Demography Special Lecture (3)　　21 April 2016

Age and sex structure

The prime objective of formal demography is to develop measures which eliminate the effects of population size and age structure so as to "pure" measures of vital events. The age and sex structure of a population depends on the previous numbers of births, deaths, and moves: In other words, population structure and vital events are interdependent each other.

To use R code given as **`http://minato.sip21c.org/demography-special/code20140424.R`**, start R (or EZR, and in this case, type R code to R script window), and read that file from "File" => "Open script...".

1. The measurement of age
    1. Ages obtained by interview or questionnaire: "25" means "after the 25th birthday, before 26th" (age last birthday). Very unusual measurement, comparing with the fact that height of "170 cm" means "between 169.5 cm and 170.5 cm". Sometimes "exact age" or "nth birthday" are used to distinguish those from "age last birthday".
    2. Age group "0-4" includes everyone up to exact age 5. Phrases involving cumulation like "the proportion dying by age 5" refer to an exact age.
2. Age misreporting
    1. Distortions in age distributions may be caused by past changes in the levels or patterns of mortality, fertility, and migration, or errors. The errors may be caused by someone omitted from the survey or by someone's ages being misreported.
    2. Age misreporting can take two basic forms: heaping and shifting. In developing countries, reporting preference on 0 or 5 in last digits (which is a typical heaping) must be paid attention (Figure 3.2a). Whipple's Index (=[Sum of the populations of ages 25, 30, ..., 60] / [Sum of the populations aged 23-62] $\times$ 100 $\times$ 5) is a measurement of heaping: "highly accurate" for under 105, "fairly accurate" for 105-110, "approximate" for 110-125, "rough" for 125-175, "very rough" for over 175.
    3. Shifting (eg. older people tend to exaggerate their ages, young men tend to understate their age to avoid military service, older single women tend to understate their ages) is harder to detect and adjust for than heaping.
3. Population pyramid
    1. Elegant and useful way to graphically present an age/sex distribution.
    2. The rules to draw pyramids are generally the same as those for producing histograms: Conventionally males should be drawn on the left, females on the right; usually last open-ended age group is omitted; the bottom scale can be graduated as either absolute numbers or percentages; the choice of scales affects greatly the final shape of the pyramid. To compare different pyramids, percentages are preferable.
    3. To draw population pyramid using computer software, pyramid package of R is useful.
4. Sex ratios
    1. The number of males divided by the number of females $\times$ 100.
    2. Sex ratios are determined by the sex ratio at birth, age/sex specific mortality rates, and age/sex specific migration rates.
    3. In most parts of the world, sex ratio at birth is higher than 100, and mortality rates are lower in females than in males. In India and some other high-mortality countries, female mortality rates tend to be continuously higher than male mortality rates, probably due to worse malnutrition among young females and the risks of maternity.
5. Dependency ratios
    1. Dependency ratio = (Children + Elderly) / (Working ages) $\times$ 100
    2. Child dependency ratio = (Children) / (Working ages) $\times$ 100
    3. Aged dependency ratio (or Old-age dependency ratio) = (Elderly) / (Working ages) $\times$ 100
        * Children: ages 0-14; Working ages: ages 15-64; Elderly: ages 65-, changing by society
6. The determinants of age structure
    1. Mortality, fertility and migration affect age structure in different ways
    2. Fertility affects the youngest age group at first, and the older age groups after many decades
    3. Variation in the level of infant mortality may produce similar effects with changes in fertility. Changes in adult mortality generally have much less effects on age structure.
    4. Migration may cause big distortions in age structure, since young adult males are easily moving. Areas with large immigration tend to have young age structure, many children, few elderly. Areas with large emigration will have top-heavy pyramids.
7. The impact of age and sex structure on vital events
    1. Vital events => age/sex structure. How about the opposite direction?
    2. If high proportion of a population is in the age groups which are at greatest risk of giving birth, dying or moving, then the numbers of these events occurring will be higher than if there are small numbers in these age groups.
    3. If a large number of births occur in a population, there will be a large number of births between 20-30 years later. The populations of most developing countries currently contain a very high proportion of young children, and these will begin childbearing during the next 20 or so years. Thus, even if the extraordinary rapid decline of fertility occur, there will be large numbers of births. It's called as "demographic momentum".

Additional materials for exercises

1. How to calculate Whipple's Index for Japanese data
    1. The Whipple's Index should be 100 if there is no preference of reporting ages ending at 0 or 5, and 500 at maximum if all ages reported are ended at 0 or 5.
    2. In the fmsb package, you can see the explanation of Jpopl by typing ?Jpopl.
    3. After loading fmsb library by typing **`library(fmsb)`** as R script, you can see the age 0 males population in 2010 by typing **`Jpopl[1,"M2010"]`** or **`Jpopl$M2010[1]`**.
    4. To calculate the sum of the males' populations of ages 25, 30, 35, 40, 45, 50, 55 and 60 in 2010 Japan, type **`sum(Jpopl[26+0:7*5, "M2010"])`**, which means **`sum(Jpopl[c(26,31,36,41,46,51,56,61),"M2010"])`**.
    5. Thus, to get the Whipple's Index for Japanese males in 2010, type as follows.
       `sum(Jpopl[26+0:7*5,"M2010"])/sum(Jpopl[24:63,"M2010"])*100*5`
    6. The result is 100.9381.
    7. Similarly, Whipple's Index for Japanese females in 2010, type as follows.
       `sum(Jpopl[26+0:7*5,"F2010"])/sum(Jpopl[24:63,"F2010"])*100*5`
    8. The result is 100.8595.
    9. Both results are very close to 100, so that we can judge almost no preference of reporting ages ending at 0 or 5 in recent Japan. For earlier years (for example, **`"M1888"`** can be used instead of **`"M2010"`**), the results are similar.
2. Make population pyramids for Japanese data
    1. To draw population pyramid, you should load the pyramid package by typing **`library(pyramid)`**. After that, the simplest way to draw the population pyramid for Japanese population in 2010 is,
       `pyramids(Left=Jpopl$M2010,Right=Jpopl$F2010,Center=Jpopl$Age)`
    2. It will generate an ugly graph. To improve this, at first, type only selected ages in center.
       `pyramids(Left=Jpopl$M2010,Right=Jpopl$F2010,Center=Jpopl$Age,Cstep=10)`
    3. Next, set the details of population axis. In this case, maximum values of age-specific males population in Japan in 2010 is **`max(Jpopl$M2010)`**, which is a little more than 100000. Probably the labels for population axis are to be 0, 200000, 400000, 600000, 800000, 1000000, 1200000. However, such too many zeros are difficult to see. Instead, the population can be divided by 1000, then the labels for population axis can be 0, 2000, 4000, 6000, 8000, 10000, 12000 (unit of population: 1000). To do so, type as follows (Caution! type in one line. no carriage-return.)
       ```
       pyramids(Left=Jpopl$M2010/1000, Laxis=0:6*200, Right=Jpopl$F2010/1000,
       Center=Jpopl$Age, Cstep=10, Llab="Males(x1000)", Rlab="Females(x1000)",
       Clab="", main="Population pyramid for Japan in 2010")
       ```
    4. When you type **`windows()`** or **`x11()`**, you can see a new graphic window. And type as follows:
       ```
       pyramids(Left=Jpopl$M1950/1000, Laxis=0:6*200, Right=Jpopl$F1950/1000,
       Center=Jpopl$Age, Cstep=10, Llab="Males(x1000)", Rlab="Females(x1000)",
       Clab="", main="Population pyramid for Japan in 1950")
       ```
    5. You can compare the two population structures by moving two windows by mouse.