

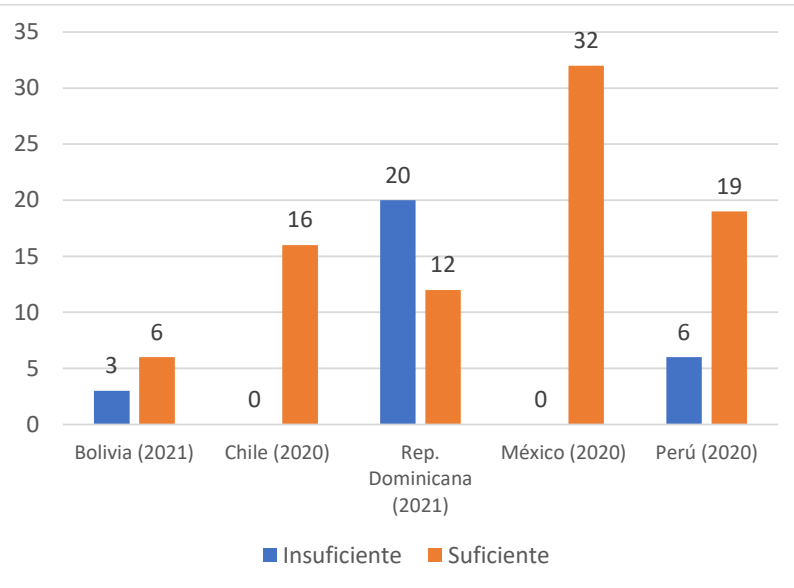
Recomendaciones metodológicas sobre la medición de la calidad de cifras provenientes de encuestas de hogares

**Methodological recommendations on
measuring the quality of estimates from
household surveys**

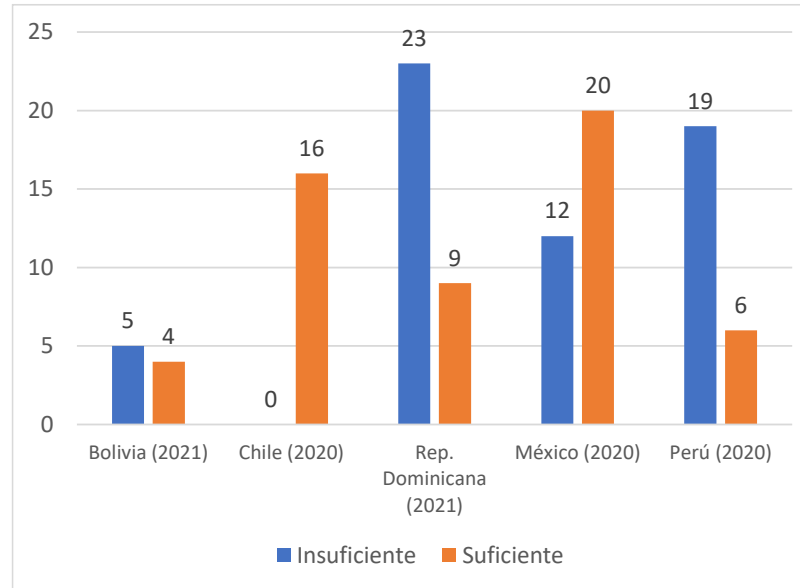
Antecedentes / Background

- Las encuestas por muestreo son una **fuentes de información** ampliamente utilizada.
 - Se espera que las estimaciones cumplan con el **principio de representatividad**: cada elemento incluido en la muestra no solo se representa a sí mismo sino también a un grupo de unidades que no pertenecen a la muestra seleccionada, brindando la posibilidad de hacer inferencias sobre la población.
 - Con la **disposición de las bases de datos** asociadas a las encuestas por muestreo, las ONE deben generar lineamientos claros de hasta donde es pertinente desagregar la información.
 - En la **XII Reunión de la CEA** (26-28 sep 2023) se aprobaron **estándares** para evaluar la calidad de cifras de encuestas de hogares.
- Sample surveys are an extensively used **source of information**.
 - It is expected that the estimates comply with the **principle of representativeness**: each element included in the sample not only represents itself but also a group of units that are not part of the selected sample, allowing the possibility to make inferences about the population.
 - With the **availability of databases** associated with sample surveys, NSOs need to provide clear guidelines on how far it is adequate to disaggregate the information.
 - At the **XII SCA Meeting** (September 26-28, 2023), **standards** were approved to evaluate the quality of household survey figures.

