Nonstable Population Relations: Applications to Aging & Gerontology, Family Demography, Labor, Health & Mortality, and Immigration

Michael Lachanski

University of Pennsylvania

Population Studies Center and Department of Sociology

October 19, 2022

Outline

Demographic Methods: Why Do I Care?

- Basic motivation and missing data
- Model populations: Stationary populations
- Live Literatures in Nonstable Population Relations:
  - Chinese fertility (Cai 2008)
  - Centenarians in a developing country (Nepomuceno and Turra 2020)

Derivation of Basic Relations

- Some math
- Visualization

Using "variable-r" Relations for Demographic Estimation

Applications: Family Demography, Health & Mortality, Aging & Gerontology, Immigration, Labor

Live Demo

Demographic Methods (Why? And Their Limitations...)

Demographic Methods

- Demographic methods relate:
  - Population-level processes to individual outcomes - mortality, e.g. individual's "exit" from region A given population-level mortality and emigration rates (no returning)
  - Individual level-risks to population level, e.g. if each individual i faces mortality risk m_i and the initial population size is P_0, how many people will survive to $T > 0, P_T$?
- These computations are of obvious interest to sociologists:
  - Population-level mortality inequality? How long do marriages last? How long do jobs last? How does segregation happen (turnover, changes to entry by a particular race, changes to exit by a particular race)? How have these processes changed over time?
  - Population projection: What will happen in T years if everything stays as it is?

Demographic Methods

- Demographic methods relate:
  - Population-level processes to individual outcomes - mortality, e.g. individual's "exit" from region A given population-level mortality and emigration rates (no returning)
  - Individual level-risks to population level, e.g. if each individual i faces mortality risk m_i and the initial population size is P_0, how many people will survive to $T > 0, P_T$?
- These computations are of obvious interest to sociologists:
  - Population-level mortality inequality? How long do marriages last? How long do jobs last? How does segregation happen (turnover, changes to entry by a particular race, changes to exit by a particular race)? How have these processes changed over time?
  - Population projection: What will happen in T years if everything stays as it is?
  - What do deviations from this trajectory tell us?

Demographic Methods

- Demographic methods relate:
  - Population-level processes to individual outcomes - mortality, e.g. individual's "exit" from region A given population-level mortality and emigration rates (no returning)
  - Individual level-risks to population level, e.g. if each individual i faces mortality risk m_i and the initial population size is P_0, how many people will survive to $T > 0, P_T$?
- These computations are of obvious interest to sociologists:
  - Population-level mortality inequality? How long do marriages last? How long do jobs last? How does segregation happen (turnover, changes to entry by a particular race, changes to exit by a particular race)? How have these processes changed over time?
  - Population projection: What will happen in T years if everything stays as it is?
  - What do deviations from this trajectory tell us?
  - Ans: The rates have changed

Demographic estimation problems: Missing Data

- Sometimes missing key data
- Example 1: need to keep track of age at death to construct p_x survival curve (Preston et al. 2001: 49).
- What if its missing?
- Solution: make some assumptions about population, fill in the gaps.
- We call the population generated by these assumptions a model population

Model populations

- Model population structures give us a set of useful relations (read: formulas, graphics) with a simplicity-realism tradeoff.
- Almost all demographers study the stationary population because it is the simplest model population structure
- Some study stable populations
- Few study nonstable populations

Stationary Populations

- The simplest model population structure, size P

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x
- $\Rightarrow B = D \Rightarrow CBR = CDR$
- Overall population growth rate r, what is it?

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x
- $\Rightarrow B = D \Rightarrow CBR = CDR$
- Overall population growth rate r, what is it?
- Overall population growth rate $r = 0 \Leftarrow B = D$

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x
- $\Rightarrow B = D \Rightarrow CBR = CDR$
- Overall population growth rate r, what is it?
- Overall population growth rate $r = 0 \Leftarrow B = D$
- Thought to be a good model for islands in contemporary period, historical populations, with approximately 0 long-run growth rate

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x
- $\Rightarrow B = D \Rightarrow CBR = CDR$
- Overall population growth rate r, what is it?
- Overall population growth rate $r = 0 \Leftarrow B = D$
- Thought to be a good model for islands in contemporary period, historical populations, with approximately 0 long-run growth rate
- Used in policy analyses over 100 years ago (Ayres 1909).

Stationary Populations

- The simplest model population structure, size P
- B births every year, $CBR = B/P = b$
- D deaths every year, $CDR = D/P = d$
- No migration
- Fixed survival probability from birth to age x: p_x
- $\Rightarrow B = D \Rightarrow CBR = CDR$
- Overall population growth rate r, what is it?
- Overall population growth rate $r = 0 \Leftarrow B = D$
- Thought to be a good model for islands in contemporary period, historical populations, with approximately 0 long-run growth rate
- Used in policy analyses over 100 years ago (Ayres 1909).
- Now of theoretical interest in its own right, e.g. Wrigley-Field and Feehan (2022) in *Demography*

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x + n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...
- Let ${}_nN_x$ be number of people alive between ages x and $x + n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be ${}_nN_x$.

