Data Evaluation with Consecutive Censuses: Adult Mortality and Census Coverage

United Nations Statistics Division
Outline

1. The Population Balancing Equation
2. Adult Mortality (Death Distribution Methods)
3. Intercensal Cohort Survival Rates
4. Cohort Component Method
Population Balancing Equation

Census coverage
Population balancing equation

If a country has:

- A relatively complete system of vital registration
- A fairly reliable estimate of the degree of under-registration

- Information on the number of intercensal births, deaths and net international migrants can be used in conjunction with the results of a previous census to evaluate the coverage of a subsequent or current census.
Population balancing equation

\[ P_1 = P_0 + B - D + M \]

Where:

- \( P_1 \) = the population enumerated in the census being evaluated
- \( P_0 \) = The population enumerated in a previous census
- \( B \) = the number of births in the period between the two censuses
- \( D \) = the number of deaths in the period between two censuses
- \( M \) = the number of net international migrants in the period
  \[ M = I \text{ (Immigrants)} - E \text{ (Emigrants)} \]
Population balancing equation

- The population balancing equation is the most fundamental equation in demographic analysis and is also used to estimate population growth.

- It is based on the logic that:
  
  The population of a country can increase or decrease between any two points in time only as a result of births, deaths and movement of population across national boundaries
  
  > Births and immigration add to the population
  
  > Deaths and emigration reduce the population
Population balancing equation

- For census evaluation purposes, there is a residual \( (e) \) needed to make the equation balance exactly.
- "\( e \)" in the equation is referred to as the "error of closure" and represents the balance of errors in the data on births, deaths, net migration, and the coverage of the two censuses:

\[
P_1 = P_0 + B - D + M + e
\]

- If a negative residual quantity \( e \), \( P_1 \) is under-enumerated relative to \( P_0 \).
- If a positive residual is required to balance the equation, \( P_1 \) is over-enumerated relative to \( P_0 \).
Population balancing equation – Data required

- The population enumerated in two consecutive censuses
  - \( P_1 \): the census under evaluation
  - \( P_0 \): previous census

- The number of births, deaths and net international migration (immigrants-emigrants) during the intercensal period, adjusted for under-registration (to the extent possible)
Population balancing equation – Computational Procedure

1. Compile registered numbers of intercensal births, deaths and migrants
   - Vital registration system
   - Immigration record system (residence permit, border records, etc.)
     - Adjustment based on under-coverage of these systems including indirect estimates

2. Calculation of the “expected” census population \( E(P_1) \)
   \[ E(P_1) = P_0 + B - D + M \]

3. Calculation of the residual error or error of closure
   \[ e = P_1 - E(P_1) \]
Population balancing equation – Interpretation of “e”

- If $P_0$ has been adjusted for net coverage error, the estimated residual error ($e$) will represent an estimate of net coverage error in $P_1$
  - If “$e$” is positive, $P_1$ is overenumerated
  - If “$e$” is negative, $P_1$ is underenumerated

- If $P_0$ is not adjusted, “$e$” will represent an estimate of the relative level of net coverage error in $P_1$ in comparison with $P_0$

For an unadjusted census:

\[ E(P_1) = P_0 \text{ (unadjusted)} + B_{\text{adj}} - D_{\text{adj}} + M_{\text{adj}} \]
\[ = 12,689,897 + 3,716,878 - 1,002,108 + (-446,911) \]
\[ = 14,957,756 \]

\[ e = P_1 - E(P_1) \]
\[ = 14,848,364 - 14,957,756 \]
\[ = -109,392 \quad 0.7\% \text{ of } E(P_1) \]

\( P_1 \text{ is under-enumerated relative to } P_0 \)

Source: U.S. Census Bureau (1985)
Population balancing equation – Example
Sri Lanka, 1971 and 1981 Censuses (2)