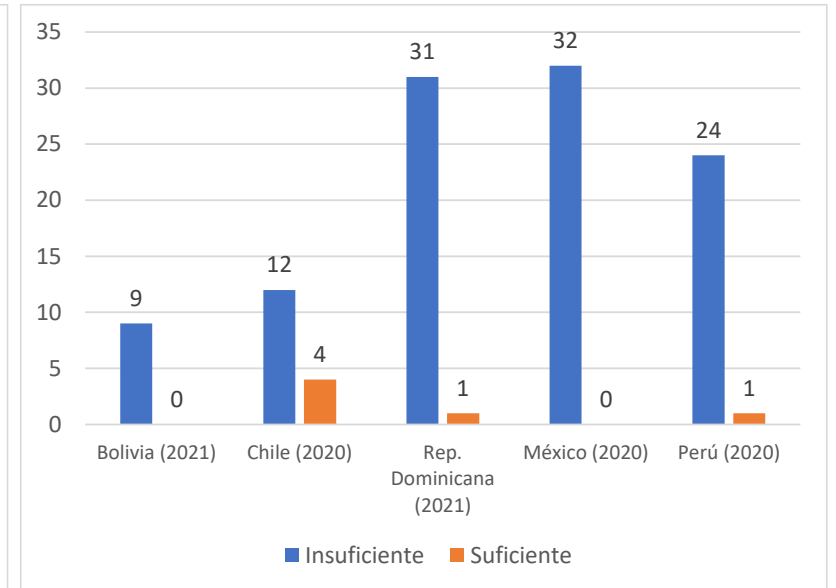
Ejemplo: Número de DAM¹ según calidad de las estimaciones / Example: Number of MAD¹ according to quality of estimates



Tasa de desocupación, 15 y más años
Unemployment rate, 15 years and older



Tasa de desocupación femenina, 15 y más años
Female unemployment rate, 15 years and older



Tasa de desocupación femenina, 25 a 29 años
Female unemployment rate, 25 to 29 years

(*) DAM = División Administrativa Mayor / MAD = Major Administrative Division

Producto / Product



RECOMENDACIONES METODOLÓGICAS SOBRE LA MEDICIÓN DE LA CALIDAD DE LAS CIFRAS PROVENIENTES DE ENCUESTAS DE HOGARES



TABLA DE CONTENIDO

1. Presentación	3
1.1 Antecedentes de aplicación de estándares de calidad por parte de las ONE 4	
2. Marco conceptual	11
2.1.1. Marco de muestreo	13
2.1.2. Tipos de muestreo	14
2.2. Diseños muestrales complejos y sus características	21
2.3. Parámetros poblacionales y sus estimadores	22
2.4. Intervalo de confianza	24
2.5. Estimación de los errores muestrales	26
2.5.1. Varianza y sus estimadores	26
2.5.2. Medidas de dispersión asociadas con la varianza	29
2.6. Tabulados y/o desagregaciones	33
3. Elementos para evaluar la calidad estadística de una estimación	34
3.1 Efecto de diseño (<i>Deff</i>)	35
3.2 Tamaño muestral mínimo para aplicar medidas de precisión 35	
3.2.1 Tamaño muestral	36
3.2.2 Tamaño de muestra efectivo	37
3.3 Grados de libertad	37
3.4 Tipología de estimadores	39
3.5 Medidas de precisión utilizadas para la publicación y difusión	42
3.5.1 Error estándar	43
3.5.2 Coeficiente de variación	44
3.5.3 Coeficiente de variación logarítmico	44
3.6 Cuento de casos no ponderado	46
4. Requerimientos de calidad para las estimaciones	46
4.1 Flujograma para la determinación de la calidad estadística	47
5. Consideraciones	50