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...
- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be ${}_nN_x$.
- For the population not to grow: $B \uparrow \Rightarrow D \uparrow$

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...
- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be ${}_nN_x$.
- For the population not to grow: $B \uparrow \Rightarrow D \uparrow$
- Let the $e_0 = \sum_x p_x \cdot x$ be the life expectancy of the population

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...
- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be ${}_nN_x$.
- For the population not to grow: $B \uparrow \Rightarrow D \uparrow$
- Let the $e_0 = \sum_x p_x \cdot x$ be the life expectancy of the population
- We have: $b \cdot e_0 = P$

Stationary Population Implications

- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be...
- Let ${}_nN_x$ be number of people alive between ages x and $x+n$ today. For a stationary population, the number of people alive in that age category T years from now ${}_nN_x^T$ will be ${}_nN_x$.
- For the population not to grow: $B \uparrow \Rightarrow D \uparrow$
- Let the $e_0 = \sum_x p_x \cdot x$ be the life expectancy of the population
- We have: $b \cdot e_0 = P$
- Less intuitive: the life expectancy of a person in a stationary population $e_0 = 1/b = 1/d$
- $B \uparrow \Rightarrow D \uparrow \Rightarrow e_0 \downarrow$, Malthusian implications

Stationary Population Applications

- Stationary population relations naturally suggest estimation strategies
- Suppose we want to estimate p_x
- $p_x \approx \frac{{}_nN_x}{B}$
- Let ${}_nd_x$ be the number dying between ages x and $x + n$
- Suppose we want to estimate the probability that a newborn in a stationary population dies between ages x and $x + n$, ${}_nq_x$, x could be 0
- For a stationary population, ${}_nq_x = {}_nd_x/B$
- These estimation strategies are often used by paleodemographers, archaeologists, anthropologists to study of small, closed, approximately zero-growth populations

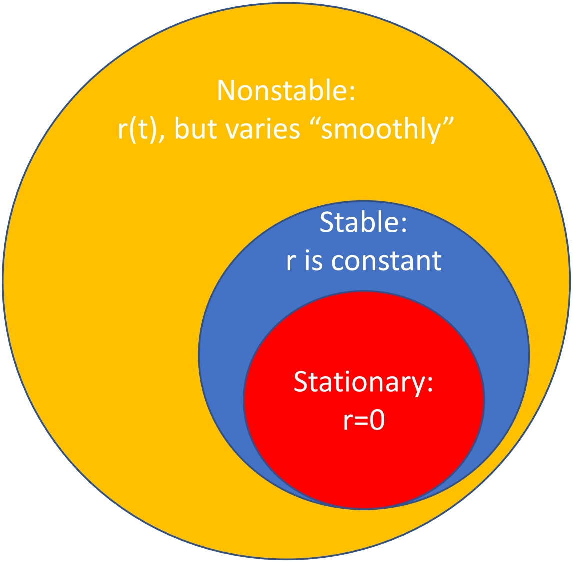
Stationary Population Problems: Assumption Violation

- Problem: every assumption violated for a typical contemporary human population, e.g. migration, birth rate changes, p_x changes and birth rates and mortality do not have the inverse relationship predicted by the stationary population model
- Assumption of $B(t) = B$ almost always false. In general, population growth rates clearly not 0.
- Partial solution: stable population, let birth rates grow at rate r; $B(t) = Be^{rt}$, fix p_x
- Historical context: Stable population models developed by Lotka (1939) under restrictive conditions
- Estimation strategies based on stable populations developed until around Coale, Demeny, and Vaughn (1983)

Nonstable Population Theory

- Little improvement on stable population methods since 1980s
- Why? Stable population theory essentially complete and fixed growth rate for $B(t)$ is objectionable.
- Solution: allow arbitrary birth rate changes b_t, fix p_x
- It turns out that you can still derive useful "nonstable population relations"
- In a stable population $r(a) = r$.
- In a nonstable population, we have: age-specific growth rates that may vary over time $r(a, t)$, "variable-r"s
- Nonstable population relations use these quantities to derive fertility proxies, p_x, and thus e_0
- Also turn out to tell us useful things about immigration.

Model Populations in Perspective



Recent Controversies: Chinese fertility I

AFRICA AFGHANISTAN DEMOCRACY MEDIA COVID-19 CRYPTO

EN English

Politics

Leaked Data Show China's Population Is Shrinking Fast

Jul 27, 2022 | **YI FUXIAN**

Because China has always massaged its demographic figures and cracked down on anyone who challenges the official line, there are endless debates about the true size and growth trajectory of the country's population. But a recent, large-scale data breach offers some sorely needed clarity.