For an adjusted census count:

\[ E(P_1) = P_0 \text{ (adjusted)} + B_{\text{adj}} - D_{\text{adj}} + M_{\text{adj}} \]
\[ = 12,849,796 + 3,716,878 - 1,002,108 + (-446,911) \]
\[ = 15,117,655 \]

\[ e = P_1 - E(P_1) = 14,848,364 - 15,117,655 \]
\[ = -269,291 \quad 1.8\% \text{ of } E(P_1) \]

\[ P_1 \text{ is underenumerated} \]

Source: U.S. Census Bureau (1985)
Population balancing equation – Limitations

- Incomplete and defective data on the components of population change are very common
  - Applicability of the method is limited to countries with good vital registration coverage and migration data

- It is generally not useful for obtaining estimates of net census coverage error for sub-national populations (for example regions, provinces).
  - In addition to the components of population change considered, internal migration has to be considered.
  - For most practical purposes, the use of the population balancing equation is limited to analysis of net coverage error at the national level.
Death Distribution Methods

Completeness of reporting of adult mortality
Death distribution methods

“Death distribution methods” apply the logic of the population balancing equation to different age groups in the population.

- E.g. for the age group 40 – 50, the only way to enter the age group in a country is through aging or immigration, the only way to exit is through death or emigration.

- By comparing our expectation for the size of an age group at the time of the census to its actual enumerated size, we can get a sense of whether we have “missing” or “extra” people in the enumeration.
Death Distribution Methods - Advantages

- Can provide timely estimates of age-specific period mortality rates – here we will use the method to check estimates of completeness of death reporting

- Data requirements:
  - Population by sex and 5-year age groups
  - Deaths by sex and 5-year age groups
  - Can be computed with data from two consecutive censuses with an estimate of the number of deaths between the two censuses

Source: Moultrie et al. (2013)
Death Distribution Methods – Assumptions and Violations (1)

- Completeness of deaths reporting is the same across ages
  - Generally does not hold for the oldest and youngest age groups
  - To avoid, usually truncate analysis to middle age ranges

- (Two-census variant) Coverage of both censuses is the same for all age groups
  - Census coverage evaluation will be discussed later in this session

- Age reporting (by 5-year age groups) is accurate
  - Can be checked through age-sex distribution techniques discussed in previous sessions

Death Distribution Methods – Assumptions and Violations (2)

- Net in-migration is limited
  - Will depend on country context

- (One-census variant) population is stable (constant growth rate over past several decades)
  - Will depend on country context – in contexts with recent fertility decline, will not hold
Data quality issues

- Common errors in data on recent deaths by age
  - Under-reporting, especially for child deaths and older age deaths
  - Reference period errors in reporting of deaths (i.e. reporting deaths that occurred prior to the usual 12-month reference period)
  - Death question omitted by interviewers
  - Household breaking up due to the death of a senior household member
    - In this case, any deaths in household will not be captured
  - Age-heaping and age exaggeration

- In addition to age-sex distribution checks discussed in previous sessions, the age and sex structure of reported deaths should be examined prior to conducting any analysis
Data quality checks:
Schedules of death rates by age and sex

Sweden 1980

Cambodia 2008
Data quality checks: Comparison with other surveys

Source: Graph produced based on IPUMS-International and DHS country report
General Growth Balance Method (GGB)

Basis: The Balancing Equation of Population Change

\[ P_2 = P_1 + B - D + G \]

Assumptions:
- a) population is closed to migration, \( G = 0 \);
- b) completeness of first census, \( k_1 \), is independent of age;
- c) completeness of second census, \( k_2 \), is independent of age;
- d) completeness of intercensal deaths, \( c \), is independent both of age and year;
**GGB regression:** \[ b(x+) - r(x+) = \beta_0 + \beta_1 d(x+) \]

- Relative completeness of censuses
  \[ \frac{k_1}{k_2} = \exp(t\hat{\beta}_0) \]

- Completeness of death registration
  \[ c = \frac{1}{(k_1 k_2)^2} \hat{\beta}_1 \]

- Adjusting observed death rates
  \[ m = \hat{\beta}_1 m^* \]