Producto / Product

Alcance: Generar recomendaciones claras y precisas para evaluar las estimaciones provenientes de encuestas de hogares por muestreo que permitan dictaminar la factibilidad de publicación.

Scope: To provide clear and precise recommendations for evaluating estimates from sample household surveys, allowing for the assessment of the feasibility of publication.

Flujograma de evaluación de estimaciones

Trabajo doméstico no remunerado para el propio hogar por sexo y tramos de edad en un día tipo.
Región S/N

	Tramos de edad	Tiempo promedio (horas)	Tamaño muestral	Grados de libertad	Coefficiente de Variación (%)
AMBOS SEXOS	12 a 24 años	1,60	278	11	6,96
	25 a 45 años	3,69	373	11	4,72
	46 a 65 años	3,48	301	11	5,83
	66 años y más	3,76	158	10	4,52
HOMBRES	12 a 24 años	1,33	123	11	19,05
	25 a 45 años	2,35	137	11	11,61
	46 a 65 años	2,20	116	10	10,89
	66 años y más	2,53	69	9	7,51
MUJERES	12 a 24 años	1,87	155	11	9,30
	25 a 45 años	4,64	236	11	6,34
	46 a 65 años	4,49	185	11	4,164
	66 años y más	5,01	89	9	5,327

Flujograma de evaluación de tabulados

Tiempo promedio en trabajo doméstico no remunerado en día tipo, Región S/N

Tramos de edad	Sexo		
	Ambos sexos	Hombres	Mujeres
12 a 24 años	1,60	1,33 (*)	1,87
25 a 45 años	3,69	2,35	4,64
46 a 65 años	3,48	2,20	4,49
66 años y más	3,76	2,53	5,01

1. Efecto de diseño (Deff)/ Design effect

- **Fórmula:**

$$Deff(\hat{\theta}) = \frac{\widehat{Var}(\hat{\theta})}{\widehat{Var}_{mas}(\hat{\theta})} \approx 1 + (\bar{m} - 1) \cdot \hat{\rho}_{\hat{\theta}}$$

Where:

$\widehat{Deff}(\hat{\theta})$: Design effect, depends on the estimator and the sample design.

$\widehat{Var}(\hat{\theta})$: Estimation of variance under complex samples

$\widehat{Var}_{mas}(\hat{\theta})$: Estimation of variance under simple random sampling

\bar{m} : Average number of households selected per PSU (Primary Sampling Unit)

$\hat{\rho}_{\hat{\theta}}$: Intra-class correlation; it can occasionally be negative (rarely occurs)

- El Deff se calcula para la variable de interés al nivel en que se requiere evaluar la estimación.
- **Deff sea menor que 1** (pocas veces ocurre): para los casos que la estimación corresponda a un dominio de estudio planificado se sugiere revisar el mecanismo de estimación*. Asimismo, estudiar la homogeneidad del fenómeno en el dominio de estudio.
- Se recomienda establecer un umbral de aceptación igual o superior a 1 para el valor del efecto de diseño.
- The Deff is calculated for the variable of interest at the level at which the estimate is required to be evaluated.
- **If Deff is less than 1** (rarely occurs): for cases where the estimate corresponds to a planned study domain, it is suggested to review the estimation mechanism*. Additionally, consider examining the homogeneity of the phenomenon in the study domain.
- It is recommended to set an acceptance threshold equal to or greater than 1 for the design effect value.