MADISON, WISCONSIN – Even though everyone knows that China's official demographic figures are systematically overestimated, the authorities have consistently cracked down on anyone who questions the data. For example, my book *Big Country with an Empty Nest* was quickly banned when it appeared in 2007, because it voiced concerns about China's one-child policy and predicted that the Chinese population would begin to shrink in 2017, not in 2033-34,

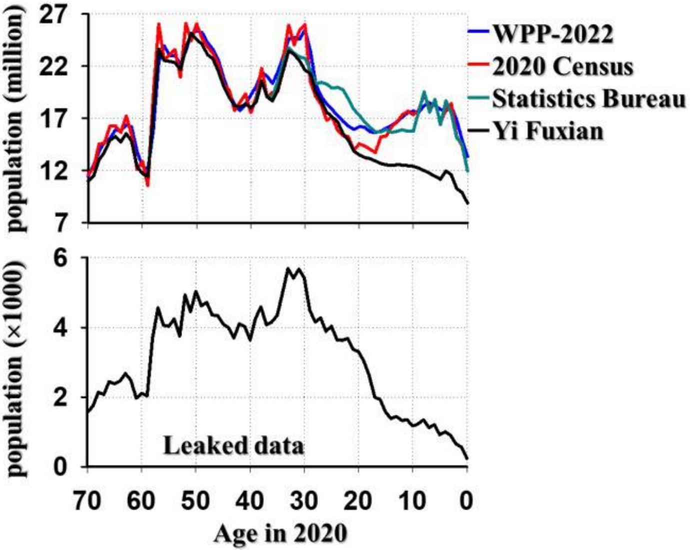
FEATURED

1 Germany's Emerging War Economy
Oct 6, 2022 | **DALIA MARIN**

2 The Global Echoes of a British Near-Collapse
Oct 11, 2022 | **ANATOLE KALETSKY**

Penn University of Pennsylvania

Recent Controversies: Chinese fertility II



Recent Controversies: Chinese Fertility III, Cai (2008)

Table 1. Computing NRR Using the Variable-r Method, China 1990–2000

Age x	${}_5P_x$ (1990) million (1)	${}_5P_x$ (2000) million (2)	${}_5B_x$ million (3)	${}_5r_x$ (4)	e^{S_x} (5)	${}_5\text{v}_x$ (6)	${}_5\text{v}_x \cdot e^{S_x}$ (7)
0	55.4	31.3		–0.0551	0.8712		
5	47.7	41.8		–0.0127	0.7354		
10	47.0	60.1		0.0236	0.7558		
15	58.5	50.2	3.7	–0.0149	0.7724	0.0222	0.017
20	61.5	46.6	74.5	–0.0268	0.6959	0.4502	0.313
25	50.8	57.4	69.2	0.0119	0.6703	0.4185	0.281
30	40.2	62.0	14.4	0.0419	0.7668	0.0868	0.067
35	41.8	53.0	2.8	0.0230	0.9020	0.0171	0.015
40	30.4	39.0	0.6	0.0242	1.0151	0.0039	0.004
45	23.2	41.6	0.2	0.0563	1.2414	0.0012	0.001
Sum			165.4			1.0000	
NRR							0.698

Recent Controversies: Centenarians I, S.J. Newman (2020)

Fraud And Poor Record-Keeping Are What Account For Many of Earth's 'Oldest People,' Study Says

By Natasha Ishak | Checked By John Kuroski
Published August 16, 2019

These researchers noticed that the areas of the world with the highest concentrations of reported 110-year-olds actually had poor average lifespans and healthcare. It just didn't add up.

Toru Yamanaka/AFP/Getty Images

A troupe of elderly women performs in Okinawa, Japan, one of the reported homes of a large number of the world's supercentenarians.

Recent Controversies: Centenarians II, Nepomuceno and Turra (2020)

FIGURE 2 Ratio of the number of centenarians recorded in the census and indirectly estimated according to different parameters: Brazil, men and women, 1900–2000

NOTE: The dotted line corresponds to an equal number of centenarians recorded in census and indirectly estimated.
SOURCE: Author's calculations.

The Variable-r Equations and Derivations

The Fundamental Nonstable Population Equation

- Denote $N(x, t)$ as the population reaching age x at time t
- ℓ_x be the proportion of the population that survives to x, born at 0
- $r(a, t)$ is the age-a to $a + da$-specific growth rate at time t to $t + dt$.
- ℓ_x/ℓ_y = probability of surviving from age y to age x during the interval t to $t + dt$

The Fundamental Nonstable Population Equation

- Denote $N(x,t)$ as the population reaching age x at time t
- ℓ_x be the proportion of the population that survives to x, born at 0
- $r(a,t)$ is the age-a to $a+da$-specific growth rate at time t to $t+dt$.
- $\ell_x/\ell_y =$ probability of surviving from age y to age x during the interval t to $t+dt$

$$N(x,t) \underset{\text{Bennett and Horiuchi 1981}}{=} N(y,t) \cdot \underbrace{\frac{\ell_x}{\ell_y}}_{\substack{\text{Probability of} \\ \text{surviving from } y \text{ to } x}} \cdot \underbrace{e^{-\int_y^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from y to x}}}$$

- Preston and Coale (1982) discovered variable-r relations

Unstable vs. Stable Population Equations I

- Let $c(x)$ be the proportion of the population aged x to $x + dx$
- Unstable:

$$c(x) = b \cdot \underbrace{\ell_x/\ell_0}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x \\ = p_x}} \cdot \underbrace{e^{-\int_0^x r(a)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