**Ideal case:**

- \( r(x+) \) - population growth rate above age \( x \)
- \( b(x+) \) - entry rate at age \( x+ \), “birth rate”
- \( d(x+) \) - open age death rates
- \( m^* \) - observed death rates
- \( m \) - adjusted death rates

---

**United Nations Workshop on Census Evaluation**

Hanoi, Viet Nam

2–6 December 2013
GGB regression with migration:

\[ b(x+) - r(x+) + g(x+) = \beta_0 + \beta_1 d(x+) \]

\[ g(x+) = \text{net migration rate, age } x+ \]

A common case >> significant emigration
If migration is not accounted for \((g(x)=0)\), estimate of the slope will be less than one, the regression line will be increasing less steep and the completeness of deaths will be overestimated (as emigration reduces the population)
GGB Method – Application

Source: Moultrie et al. (2013)
GGB Method – Estimating intercensal deaths

If accurate data for intercensal deaths is not available, they can be estimated if deaths for two other well-defined periods are available – e.g., deaths in the year prior to two different censuses.

Worksheet will compute growth rates and deaths for the intercensal period.
GGB Method – Setting the age range

1. **Upper age range**
   - Note the lower age bound – we may adjust later

2. **Open age group**
   - Upper age range must be 1 less than start of open age group

---

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
GGB Method - Diagnostics

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
1) Ideally we want all points to fall on the line.
2) If, right-hand points (for older age groups) are falling below the line → age exaggeration. 
> We must then lower the upper age limit so to exclude these points. Should do this progressively, by 5-year age groups, until all points are on line.
3) Recommended not to select a last age group ending in ‘0’
4) If left-hand points (younger age groups), particularly ages 15–30 are deviating from the line this likely indicates that our migration data are insufficient (or, if they are missing entirely, that there is indeed significant migration that we have missed). We should then raise the lower limit to age 30 or 35.
GGB Method - Diagnostics

Residuals should not exceed 0.01
GGB Method – Two census (6)
Interpretation

Check the estimate of completeness of death reporting and reasonableness of the analysis

Compare with the results for the opposite sex – unless we have reason to believe that completeness will vary significantly by sex, should be fairly close

Compare with results of Synthetic Extinct Generations approach (worksheets also available through IUSSP)
Synthetic Extinct Generations method (Bennett & Horiuchi)

• Used for estimating completeness of death registration, with different inputs

• Population at exact age, $N$, can be computed from registered deaths, $D$, and intercensal rates of increase $r$:

$$N(a) = \int_a^\infty D(x)e^{\int_a^x r(u)du} \, dx$$

• Software:
  Ken Hill’s spreadsheet, *Death_dist_method all template.xls*
  IUSSP Tools for Demographic Estimation’s spreadsheet, *AM_SEG_. . . .xls*
Cohort Survival Ratios

Mortality and census coverage
Cohort survival ratios

This technique is based on a comparison of the size of birth cohorts enumerated in successive censuses.

In the absence of census errors, the ratio of the number of persons in a cohort enumerated in the second census to the number enumerated in the first census should approximate the survival rate that would be expected on the basis of mortality conditions.

E.g. we have a cohort of males aged 40 – 44 at the time of the first census, say in 2000.

If the next census is held exactly 10 years later, in 2010, this cohort will be aged 50 – 54.

In the absence of other factors, we expect their numbers to have been reduced only by the life table quantity $10d_x = l_x - l_{x+n}$, the number of deaths to those aged $x$ over the subsequent 10 years.