**For the calculation of the variance under SRS (Simple Random Sampling) used in the Deff computation, it is essential to ensure the use of the finite population correction and to consider whether the estimate corresponds to a stratum or groupings of these, or if it pertains to a subpopulation.*

2. Tamaño muestral / Sample Size

Se considera como uno de los elementos más importantes a la hora de decidir la ruta de publicación de una cifra, puesto que los desarrollos teóricos en términos de inferencia estadística para encuestas dependen de este término. La cobertura de los intervalos de confianza y la distribución de los estimadores dependen de que tanto el tamaño de la subpoblación como su tamaño de muestra asociado no sean pequeños.

En este espíritu, Barnett-Walker et al. (2003) proponen que todas las estimaciones basadas en un tamaño de muestra menor a que 100 unidades deberían ser suprimidas o marcadas como no confiables.

Se recomienda un Tamaño muestral mínimo (TM_Min) de
100

It is considered one of the most critical elements when deciding on the publication path of a figure, since theoretical developments in terms of statistical inference for surveys depend on this term. The coverage of confidence intervals and the distribution of the estimators depend on both the size of the subpopulation and its associated sample size not being small.

In this spirit, Barnett-Walker et al. (2003) suggest that all estimates based on a sample size smaller than 100 units should be suppressed or marked as unreliable.

A minimum sample size (TM_Min) of 100 is suggested

** Each NSO must determinate the minimum sample size admissible to define if the estimates are reliable, based on the characteristics of their designs.*

3. Effective Sample Size (TME)

The effective sample size of the estimate should be compared with what is denoted as **Minimum Effective Sample Size (TME_Min)**, which will be obtained according to the following steps:

Example when the unit of analysis is a household*:

1. Efecto diseño de la variable a evaluar a nivel de vivienda	$d = 2,5$
2. Total de las UPM planificadas	$x = 15$ UPM
3. Viviendas promedio por UPM planificadas	$y = 12$ viviendas promedio por UPM
<i>Tamaño de muestra planificado -></i>	$x \cdot y = 12 \cdot 15 = 180$
4. Tasa de efectividad (Viviendas con respuesta para la estimación /Viviendas planificadas)	$z = 0,85$
<i>Tamaño de muestra planificado ajustado por efectividad -></i>	$x \cdot y \cdot z = 12 \cdot 15 \cdot 0,85 = 153$
Tamaño de muestra efectivo mínimo (TME_Min) ->	$(x \cdot y \cdot z) / d = (12 \cdot 15 \cdot 0,85) / 2,5 = 61,2 \sim 61$

**Calculations will be carried out at the level under which the estimate to be evaluated is calculated, which, in the context of household surveys, could be dwellings, households, or individuals.*

Based on the acceptance thresholds of the INE-Chile and ECLAC standards (60 and 100, respectively), it is suggested to use acceptance thresholds between 60 and 100, with the aim of ensuring the convergence of the distribution of the estimators, and thus, being able to make inferences about the population.

4. Grados de libertad (GL)/ Degrees of Freedom

Calculation of Degrees of Freedom:

- Variables:**

$$df_d(V) = \begin{array}{l} \text{\#PSU with observations in subpoblation} \\ \text{minus} \\ \text{\#Strata with observations in subpoblation} \end{array}$$