- Stable:

$$c(x) = b \cdot \underbrace{p_x}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-rx}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x} \\ \text{constant } r}}$$

Unstable vs. Stable Population Equations II

- Computing birthrates
- Unstable:

$$b = \frac{1}{\int_0^\infty e^{-\int_0^x r(a)da} p(x)dx}$$

- Stable:

$$b = \frac{1}{\int_0^\infty e^{-rx} p(x)dx}$$

Unstable vs. Stable Population Equations III

- Let $m(x)$ be the age-specific maternity rate (female births for mothers aged x to $x + dx$). Let the fertile ages range from age α to age β
- Unstable:

$$1 = \int_{\alpha}^{\beta} e^{-\int_0^x r(a)da} p(x)m(x)dx$$

- Stable:

$$1 = \int_{\alpha}^{\beta} e^{-rx} p(x)m(x)dx$$

Two Important Lessons

- The nonstable population equations I have presented hold for any population in which $r(a)$ varies smoothly
- Essentially every social phenomena
- Implementing these for actual estimation requires non-trivial discretization
- Q: How can all contemporary demographic functions be related to one another in such a simple way?

Two Important Lessons

- The nonstable population equations I have presented hold for any population in which $r(a)$ varies smoothly
- Essentially every social phenomena
- Implementing these for actual estimation requires non-trivial discretization
- Q: How can all contemporary demographic functions be related to one another in such a simple way?
- Lesson 1: Age distributions are the products of historical patterns. All necessary history is contained in the $r(a, t)$ function

Two Important Lessons

- The nonstable population equations I have presented hold for any population in which $r(a)$ varies smoothly
- Essentially every social phenomena
- Implementing these for actual estimation requires non-trivial discretization
- Q: How can all contemporary demographic functions be related to one another in such a simple way?
- Lesson 1: Age distributions are the products of historical patterns. All necessary history is contained in the $r(a, t)$ function
- Lesson 2: To reestablish the many useful relations in a stationary populations, one can simply apply a "growth correction". (But migration rates may also be needed.)

Why Are Age-Specific Growth Rates Powerful?

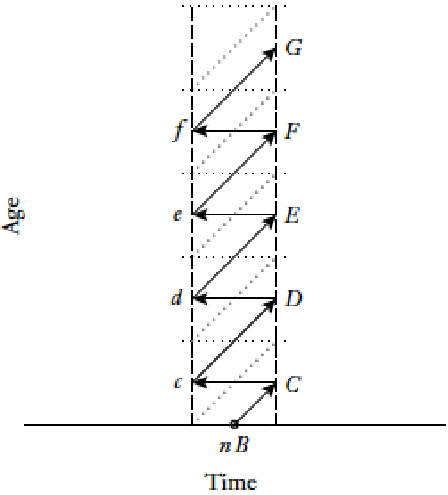


Figure 5.3 Zigzag paths with Variable r

Variable-r for Demographic Estimation

Application of variable-r relations for indirect estimation

- How to estimate a survival curve $p(x)$ when missing ${}_nD_x$
- Preston and Bennett (1983): estimation of intercensal life table using age distributions

$$N(x)e^{\int_0^x r(a)da} = Bp(x)$$

$$\frac{N(x)}{B}e^{\int_0^x r(a)da} = p(x)$$

- Notice that this is essentially the stationary population approximation with the growth correction mentioned earlier $p_x \approx \frac{{}_nN_x}{B}$

Preston and Bennett (1983)

Table 1. *Application of census-based method to Swedish females, 1966–1970*

Start of age interval (x)	Mean person years lived in interval, 1966–70 ${}_5N_x$	Average annual growth rate in interval 1966–70 ${}_5r_x$	Sum of age-specific growth rates from age 5 to mid-point of interval S_x	Stationary population in interval ${}_5L_x = {}_5N_x e^{S_x}$	Stationary population above age x, T_x	Number surviving to age x in stationary population, l_x	Estimated life expectancy at age x, e_x	Life expectancy from official Swedish life table	Ratio, estimated to official e_x
0	277,957	−0.001097	0.00274	278,721	4,300,267	—	—	—	—
5	267,024	0.015746	0.03937	277,745	4,021,546	55,647	72.27	72.52	0.997
10	259,265	−0.005044	0.06612	276,987	3,743,801	55,473	67.49	67.64	0.998
15	285,619	−0.033251	−0.02962	277,284	3,466,815	55,427	62.55	62.71	0.997
20	309,826	0.003975	−0.10281	279,557	3,189,530	55,684	57.28	57.83	0.990
25	267,775	0.048595	0.02862	275,550	2,909,973	55,511	52.42	52.96	0.990
30	226,859	0.014550	0.18648	273,366	2,634,428	54,892	47.99	48.10	0.998
35	226,394	−0.016020	0.18281	271,806	2,361,057	54,517	43.31	43.27	1.001
40	249,171	−0.025007	0.08024	269,990	2,089,252	54,180	38.56	38.49	1.002
45	263,172	−0.001657	0.01358	266,771	1,819,262	53,676	33.89	33.80	1.003
50	260,557	0.001220	0.00639	262,228	1,552,491	52,900	29.35	29.21	1.005
55	254,801	0.000380	0.00429	255,897	1,290,263	51,813	24.90	24.75	1.006
60	235,205	0.016034	0.04533	246,111	1,034,366	50,201	20.60	20.43	1.009
65	201,225	0.020365	0.13632	230,614	788,254	47,673	16.53	16.35	1.011
70	160,846	0.022252	0.24287	205,062	557,640	43,568	12.80	12.58	1.017
75	114,792	0.025615	0.36254	164,953	352,578	37,002	9.53	9.32	1.022
80	66,452	0.033675	0.51076	187,625	187,625	27,570	6.81	6.69	1.017
85	37,059	—	—	—	—	—	—	—	—