Source: U.S. Census Bureau (1985); Moultrie et al. (2013)
Cohort Survival Ratios: Caveats

The method is less useful when other factors make it difficult to determine whether deviations from the expected CSR are due to census error or something else

- Substantial net migration (unless there are accurate estimates of net migration by age)
- Changes in country borders between censuses
- Changes in the population groups included in the two censuses (e.g. active military, nomadic groups) if the size of these groups is substantial
Cohort Survival Ratios: DPR Korea


United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
Calculating CSRs (1)

Intercensal cohort survival rates are defined as:

\[ n \text{CSR}_x (a) = \frac{nP_{x+a} (t+a)}{nP_x (t)} \]

Where:
- \( t \) = time of first census
- \( a \) = number of years between censuses
- \( nP_x (t) \) = size of the cohort at the time of the first census
- \( nP_{x+a} (t+a) \) = size of the cohort at the time of the second census
Calculating CSRs (2)

The ratio of the observed intercensal cohort survival rate to the corresponding life-table survival rate

\[ nR_x = \frac{nP_{x+a}(t+a)}{nP_x(t)} \]

Where:

- \( nP_{x+a}(t+a) \) = size of the cohort at time of the second census
- \( nP_x(t) \) = size of the cohort at the time of the first census
- \( nL_{x+a} \) = the life table number of person-years lived in the age interval \( x+a \) to \( x+a+n \) years
- \( nL_x \) = the life table number of person-years lived in the age interval \( x \) to \( x+n \) years
Cohort Survival Ratio - Interpretation

In the absence of census error, the expected value of the ratio \( \frac{nR_x}{n_{x}} \) would be 1.0.

Ratio values for any particular cohort which exceed 1.0 would indicate over-enumeration of the cohort in the second census relative to the first census.

Ratio values of less than 1.0 would indicate under-enumeration of the cohort in the second census relative to the first census.
Cohort survival ratios – Example (1)

**Step 1: Adjustment for migration (if appropriate)**

- In countries experiencing significant levels of net intercensal immigration, the number of net immigrants in each cohort may either added to the cohort enumerated in the first census or subtracted from the cohort enumerated in the second census.

- In cohorts experiencing net intercensal emigration, the number of net intercensal emigrants can either added to the second census or subtracted from the first census.

- Should be confident that migration data is reasonably accurate before making any adjustments.
Cohort survival ratios – Example (2)

**Step 2:** Calculation of census survival rates using two consecutive censuses $n\bar{P}_{x+a}(t+a) / n\bar{P}_x(t)$
Cohort survival ratios – Example (3)

Step 3: Calculation of life table survival rates based on the expected level of mortality

\[ nS_x = \frac{nL_x + a}{nL_x} \]

Step 4: Calculation of cohort survival ratios (nRx)
Cohort survival ratios – Example (4)

Male and female census survival rates, Brazil 2000 - 2010

Census and life table survival rates for Brazilian females 2000 - 2010

Census and life table survival rates for Brazilian males, 2000 - 2010
Cohort Survival: Uses and limitations

It is a widely applicable approach for examining error in consecutive censuses.

Method requires relatively little information.

Information on the level of fertility is not required since the method does not assess the coverage of the population born between two censuses.

Method is complicated by migration etc. as discussed.

When only two censuses are available, the method suffers from the limitations shared by many demographic methods, namely difficulties in separating census errors from real irregularities caused by extraordinary events.

The utility of census survival approaches increases significantly when three or more censuses are available.
Cohort Component Method

Census coverage
Overview of cohort-component method


United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
Cohort component method

In this approach, the population enumerated in the first census is projected to the reference date of the second census based on estimated levels and age schedules of fertility, mortality and migration during the intercensal period.

The expected population from the projection is then compared with the actual population enumerated in the second census.

Data for intercensal births, deaths and migration are taken from estimates and/or assumptions regarding the level and age schedules of these parameters rather than directly available data based on registration systems.

