# Demography (15)
# Microsimulation models

Minato Nakazawa, Ph.D.
16 July 2020

# Meaning of microsimulation in demography

- Bennett Dyke (1981) Computer simulation in anthropology. *Ann. Rev. Anthropol.*, 10: 193-207.
  - 2 meaningful ways to use microsimulation
    - Artificial data generation: Unobservable data can be generated by microsimulation (eg.) Complete rebuilding of vital events history for 200 years at individual level is impossible. We can evaluate the possible bias caused by unreliable data collection. A kind of computer intensive statistics.
    - Complicated phenomenon, impossible to build analytic model, can be evaluated by microsimulation. (eg.) mating process

# Microsimulation model for human reproduction

- Definition of individual: Up to the purpose
  - Here each individual has ID, SEX, AGE, MRD (initially 0, spouse's ID after marriage), NC (number of children ever born), LCB (year at latest childbirth) as variables in data.frame
  - If we consider the postpartum amenorrhea, LCB should be recorded at month
  - Initial population was calculated by stable population theory
- Mortality: Lifetable is needed.
  - Actual life table, model life table, or mechanism-based model (Avalanche model).
  - Here qx from 2010 Japan's lifetable for males and females
- Mating: Coale-McNeil's model was applied
  - CM(scale=0.9, a0=16, k=1)
- Fertility: Coale-Trussell's model  was applied
  - CT(M=0.9, m=1.0)

# Stochastic occurrence of vital events

- For deterministic macro simulation, if N people have an event at probability p, the expected number of events is Np.  If considering stochasticity in R, rbinom(1, N, p) generates expected number with binomial distribution.