Source: Statistiska Centralbyrån, 1966–72 and 1974.

Preston and Bennett 1983

Formulas for Certain Functions in Model Populations

Function	Notation	Formula for: Stationary Population	Formula for: Stable Population	Formula for: Any Population
Proportionate age distribution	$c(a)$	$bp(a)$	$be^{-ra}p(a)$	$be^{-\int_0^a r(x)dx}p(a)$
Ratio of population at two ages	$\frac{c(a+n)}{c(a)}$	nPa	$e^{-rn}nPa$	$e^{-\int_a^{a+n} r(x)dx}nPa$
Life expectancy at birth	$e_o^o = \int_0^\infty p(a)da$	$\frac{\int_0^\infty c(a)da}{b} = \frac{1}{b}$	$\frac{\int_0^\infty c(a)e^{ra}da}{b}$	$\frac{\int_0^\infty c(a)e^{\int_0^a r(x)dx}}{b}$
Birth rate	b	$\frac{1}{\int_0^\infty p(a)da}$	$\frac{1}{\int_0^\infty p(a)e^{-ra}da}$	$\frac{1}{\int_0^\infty p(a)e^{-\int_0^a r(x)dx}}$
Proportionate age distribution of mothers at childbirth	$v(a)$	$p(a)m(a)$	$p(a)m(a)e^{-ra}$	$p(a)m(a)e^{-\int_0^a r(x)dx}$
Net reproduction rate	$NRR = \int_\alpha^\beta p(a)m(a)da$	$\int_\alpha^\beta v(a)da = 1$	$\int_\alpha^\beta v(a)e^{ra}da$	$\int_\alpha^\beta v(a)e^{\int_0^a r(x)dx}da$
Expected years of life to be spent in state G with incidence at age a g(a)	$G'_L = \int_0^\infty g(a)p(a)da$	$\frac{\int_0^\infty g(a)c(a)da}{b}$	$\frac{\int_0^\infty g(a)c(a)e^{ra}da}{b}$	$\frac{\int_0^\infty g(a)c(a)e^{\int_0^a r(x)}da}{b}$
Number of persons at age a^* in terms of deaths above age a^*	$N(a^*)$	$\int_{a^*}^\infty D(a)da$	$\int_{a^*}^\infty D(a)e^{r(a-a^*)}da$	$\int_{a^*}^\infty D(a)e^{\int_{a^*}^a r(x)dx}da$
Number of persons at age a^* in terms of deaths below age a^*	$N(a^*)$	$N(0) - \int_0^{a^*} D(a)da$	$e^{ra^*}\left[N(0)-\int_0^{a^*}D(a)e^{ra}da\right]$	$e^{\int_0^{a^*}r(x)dx}\left[N(0)-\int_0^{a^*}D(a)e^{\int_0^a r(x)dx}da\right]$
Probability of survival from a^* to a^*+n in terms of deaths	nPa^*	$\frac{\int_{a^*+n}^\infty D(a)da}{\int_{a^*}^\infty D(a)da}$	$\frac{\int_{a^*+n}^\infty D(a)e^{r(a-a^*)}da}{\int_{a^*}^\infty D(a)e^{r(a-a^*)}da}$	$\frac{\int_{a^*+n}^\infty D(a)e^{\int_{a^*}^a r(x)dx}}{\int_{a^*}^\infty D(a)e^{\int_0^a r(x)dx}da}$

Applications: Family Demography

Basic Idea for Applications: X as a Population

The tools used for thinking about human populations can be applied to the population of marital relations.

- Target: first transition from married state
- "birth" $\rightarrow$ marriage; "death" $\rightarrow$ separation, death;
- "Life expectancy" $\rightarrow$ expected years married
- "Birth cohort" $\rightarrow$ year of marriage

Basic Idea for Applications: X as a Population

The tools used for thinking about human populations can be applied to the population of marital relations.

- Target: first transition from married state
- "birth" $\rightarrow$ marriage; "death" $\rightarrow$ separation, death;
- "Life expectancy" $\rightarrow$ expected years married
- "Birth cohort" $\rightarrow$ year of marriage
- "Causes of death" $\rightarrow$ reason for separation: divorce, death
- Demographic methods yield information about synthetic cohorts: what **would** happen to marriages if all sources of decrement were held constant.
- Demographic decomposition can be used to study geographic, race, and class inequalities in the "life" of a marriage.

