصول إلى جميع الوحدات في المجموعة — في هذه الحالة، الأسر المعيشية.

٤٧ - يمكن توضيح اختيار العيّنات من المجموعات كما يلي. افترض أن عيّنة من مجموعة من القرى قد اختيرت باحتمالات متساوية. فلمعاينة مجموعات في مرحلة واحدة، ستكون جميع الأسر المعيشية التي من القرى المختارة مشمولة بالعيّنة.

نظراً إلى أن:

$$A = \text{مجموع عدد القرى}$$

$$B = \text{مجموع عدد الأسر المعيشية في المجموعة}$$

$$a = \text{عيّنة من القرى}$$

ولذلك، أن

$$aB = n \text{ تمثل عدد الوحدات الابتدائية (الأسر المعيشية) في مجموع العيّنة}$$

و

$$AB = N \text{ هي مجموع عدد الأسر المعيشية في جميع القرى}$$

ثم يُعطى احتمال اختيار عنصر، مع تساوي الاحتمالات، بما يلي:

(ميم - ٣٠)

$$\frac{a}{A} \times \frac{B}{B} = \frac{n}{N} = f$$

حيث  $N$  هو مجموع عدد الوحدات الابتدائية و  $f$  هو كسر المعاينة. في هذه الحالة يكون الاحتمال، ببساطة  $\frac{a}{A}$ .

ميم - ٤ - ٤ - ٣ معادلات لمتوسط العينة والتباين

٤٨ - ترد أدناه معادلات لمتوسط العينة والتباين:

متوسط العينة

(ميم - ٣١)

$$\bar{y} = \frac{1}{aB} = \sum_{\alpha=1}^{\alpha} \sum_{\beta=1}^{\beta} \bar{y}_{\alpha\beta} = \frac{1}{a} \sum_{\alpha=1}^{\alpha} \bar{y}_{\alpha}$$

متوسط العينة تقدير غير متحيز لمتوسط السكان:

(ميم - ٣٢)

$$E(\bar{y}) = \frac{1}{A} \sum_{\alpha=1}^{\alpha} \bar{y}_{\alpha} = \bar{Y}$$

لأن حجم العينة في الواقع ثابت ( $aB = n$ ) وأن الاختيار مبني على أساس التساوي في الاحتمالات، فإن المتوسط ( $\bar{y}$ ) هو تقدير غير متحيز لمتوسط السكان  $\bar{Y}$ .

التباين

إذا اختيرت المجموعات باستخدام اختيار عشوائي بسيط، يمكن تقدير التباين كما يلي:

(ميم - ٣٣)

$$V(\bar{y}) = (1-f)s_{\alpha}^2$$

$$s_{\alpha}^2 = \frac{1}{a-1} \sum_{\alpha=1}^{\alpha} (\bar{y}_{\alpha} - \bar{y})^2 \quad \text{حيث:}$$

٤٩ - من المهم ملاحظة أن القيم خالية من أخطاء المعاينة، لأنها تستند إلى قيم جميع العناصر في  $B$  لا على أساس عينة. أما تباين متوسط العينة فمرده إلى التباينات بين متوسطات المجموعات.



## المرفق الثاني قائمة بأسماء الخبراء

الخبراء المشاركون في اجتماع فريق الخبراء التابع للأمم المتحدة لاستعراض  
مسودة الدليل المتعلق بتصميم استقصاءات الأسر المعيشية، نيويورك،  
٣ - ٥ كانون الأول/ديسمبر ٢٠٠٥ أ

الاسم	اللقب الوظيفي والانتماء
أولاديو أوبيليكه أجايي	خبير استشاري إحصائي، نيجيريا
بفرلي كارلسون	شعبة الإنتاج والإنتاجية والإدارة، اللجنة الاقتصادية لأمريكا اللاتينية والكاريبي، سنتياغو، شيلي
سمير فريد	خبير استشاري إحصائي، مصر
مافيون م. جامبوا	مستشار فني، الجماعة الإنمائية للجنوب الأفريقي/الاتحاد الأوروبي، غابروني، بوتسوانا
أودايا شانكار ميشرا	زميل مشارك، جامعة هارفارد، بوستون، ماساشوسيتس، الولايات المتحدة الأمريكية
جان كوردوس	أستاذ، مدرسة وارسو للاقتصاد، وارسو، بولندا
إدوين سانت كاترين	مدير، مكتب الإحصاءات الوطني، سانت لوسيا
أنتوني تيرنر	خبير استشاري في اختيار العينات، الولايات المتحدة الأمريكية
شيام أوباديايا	مدير، الخدمات الإحصائية المتكاملة (INSTAT)، نيبال
إبراهيم يانسانة	نائب الرئيس، شعبة تكاليف المعيشة، لجنة الخدمة المدنية الدولية، نيويورك، الولايات المتحدة الأمريكية

أ انظر الوثيقة: ESA/STAT/AC.93/L.4 للاطلاع على تقرير اجتماع فريق الخبراء.





---

### كيفية الحصول على منشورات الأمم المتحدة

يمكن الحصول على منشورات الأمم المتحدة من المكتبات ودور التوزيع في جميع أنحاء العالم . استعلم عنها من المكتبة التي تتعامل معها أو اكتب إلى : الأمم المتحدة ، قسم البيع في نيويورك أو في جنيف .

#### 如何购取联合国出版物

联合国出版物在全世界各地的书店和经售处均有发售。请向书店询问或写信到纽约或日内瓦的联合国销售组。

#### HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

#### COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre libraire ou adressez-vous à : Nations Unies, Section des ventes, New York ou Genève.

#### КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

#### COMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.

---