Cohort component method – data required

1. The population enumerated in two censuses by age and sex
2. Age specific fertility rates for women aged 15 to 49 (in 5-year age groups), assumed to represent the level and age structure of fertility during the intercensal period
3. Life table survival rates for males and females, assumed to be representative of mortality conditions during the intercensal period
4. An estimate of sex ratio at birth
5. Estimates of the level and age pattern of net international migration during the intercensal period if the level of net migration is substantial
Cohort component method – overview of computational steps

1. “Survive” the age distribution at the initial census to the time of the second census
   1. Multiply each age group population by life table survival rates
   2. Open-ended interval requires special handling
2. Make any necessary adjustments for migration
3. Calculate the number of births during the period
   1. Average initial and projected population for each age group between 15 – 49 to estimate mid-period female population
   2. Apply age-specific birth rates to these populations to generate total numbers of births during time period
   3. Apply sex ratio factor to get female and male births from total births
4. Apply life table survivorship to these births to determine number that survive to time of the second census
5. Compare the estimated female population by age group with the enumerated female population
Cohort component method – Step 1 (survive initial age distribution)

1. “Survive” the age distribution at the initial census to the time of the second census

\[ nS_x = \frac{nL_{x+a}}{nL_x} \]

Oldest age group requires special treatment
2. For the oldest age category (open-ended)

\[ wS_x = \frac{wT_{x+a}}{wT_x} \]

- \( w = \) the oldest age attainable in the population
- \( a = \) the length of the projection interval
- \( wS_x = \) the life table survival ratio for the population aged \( x \) and above
- \( wT_x = \) the number of life table persons lived at ages \( x \) and above
- \( wT_{x+a} = \) the number of life table person-years lived at ages \( x+a \) and above

Still need to estimate youngest cohorts based on fertility data.
Cohort component method – Step 2 (adjust for migration)

If net international migration is substantial, the “survived” cohort population must be adjusted to reflect the effects of migration.

- The introduction of net migrants by age group at the mid-point of the projection period and the survival of net migrants to the end of the period:

\[ n \hat{M}_{x+i} = \frac{1}{4} n M_x (1 + n S_x) + \frac{1}{4} n M_{x+i} (1 + n S_{x+i}) \]

**Assumptions:** i) An equal distribution of net migrants across years of the intercensal period, ii) Migrants have the same fertility and mortality level as the enumerated population.
Cohort component method – Step 3 (calculate births)

1. Calculate the average number of women in each childbearing age group (15 – 49) during the intercensal period in order to estimate the number of births during the projection period

\[
\bar{n}_{P_x} = \frac{n \cdot P^0_x + n \cdot P^1_x}{2}
\]

\( n_{P_x} \) = average number of females aged \( x \) to \( x+n \) in the projection period

\( n_{P^0_x} \) = number of females aged \( x \) to \( x+n \) at the beginning of the projection period

\( n_{P^1_x} \) = projected number of females aged \( x \) to \( x+n \) at the end of the projection period
2. Calculate total births during the period

\[ B = \sum_{x=15}^{49} (n_{P_x} * n_{f_x}) \] for 1-year projection

\[ B = 5 \sum_{x=15}^{49} (n_{P_x} * n_{f_x}) \] for 5-year projection period

\( B \) = the estimated number of births during the projection period

\( \bar{n}_{P_x} \) = the average number of women in the age group \( x \) to \( x+n \) years during the projection period

\( n_{f_x} \) = the age specific fertility rate (per woman) for women age \( x \) to \( x+n \) years during the projection period
Cohort component method –
Step 3 (calculate births)

3. Calculate proportion of male and female births

\[
B^f = 1 - \frac{\text{SRB}}{1 + \text{SRB}} \\
B^m = \frac{\text{SRB}}{1 + \text{SRB}}
\]

\[
B^f = 1 - \frac{1.05}{1 + 1.05} = 0.488 \\
= 0.488 \times 17,521,248 = 8,550,369
\]

\[
B^m = \frac{1.05}{1 + 1.05} = 0.512 \\
= 0.512 \times 17,521,248 = 8,970,879
\]
3. Apply life table survivorship to these births to determine number that survive to time of the second census

\[ 5P_0^f = B_0^f \times 5S_0 \]

\[ 5S_0 = \frac{5L_0}{5 \times l_0} \]

\[ 5S_0 = \frac{484,129}{500,000} = 0.968 \]

\[ 5P_0^f = 0.968 \times 8,550,369 = 8,276,757 \]
Cohort component method – Step 5 (compare with enumerated population)