U.S. Marriage Survival in the 1970s

Example: United States, females, life table for first marriages corresponding to divorce and mortality rates in 1975–80; $l_0 = 1{,}534$

Duration x	${}_nN_x(1975)$	${}_nN_x(1980)$	${}_nN_x^*$	${}_nr_x$	S_x	${}_nL_x$	l_x	T_x	e_x^o
0	1,428	1,395	1,411	−0.00468	−0.00234	1,408	1,534	43,216	28.17
1	1,448	1,432	1,440	−0.00222	−0.00579	1,432	1,420	41,808	29.44
2	1,457	1,411	1,434	−0.00642	−0.01011	1,419	1,426	40,377	28.32
3	1,436	1,246	1,338	−0.02838	−0.02751	1,301	1,360	38,957	28.64
4	1,315	1,288	1,301	−0.00415	−0.04377	1,246	1,274	37,656	29.57
5	1,462	1,242	1,348	−0.03262	−0.06216	1,266	1,256	36,410	28.99
6	1,383	1,180	1,277	−0.03175	−0.09434	1,162	1,214	35,144	28.94
7	1,258	1,317	1,287	0.00917	−0.10563	1,158	1,160	33,981	29.29
8	1,104	1,217	1,159	0.01949	−0.09130	1,058	1,108	32,823	29.62
9	1,020	1,161	1,088	0.02590	−0.06861	1,016	1,037	31,765	30.63
10	1,039	1,143	1,090	0.01908	−0.04612	1,041	1,028	30,749	29.90
11	975	1,118	1,044	0.02737	−0.02290	1,020	1,031	29,709	28.83
12	943	1,071	1,005	0.02546	0.00352	1,009	1,014	28,688	28.28
13	943	908	925	−0.00756	0.01246	937	973	27,680	28.46
14	954	1,022	987	0.01377	0.01557	1,003	970	26,743	27.57
15	4,477	4,477	4,477	0.00000	0.02245	4,579	988	25,740	26.04
20	4,424	4,200	4,311	−0.01039	−0.00353	4,295	887	21,161	23.85
25	4,712	3,984	4,333	−0.03357	−0.11342	3,868	816	16,866	20.66
30	6,475	7,553	6,993	0.03080	−0.04334	6,697	767	12,998	16.95
40	3,277	3,621	3,445	0.01996	0.21048	4,252	547	6,301	11.51
50	1,351	1,343	1,347	−0.00119	0.30436	1,826	304	2,049	6.74
60	170	167	168	−0.00356	0.28062	223	102	223	2.18

Probability of a Marriage Ending in Divorce

Box 8.3 *Computation of the Probability that a Marriage will End in Divorce*

${}_n r_x$ = duration-specific growth rate in number of intact marriages (any order) between t_1 and t_2

${}_n D_x^i$ = total number of divorces by duration of marriage between t_1 and t_2

S_x = sum of growth rates from duration 0 to midpoint of interval

$\frac{{}_n d_x^i}{l_o} = \frac{{}_n D_x^i \cdot e^{S_x}}{N(0)}$ = probability that a marriage just contracted will end in divorce between durations x and $x+n$.

$N(0)$ is the total number of marriages between t_1 and t_2

$\sum_{x=0,n}^{\infty} \frac{{}_n d_x^i}{l_o}$ = probability that a marriage will eventually end in divorce

Example: United States, 1975–80; $N(0)$ = 11,218,240

Duration x	${}_n r_x$	${}_n D_x^i$	S_x	$\frac{{}_n d_x^i}{l_o}$
0	0.00603	251,888	0.00302	0.0225
1	0.01270	458,995	0.01238	0.0414
2	0.00558	506,574	0.02152	0.0461
3	−0.01319	506,574	0.01772	0.0460
4	0.01300	464,592	0.01762	0.0422
5	−0.02505	405,819	0.01159	0.0366
6	−0.01454	352,642	−0.00820	0.0312
7	0.01799	323,460	−0.00648	0.0286
8	0.02607	265,881	0.01555	0.0241
9	0.01948	229,497	0.03833	0.0213
10	0.01649	766,858	0.08930	0.0747
15	−0.00164	442,202	0.12642	0.0447
20	−0.01092	299,466	0.09502	0.0294
25	−0.03556	179,120	−0.02118	0.0156
30	0.02494	156,730	0.13932	0.0161
Sum				0.5205

Preston's Handwritten Notes on Using Variable-r to Compute Age-specific Risk of Marriage I

Proportion leaving a state as a function of duration of time spent in the state: First marriage from the single state

$$N(a,t) = B\, e^{-\int_0^a r(x,t)\,dx}\, p(a,t)$$

$$S(a,t) = B\, e^{-\int_0^a r_S(x,t)\,dx}\, p_S(a,t) \cdot \underbrace{p_M^*(a,t)}_{e^{-\int_0^a \mu_M(x,t)\,dx}}$$