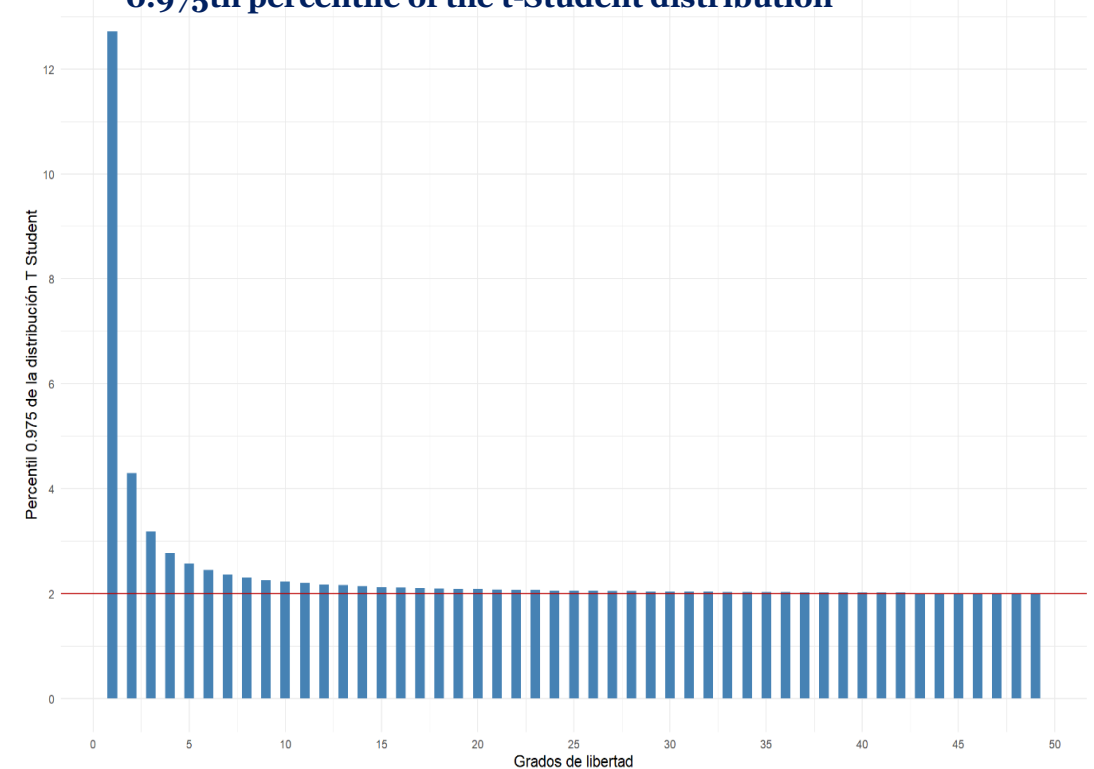
- Fixed:**

$$df_d(F) = \begin{array}{l} \text{N\# PSU minus} \\ \text{N\# Strata} \end{array}$$

*Nota: Proporción y razones definidas entre 0 y 1, el cálculo de los df y tm se realiza sobre las unidades de análisis en la subpoblación que participan en el "Denominador". Resto, el cálculo se realiza tomando en cuenta las unidades de análisis que participan en la estimación.

*Note: For proportions and ratios defined between 0 and 1, the calculation of df (degrees of freedom) and tm is carried out on the units of analysis in the subpopulation that participate in the "Denominator." For the rest, the calculation is done taking into account the units of analysis that participate in the estimate.

0.975th percentile of the t-Student distribution



-Parker et al. (2017) la siguiente recomendación: si los grados de libertad para alguna proporción de interés son menos de ocho, entonces la proporción debe ser marcada para revisión, cuyo resultado deriva en publicación o eliminación de la cifra.

-Parker et al. (2017) make the following recommendation: if the degrees of freedom for any proportion of interest are fewer than eight, then the proportion should be flagged for review, the outcome of which results in either publication or removal of the figure.