Final step in procedure is to compare the enumerated population by age and sex in the second population with the expected population.
Cohort component method in MortPak (1)
Cohort component method in MortPak (2)

<table>
<thead>
<tr>
<th>SELECT YEAR TO DISPLAY</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population by single year</td>
<td>2013</td>
<td>2014</td>
<td>2015</td>
</tr>
<tr>
<td>Age</td>
<td>0-4</td>
<td>5-9</td>
<td>10-14</td>
</tr>
<tr>
<td>0</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>1</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>2</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>3</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>4</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>5</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>6</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>7</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>8</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>9</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>10</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>11</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>12</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>13</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>14</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>15</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>16</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>17</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>18</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>19</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>20</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>21</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>22</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>23</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>24</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>25</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>26</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>27</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>28</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>29</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>30</td>
<td>168565</td>
<td>168565</td>
<td>168565</td>
</tr>
<tr>
<td>Total</td>
<td>5072981</td>
<td>5072981</td>
<td>5072981</td>
</tr>
<tr>
<td>Births</td>
<td>1094259</td>
<td>1094259</td>
<td>1094259</td>
</tr>
<tr>
<td>Deaths</td>
<td>1928658</td>
<td>1928658</td>
<td>1928658</td>
</tr>
</tbody>
</table>

Single-year population projection based on cohort-component technique.

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
Brazil Example with MortPak - Results

<table>
<thead>
<tr>
<th>Age</th>
<th>Acutal Enumeration 10 Aug 2010</th>
<th>MortPak Projections</th>
<th>Absolute Difference (Enumerated - Expected)</th>
<th>Percent Difference (Absolute Difference/Expected * 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>0-4</td>
<td>7,010,937</td>
<td>6,779,172</td>
<td>7,023,614</td>
<td>7,053,633</td>
</tr>
<tr>
<td>5-9</td>
<td>7,554,231</td>
<td>7,546,234</td>
<td>8,337,924</td>
<td>8,059,780</td>
</tr>
<tr>
<td>10-14</td>
<td>8,725,413</td>
<td>8,441,348</td>
<td>8,208,591</td>
<td>8,006,289</td>
</tr>
<tr>
<td>15-19</td>
<td>8,559,808</td>
<td>8,482,002</td>
<td>9,344,660</td>
<td>9,111,077</td>
</tr>
<tr>
<td>20-24</td>
<td>8,300,227</td>
<td>8,614,963</td>
<td>8,942,999</td>
<td>8,524,649</td>
</tr>
<tr>
<td>25-29</td>
<td>8,459,985</td>
<td>8,454,418</td>
<td>8,770,813</td>
<td>8,802,289</td>
</tr>
<tr>
<td>30-34</td>
<td>7,717,657</td>
<td>8,026,855</td>
<td>7,780,340</td>
<td>8,013,690</td>
</tr>
<tr>
<td>35-39</td>
<td>6,755,585</td>
<td>7,121,146</td>
<td>6,559,571</td>
<td>6,941,450</td>
</tr>
<tr>
<td>40-44</td>
<td>5,320,570</td>
<td>6,665,797</td>
<td>6,107,012</td>
<td>6,558,418</td>
</tr>
<tr>
<td>45-49</td>
<td>5,002,053</td>
<td>6,141,558</td>
<td>5,655,145</td>
<td>6,132,654</td>
</tr>
<tr>
<td>50-54</td>
<td>4,934,995</td>
<td>5,305,407</td>
<td>4,744,148</td>
<td>5,206,109</td>
</tr>
<tr>
<td>55-59</td>
<td>3,902,344</td>
<td>4,273,675</td>
<td>3,807,072</td>
<td>4,283,113</td>
</tr>
<tr>
<td>60-64</td>
<td>3,041,034</td>
<td>3,466,085</td>
<td>2,945,144</td>
<td>3,323,514</td>
</tr>
<tr>
<td>65-69</td>
<td>2,249,060</td>
<td>2,618,745</td>
<td>2,032,200</td>
<td>2,489,215</td>
</tr>
<tr>
<td>70-74</td>
<td>1,067,373</td>
<td>2,074,246</td>
<td>1,150,776</td>
<td>1,990,408</td>
</tr>
<tr>
<td>75-79</td>
<td>1,090,518</td>
<td>1,472,920</td>
<td>1,037,825</td>
<td>1,426,400</td>
</tr>
<tr>
<td>80+</td>
<td>655,522</td>
<td>866,349</td>
<td>1,040,648</td>
<td>1,693,618</td>
</tr>
<tr>
<td>Total</td>
<td>92,942,491</td>
<td>96,544,695</td>
<td>93,778,618</td>
<td>97,211,031</td>
</tr>
</tbody>
</table>