μ_M: "force of first marriage to the single pop"

$$\pi(a,t) = \frac{S(a,t)}{N(a,t)} = e^{-\int_0^a [r_S(x,t) - r(x,t)]} \cdot \frac{p_S(a,t)}{p(a,t)} \cdot p_M^*(a,t)$$

$$r_\pi(a,t) = \frac{d \ln \pi(a,t)}{dt} = \frac{d \ln \left[\frac{S(a,t)}{N(a,t)}\right]}{dt} = r_S(a,t) - r(a,t)$$

Preston's Handwritten Notes on Using Variable-r to Compute Age-specific Risk of Marriage II

So $\pi(a,t) = e^{-\int_0^a r_\pi(x,t)\,dx} \, \frac{P_s(a,t)}{P(a,t)} \cdot P^*_M(a,t)$

↑ differential mortality btw single + total pop

Assume = 1. Then

$$\underbrace{\pi(a,t)}_{\text{proportion single by age}} \; e^{\underbrace{-\int_0^a r_\pi(x,t)\,dt}_{\text{growth rate of proportion single by age}}} = \underbrace{P^*_M(a,t)}_{\text{life table of entry into marriage from single state}}$$

Expected Single Years by Age 50

Table 3.3. Expected years lived in the single state by age 50, percent single at age 50, and singulate mean age at marriage in gross nuptiality tables computed from cross-sectional observations. Selected countries

		Males			Females		
Country	Year	Years single	% Single (age 50)	SMAM	Years single	% Single (age 50)	SMAM
Kenya	1969	26.87	6.15	25.35	20.09	2.90	19.20
	1979	26.75	5.05	25.51	20.92	2.18	20.28
	1969–79	26.74	4.46	25.66	21.23	3.58	20.16
Liberia	1962	27.79	6.40	26.27	18.58	1.96	17.96
	1974	28.20	6.80	26.61	20.13	2.79	19.27
	1962–74	28.30	7.09	26.64	20.63	6.96	18.44
Egypt	1960	26.71	2.05	26.22	20.63	1.24	20.26
	1976	27.79	3.96	26.87	22.83	4.18	21.65
	1960–76	27.69	5.04	26.50	23.00	11.71	19.41
Jordan	1961	25.64	3.50	24.75	21.21	2.79	20.39
	1976	26.38	0.87	26.18	21.97	1.57	21.53
	1961–76	26.29	0.61	26.14	21.90	1.58	21.45
Kuwait	1965	27.82	4.69	26.72	19.81	2.47	19.05
	1975	27.17	3.34	26.38	21.41	3.16	20.48
	1965–75	27.12	1.99	26.65	22.02	5.57	20.37
Libya	1954	27.43	3.53	26.61	19.59	1.32	19.18
	1973	24.94	1.35	24.60	18.88	0.49	18.73
	1954–73	25.31	0.56	25.18	19.01	0.25	18.93
Morocco	1960	24.79	2.56	24.12	18.09	1.66	17.55
	1971	25.71	2.94	24.98	19.89	2.48	19.12
	1960–71	25.85	4.31	24.76	20.20	6.93	17.98
Syria	1960	26.91	4.38	25.85	20.99	2.70	20.19
	1970	26.56	2.71	25.91	21.41	2.52	20.67
	1960–70	26.46	1.47	26.21	21.51	2.99	20.63
Tunisia	1966	27.80	3.44	27.01	21.32	1.54	20.87
	1975	27.80	2.89	27.14	22.99	1.54	22.57
	1966–75	27.83	2.01	27.37	23.42	3.09	22.57

Applications: Health and Mortality

Variable-r Relations in Epidemiology: I

growth rate of # people diagnosed a years earlier

(1) $N(x) = N(0)\, e^{-\int_0^x r(a)\,da} \; p(x)$

- $N(x)$: # people with cancer diagnosed x years earlier
- $N(0)$: # of new cases of cancer
- $p(x)$: proportion surviving x years after diagnosis

(2) $p(x) = \dfrac{N(x)}{N(0)}\, e^{\int_0^x r(a)\,da}$ — probability of surviving x years

(3) $e_0^\circ = \int_0^\infty p(x)\,dx = \int_0^\infty \dfrac{N(x)}{N(0)}\, e^{\int_0^x r(a)\,da}$ — expected length of life after diagnosis

Multiply both sides of (1) by $\mu^i(x)$, the death rate from cancer for someone diagnosed x years ago

Variable-r Relations in Epidemiology: II

(4) $N(x)\mu^i(x) = N(0)\, e^{-\int_0^x r(a)\,da}\; p(x)\,\mu^i(x)$

Number of cancer deaths in population at duration x since diagnosis

probability of dying from cancer x years after diagnosis — sums to case fatality rate, CFR^i

(5) $CFR^i = \int_0^\infty p(x)\,\mu^i(x)\,dx = \int_0^\infty \frac{D^i(x)}{N(0)}\, e^{\int_0^x r(a)\,da}$

Analogy of (2) + (3) to survival of a marriage, length of employment

Analogy of (4) + (5) to probability of divorce, lifetime probability of marriage, probability of being fired

Preston, America J Epidemiology 1987.