4. Grados de libertad (GL) / Degrees of Freedom (DF)

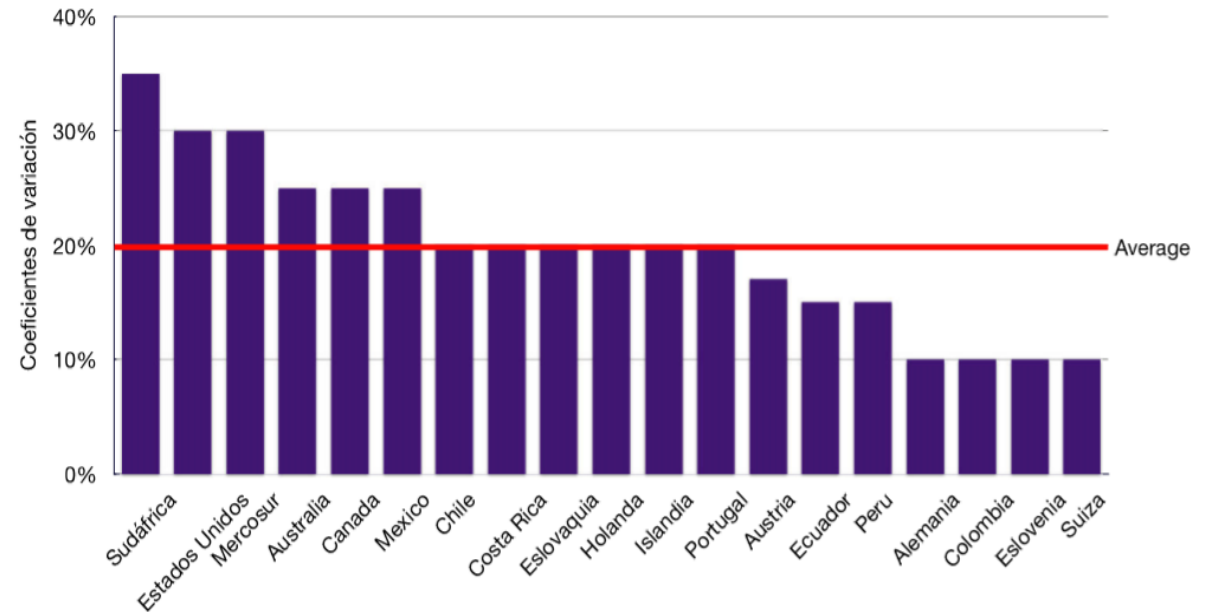
- Se sugiere como **Grado de libertad mínimo (GL_Min)** valores entre 9 y 13.
- Si la estimación no tiene suficientes grados de libertad, no podrá ser publicada, a excepción de los casos que correspondan a dominios de estudio y se justifique tener pocos grados de libertad.
- Se debe emplear el cálculo de los grados de libertad variables, es decir: (#UPM con observaciones en la subpoblación) menos (#Estratos con observaciones en la subpoblación)*.
- A minimum **Degrees of Freedom (GL_Min)** value between 9 and 13 is suggested.
- If the estimate does not have enough degrees of freedom, it cannot be published, except in cases that correspond to study domains and where having a few degrees of freedom is justified.
- The calculation of variable degrees of freedom should be used, that is: (#PSUs with observations in the subpopulation) minus (#Strata with observations in the subpopulation)*.

***Nota: Proporción y razones definidas entre 0 y 1**, el cálculo de los *df* y *tm* se realiza sobre las unidades de análisis en la subpoblación que participan en el "Denominador". **Resto**, el cálculo se realiza tomando en cuenta las unidades de análisis que participan en la estimación.

***Note: For proportions and ratios defined between 0 and 1**, the calculation of *df* (degrees of freedom) and *tm* is carried out on the units of analysis in the subpopulation that participate in the "Denominator." **For the rest**, the calculation is done taking into account the units of analysis that participate in the estimate.

5. Coefficient of Variation(CV)

- Las **mayorías de las ONE** establecen como dato “no publicable” cuando el error relativo supera un 20%; sin embargo, este no es el único elemento a ser considerado.
- **Most NSO** consider data "non-publishable" when the relative error exceeds 20%. However, other elements should also be taken into account.
- Este elemento se empleará para evaluar estimaciones que **NO** son del tipo Proporción **N_i** del tipo Razón entre 0 y 1.
- Se recomienda los siguientes valores según nivel de aceptación: **CV_a**: {15,20} ; **CV_b**: {25, 30}



Source: Presentation on the Limitations of Household Surveys in Measuring Social Indicators, Regional Workshop on the Disaggregation of Social Statistics through Small Area Estimation Methodology by Andrés Gutiérrez, Regional Expert on Social Statistics - ECLAC .

