USA TFR by sex

Carl Schmertmann

6 Nov 2019

Notation

Suppose that there are A reproductive ages $a = 1 \ldots A$ in which either men or women may have children. At some of these ages, notably 50+, male fertility may be possible even though female fertility is not.

Suppose also that there are a total of B births in a given year, and that p_{xy} is the fraction of those births for which the mother is x years old and the father is y years old.

Call the $\mathbf{P}$ the $A \times A$ matrix of p_{xy} values, with mother's age corresponding to rows and father's age to columns. The marginal distribution of mothers' ages for this year's births is $\mathbf{P1} \in \mathbb{R}^A$, and the marginal distribution of fathers' ages is $\mathbf{P'1} \in \mathbb{R}^A$, where $\mathbf{1}$ is an $A \times 1$ vector of ones.

The vector of births by mother's age is thus

$$\mathbf{b}_{female} = B \cdot \mathbf{P}\,\mathbf{1} \in \mathbb{R}^A$$

and the vector of births by father's age is

$$\mathbf{b}_{male} = B \cdot \mathbf{P'}\,\mathbf{1} \in \mathbb{R}^A$$

Define $\mathbf{W} = diag(W_1 \ldots W_A)$ as a diagonal matrix with population counts of *women*, and $\mathbf{M} = diag(M_1 \ldots M_A)$ as a diagonal matrix with population counts of *men*.

Age-specific birth rates and TFR for women are therefore

$$\mathbf{f}_{female} = \mathbf{W}^{-1}\mathbf{b}_{female} = B \cdot \mathbf{W}^{-1}\mathbf{P}\,\mathbf{1}$$

and

$$TFR_{female} = \mathbf{1}'\mathbf{f}_{female} = B \cdot \mathbf{1}'\mathbf{W}^{-1}\mathbf{P}\,\mathbf{1}$$

Because $\mathbf{P1} \in \mathbb{R}^A$ is the marginal distributions of mother's ages for this year's births, we can interpret this as

$$TFR_{female} = B \ \cdot \ E_x\left(\frac{1}{W_x}\right)$$

where the expectation is weighted by mothers' ages of this year's births.

An analogous derivation for men produces age-specific rates

$$\mathbf{f}_{male} = \mathbf{M}^{-1}\mathbf{b}_{male} = B \cdot \mathbf{M}^{-1}\mathbf{P'}\,\mathbf{1}$$

and

$$TFR_{male} = \mathbf{1}'\mathbf{f}_{male} = B \cdot \mathbf{1}'\mathbf{M}^{-1}\mathbf{P'}\,\mathbf{1}$$

We can interpret this as

$$TFR_{male} = B \ \cdot \ E_y\left(\frac{1}{M_y}\right)$$

where the expectation is weighted by fathers' ages of this year's births.

The difference of the sex-specific TFRs is

$$TFR_{male} - TFR_{female} = B \cdot \left[E_y\left(\frac{1}{M_y}\right) - E_x\left(\frac{1}{W_x}\right)\right]$$

In very broad terms, we can see that this difference is more likely to be positive (higher male TFR) if there are more women and fewer men in the reproductive-age population. More subtly, the age distributions of males and females also affect the difference in sex-specific TFRs.

Example: US TFRs in 1975

```
# USA male fertility
# from Dudel, C. and S. Klüsener. Male fertility data for high-income
# countries [unpublished data]. Submitted to the HFC by C. Dudel # on 15.05.2019.
#
# downloaded from Human Fertility Collection 2 Nov 2019
# https://www.fertilitydata.org/data/RAW_DATA/m_USA_51.zip


# Male 1975 rates from HFC, as described above
asfr_male = read.table(file='m_ASFR_USA.txt', header=TRUE) %>%
              filter(Year=='1975') %>%
              pull(ASFR)

# male data pulled is for ages 15-59. Add zero rates for 12-14
asfr_male = c( rep(0,3), asfr_male)

# Female 1975 rates from HFD
asfr_female = read.table(file='USAasfrTR.txt', skip=2, header=TRUE) %>%
       mutate(Age = 11 + as.numeric(Age)) %>%
       group_by(Year,Age) %>%
       summarize(ASFR = mean(ASFR)) %>%
       filter(Year=='1975') %>%
       pull(ASFR)

# female data pulled is for ages 12-54. Add zero rates for 56-59
asfr_female = c( asfr_female, rep(0,4))

# 1975 US populations

pop = read.table('Population.txt',skip=2, header=TRUE) %>%
        filter(Year=='1975', Age %in% 12:59) %>%
        select(Age,Female,Male)

Bx = pop$Female * asfr_female
By = pop$Male   * asfr_male

px = prop.table(Bx)
py = prop.table(By)

female_factor = weighted.mean( 1/pop$Female, w=px)
male_factor   = weighted.mean( 1/pop$Male  , w=py)
```