- For microsimulation, if N people have an event at probability p, generate uniform random numbers in (0, 1) N times, if the number is equal to or less than p, the event occur.  It should be recorded for each individual.  Otherwise, rbinom(N, 1, p) also simulate the event's occurrence for each individual (1=event, 0=no event).
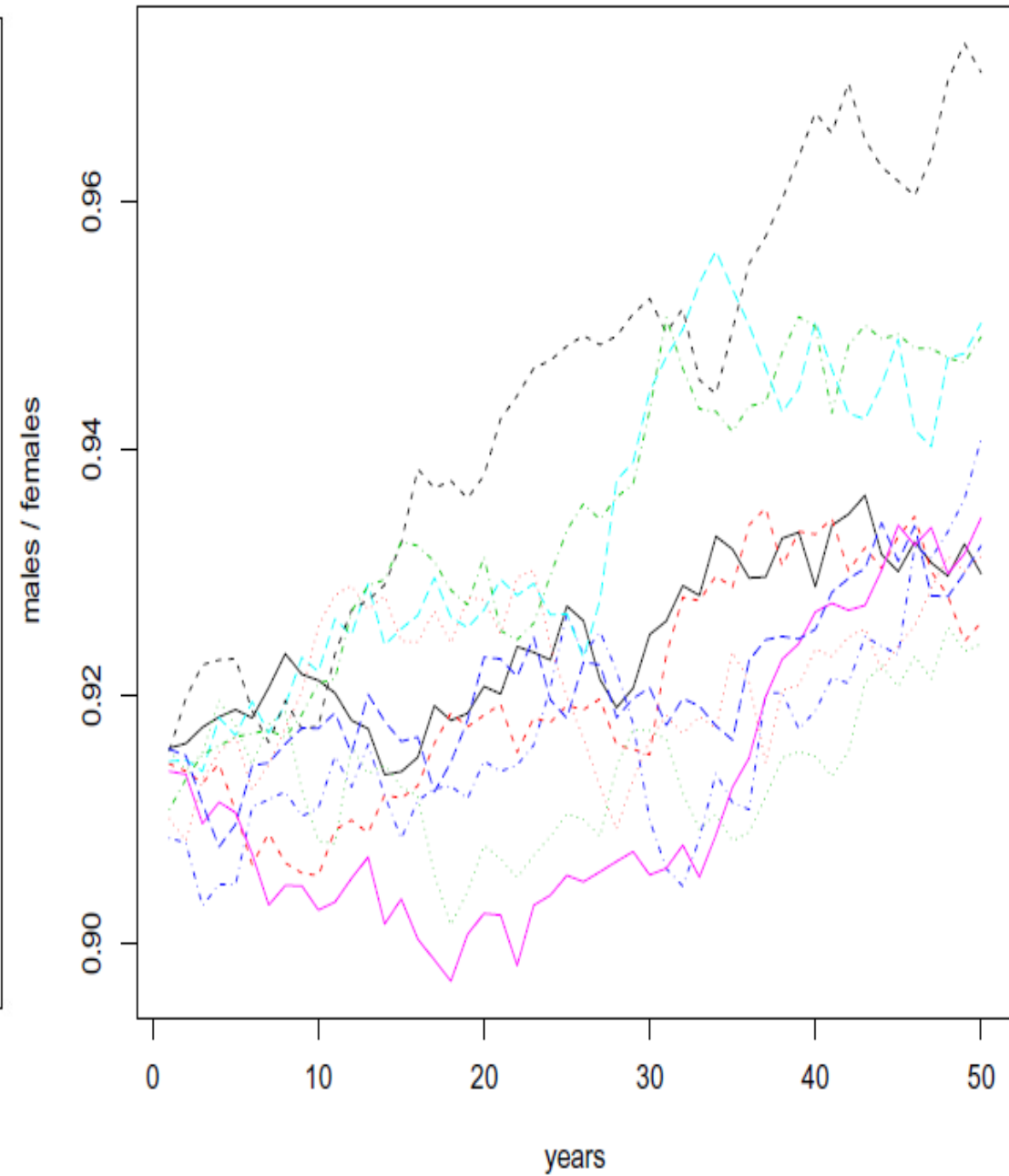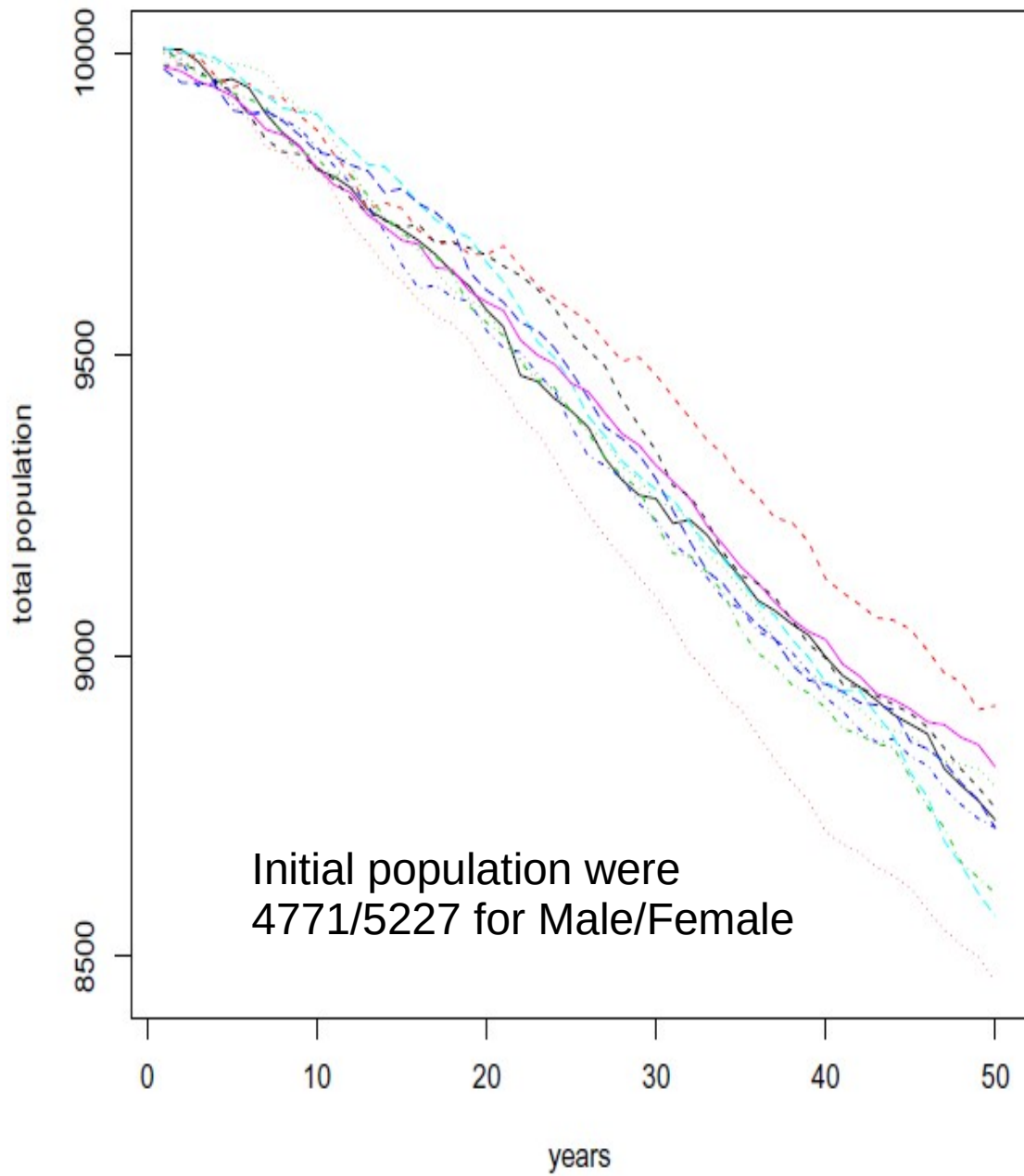
# Pseudo random numbers

- In computer simulation, random (by-chance) events are simulated by pseudo random numbers.

- Many kinds of pseudo-random number generators (PRNG) have been ever suggested.
  - Traditionally, Lehmer's linear congruential method was widely used (eg. Park and Miller, 1988 [https://www.researchgate.net/publication/220420979_Random_Number_Generators_Good_Ones_Are_Hard_to_Find]).
  - Nowadays, Mersenne-Twister (Matsumoto and Nishimura, 1998 [https://ir.lib.hiroshima-u.ac.jp/files/public/1/15032/20141016122634147579/ACMTraModel_8_3.pdf]) is *de facto* standard.
  - In R, RNGkind() can specify various PRNG, but usually leave it as default Mersenne-Twister.
  - In R, set.seed() can specify the initial seed of pseudo-random numbers. The same seed generate same pseudo-random numbers, so that it's reproducible.
  - The default values for pseudo-random numbers have been changed since R-3.6.0, so that the same code using pseudo-random numbers may generate different results from the previous version R, unless you specify set.seed() properly.

# How to evaluate results

- Considering by-chance stochasticity, simulations should be repeated at least 100 times, usually 1000 times or more for each parameter set. Median values, 95% confidence intervals, minimum and maximum should be calculated.

- Checking the longitudinal (carry-over) effects, selected indicators for each simulation can be drawn as line.
    - Here I recorded changes of total population size, sex ratio, total fertility rate, dependency ratio, and aged proportion.

# Results



Initial population were
4771/5227 for Male/Female

# Results (cont'd)



Theoretical TFR = 1.921