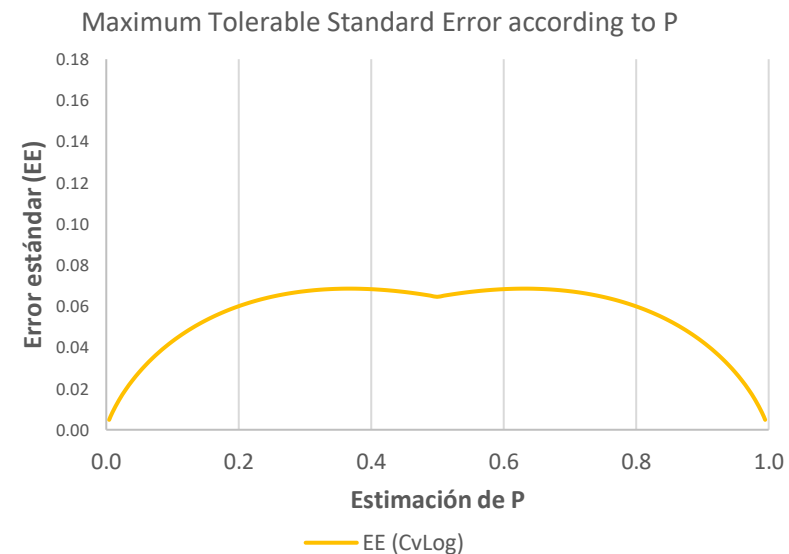
- This element will be used to evaluate estimates that are NOT of the Proportion type NOR of the Ratio type between 0 and 1.
- The following values are suggested according to the level of acceptance: **CV_a**: {15,20} ; **CV_b**: {25, 30}

5. Coeficiente de variación logarítmico / Logarithmic Coefficient of Variation (Cvlog)

Este elemento se empleará para evaluar estimaciones del tipo Proporción y Razón entre 0 y 1.

This element will be used to evaluate estimates of the Proportion and Ratio type between 0 and 1.

$$CVlog = \begin{cases} \frac{CV(p)}{-\log(p)}, p < 0,5 \\ \frac{CV(p)}{-\log(1-p)}, p \geq 0,5 \end{cases}$$



Logarithmic function

6. Conteo de casos no ponderados / Count of unweighted case

El conteo de casos no ponderados en una muestra corresponde al conteo de unidades dentro de la muestra que son afectados por un fenómeno de interés en estudio.

The count of unweighted cases in a sample corresponds to the count of units within the sample that are affected by a phenomenon of interest under study.

Se sugiere los siguientes umbrales de aceptación:

- Umbral 1: CCNP_a: {30}
- Umbral 2: CCNP_b: {40,50}

The following acceptance thresholds are suggested:

- Threshold 1: CCNP_a: {30}
- Threshold 2: CCNP_b: {40,50}

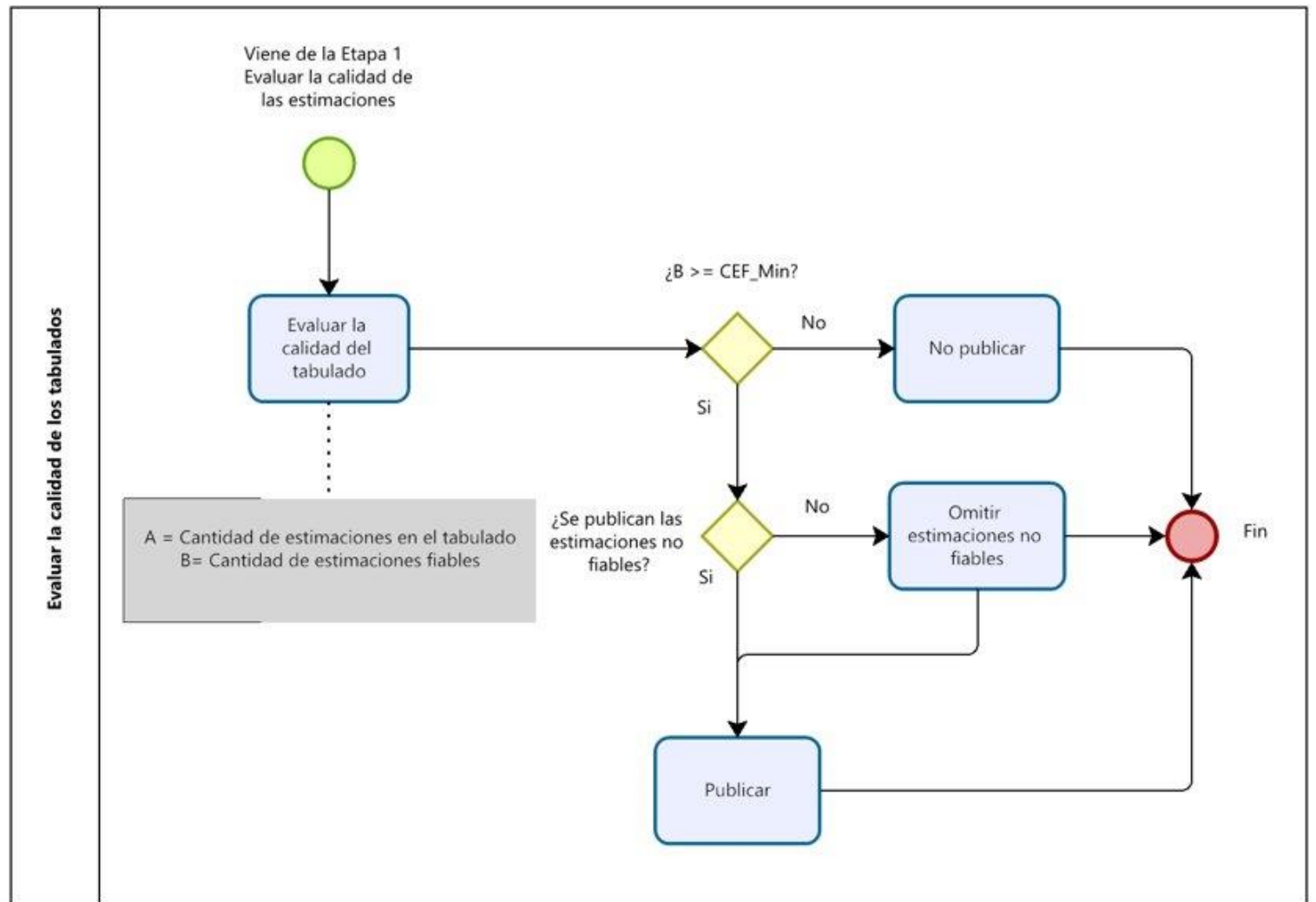
Producto / Product

Flujograma armonizado para evaluar calidad de los tabulados

Harmonized flow chart to evaluate the quality of tabulates

Se recomienda no publicar el tabulado si
It is recommended not to publish the table if

$$B \leq A/2 + 1$$



Resumen / Summary

Estas recomendaciones metodológicas posibilitan que las ONE generen lineamientos estandarizados para evaluar los resultados de las encuestas de hogares por muestreo del país.

Estas recomendaciones permiten complementar el CV, métrica mayormente empleada por los entes del SEN y las ONE.

These methodological recommendations enable the NSO to generate standardized guidelines for evaluating the estimates obtained from household sample surveys.

The recommendations complements the CV, a metric mostly used by the entities of the NSS and the NSO.