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2-6 December 2013
Brazil Example with MortPak - Results

Percent difference in expected and enumerated Brazilian population, 2010 census

Age Group

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
Main findings from Brazil example

Suggests underenumeration by about 1.5 million people, or 0.8% of the population

Significant underenumeration of youngest two age groups, particularly children 0 - 4

Some overenumeration of 10 – 14 year olds – could be a result of the underenumeration of this group (as 0 – 4 year olds) in the original census in 2000

Seeing same under-enumeration of 25 – 29 year olds of both sexes as when we calculated by hand – needs to be explored

Consistent but fairly low level of overenumeration of adults age 35 – 39 to 75 – 79

Could potentially indicate in-migration, ideally want to incorporate migration data

Significant overenumeration of older people

Very likely that there is age exaggeration

Also might consider that our life table is not accurate for these ages
<table>
<thead>
<tr>
<th>Age Group</th>
<th>Male</th>
<th>Female</th>
<th>Males &amp; Female</th>
<th>Net-Pop Projections</th>
<th>Absolute Difference (Enumerated - Expected)</th>
<th>Percent Difference (Absolute Difference/Expected * 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>2,042,676</td>
<td>2,266,559</td>
<td>2,044,061</td>
<td>2,268,061</td>
<td>3,615,657</td>
<td>3,082,901</td>
</tr>
<tr>
<td>5-9</td>
<td>1,687,905</td>
<td>1,973,616</td>
<td>1,769,212</td>
<td>1,975,212</td>
<td>3,615,305</td>
<td>2,983,152</td>
</tr>
<tr>
<td>10-14</td>
<td>1,295,519</td>
<td>1,486,447</td>
<td>1,456,133</td>
<td>1,546,133</td>
<td>3,729,941</td>
<td>2,961,239</td>
</tr>
<tr>
<td>15-19</td>
<td>1,740,730</td>
<td>1,819,410</td>
<td>1,791,380</td>
<td>1,891,380</td>
<td>3,558,414</td>
<td>2,701,380</td>
</tr>
<tr>
<td>20-24</td>
<td>1,372,049</td>
<td>1,509,781</td>
<td>1,441,830</td>
<td>1,541,830</td>
<td>3,477,780</td>
<td>2,599,860</td>
</tr>
<tr>
<td>25-29</td>
<td>1,554,766</td>
<td>1,729,981</td>
<td>1,652,542</td>
<td>1,729,542</td>
<td>3,468,658</td>
<td>2,770,178</td>
</tr>
<tr>
<td>30-34</td>
<td>885,760</td>
<td>905,908</td>
<td>899,350</td>
<td>909,350</td>
<td>1,252,255</td>
<td>1,451,682</td>
</tr>
<tr>
<td>35-39</td>
<td>705,401</td>
<td>715,140</td>
<td>703,140</td>
<td>713,140</td>
<td>1,283,470</td>
<td>1,312,682</td>
</tr>
<tr>
<td>40-44</td>
<td>596,420</td>
<td>611,420</td>
<td>599,420</td>
<td>613,420</td>
<td>1,289,060</td>
<td>1,337,725</td>
</tr>
<tr>
<td>45-49</td>
<td>410,646</td>
<td>432,620</td>
<td>412,620</td>
<td>434,620</td>
<td>1,286,720</td>
<td>1,319,640</td>
</tr>
<tr>
<td>50-54</td>
<td>232,763</td>
<td>234,748</td>
<td>227,748</td>
<td>232,748</td>
<td>1,305,460</td>
<td>1,317,748</td>
</tr>
<tr>
<td>55-59</td>
<td>254,942</td>
<td>270,412</td>
<td>256,412</td>
<td>270,412</td>
<td>1,314,860</td>
<td>1,340,460</td>
</tr>
<tr>
<td>60-64</td>
<td>142,269</td>
<td>165,383</td>
<td>146,383</td>
<td>165,383</td>
<td>1,270,460</td>
<td>1,308,540</td>
</tr>
<tr>
<td>65-69</td>
<td>118,911</td>
<td>135,878</td>
<td>123,878</td>
<td>135,878</td>
<td>1,345,811</td>
<td>1,378,791</td>
</tr>
<tr>
<td>70-74</td>
<td>117,822</td>
<td>137,036</td>
<td>116,036</td>
<td>137,036</td>
<td>1,332,252</td>
<td>1,368,082</td>
</tr>
<tr>
<td>75-79</td>
<td>51,970</td>
<td>59,346</td>
<td>60,346</td>
<td>59,346</td>
<td>983,752</td>
<td>1,042,178</td>
</tr>
<tr>
<td>80-84</td>
<td>91,925</td>
<td>120,410</td>
<td>99,410</td>
<td>120,410</td>
<td>1,078,717</td>
<td>1,149,720</td>
</tr>
<tr>
<td>Total</td>
<td>19,772,823</td>
<td>20,135,125</td>
<td>20,243,983</td>
<td>20,243,983</td>
<td>39,830,190</td>
<td>40,269,723</td>
</tr>
</tbody>
</table>

United Nations Workshop on Census Evaluation
Hanoi, Viet Nam
2–6 December 2013
Kenya (2)

The graph shows the percent difference between expected and enumerated populations of Kenya, 2009 census, by age group. The data is broken down by gender, with blue lines representing males and pink lines representing females. The percent difference is plotted against age groups ranging from 0-4 years to 80+ years. The graph highlights the variability in population counts across different age groups and genders, with some age groups showing significant discrepancies.
Kenya (3)

Overall suggests net undercount of 1.1%, about 460,000 people

Most of undercount is coming from males aged 20-45 and females aged 25–59

Migration may account for some of this difference

The lifetable used (based on Kenya 1999) census may not accurately represent changing mortality conditions over the 10 year period due to the HIV/AIDS epidemic

There may be a ‘true’ undercount of these age-sex groups
Cohort component method – uses and limitations

- It is applicable when registration data are not-existent or deficient to such an extent that satisfactory adjustment is not possible.

- Sufficient information to derive estimates of fertility and mortality levels should be available.
  - Mortality estimates can be complicated by HIV/AIDS – with a generalized epidemic, one life table is generally not sufficient to model mortality patterns over a 10 year period.

- Lack of information on international migration is often a problematic issue when applying this method.

- In case where sufficient information exists to derive reliable estimates of demographic parameters, the method is perhaps the most powerful among the alternative demographic approaches for the evaluation of censuses, since it provides age and sex specific estimates of net census error.
Tools

- In addition to MortPak, the DemProj module of Spectrum can be used for population projections
- [http://www.futuresinstitute.org/spectrum.aspx](http://www.futuresinstitute.org/spectrum.aspx)
  - DemProj is recommended for projection in contexts in which HIV/AIDS prevalence exceeds a few percent – better modeling of mortality conditions
  - Requires more data input, including prevalence and treatment estimates for HIV/AIDS
  - Data input options somewhat less flexible than MortPak
References


THANK YOU...

धन्यवाद  terima kasih
Ташаккур  Рахмат!
cảm ơn bạn ขอบขอบคุณคุณ
谢谢  Kaadinchhey La
танд баярлалаа