Suggestion

- These relations relating expected survival time to initial diagnosis could not be tested when first proposed by Preston (1987) *American Journal of Epidemiology*
- But now the **National Health Interview Survey** records years since diagnosis and cause of death
- National Cancer Database Records growth of cancer population, type of cancer, and years until death.

Applications: Aging and Gerontology

Methods for Studying Sources of Change in Age Structure

- *Stable population analysis:* comparative statics. Comparing populations "at least", in equilibrium, with different levels of fertility, mortality [migration]
- *Counter-factual projection.* What would happen over some defined period if mortality or fertility [migration] pursued one path or another (e.g., what if fertility hadn't changed?)
- *Variable-r analysis.* Why is the population aging NOW? Decomposes aging into the history of births, mortality, and migration?

Murphy (2017): Demographic Determinants of European Aging Since 1850 I

Demographic Sources of Aging

The growth rate of the population at a particular age can be traced in a simple fashion to processes of birth, death, and migration. Designate the number of persons aged x last birthday at time t as N(x,t).

$$N(x,t) \quad = B(t-x)\,p(x,t-x)\,j(x,t-x)\ , \qquad \text{where}$$

$B(t-x)$ = Number of births in the population in the year ending at time t - x

$p(x,t-x)$ = Proportion of birth cohort born in year ending at t-x that survived to time t

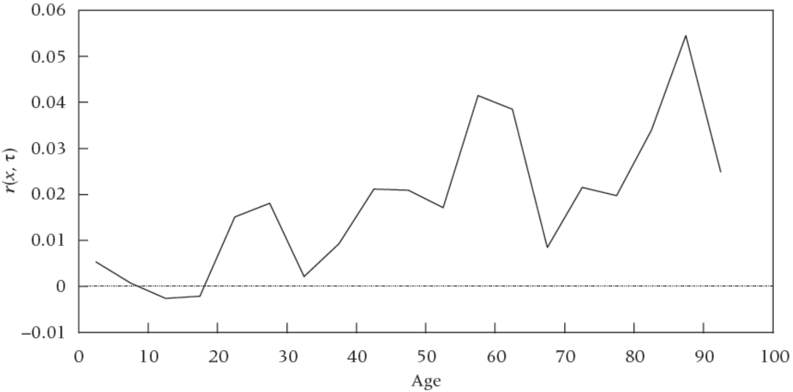
$j(x,t-x)$ = Factor by which birth cohort born in year ending at t-x changed in size by time t as a result of migration

Then the growth rate in the number of persons age x between times t and $t+1$ can be expressed as

$$r(x,t,t+1) = \ln\frac{B(t-x+1)}{B(t-x)} + \ln\frac{p(x,t-x+1)}{p(x,t-x)} + \ln\frac{j(x,t-x+1)}{j(x,t-x)} \ , \quad \text{or}$$

Murphy (2017): Demographic Determinants of European Aging Since 1850 II

FIGURE 1 Age-specific growth rates for the world, 2005–10



Murphy (2017): Demographic Determinants of European Aging Since 1850 III

TABLE 5 Change in components of overall population aging by time period, overall European series, 1900–2012

Component	1900–1924		1925–1949		1950–1974		1975–2012		1900–2012	
	Years	Percent	Years	Percent	Years	Percent	Years	Percent	Years	Percent
Total	2.09	100.0	2.75	100.0	1.67	100.0	5.94	100.0	12.45	100.0
Survival	0.29	13.8	0.95	34.5	2.33	139.2	4.68	78.8	8.25	66.2
Net migration	−0.56	−26.8	0.41	14.9	0.63	37.7	−0.22	−3.7	0.26	2.1
Birth cohort	2.36	113.0	1.39	50.6	−1.29	−76.9	1.48	24.9	3.95	31.7
of which:										
CBR	*2.21*	*105.8*	*1.36*	*49.4*	*−1.24*	*−74.2*	*0.77*	*12.9*	*3.10*	*24.9*
Total population	*0.15*	*7.1*	*0.03*	*1.2*	*−0.05*	*−2.7*	*0.71*	*12.0*	*0.85*	*6.8*

Penn
University of Pennsylvania

Preston and Stokes (2014): Variable-r Aging Equation

$$r_{x+}(\tau) = \underbrace{b_{x+}(\tau)}_{\text{“Births” above age } x} - \underbrace{d_{x+}(\tau)}_{\text{deaths above age } x} + \underbrace{m_{x+}(\tau)}_{\text{migration above age } x}$$

$$b_{x+}(\tau) = b_0(\tau - x)\,{}_x p_0^{\tau - x}\,{}_x f_0^{\tau - x}$$

- $b_0(\tau - x)$ is the number of births in the population over the interval $\tau - x$ divided by person-years lived above age x during period τ
- ${}_x p_0^{\tau - x}$ is the probability of survival between ages 0 and x for a member of the birth cohort born during the interval $\tau - x$
- ${}_x f_0^{\tau - x}$ is the factor by which the birth cohort's size is modified through migration between ages 0 and x

Applications: NRR & Immigration

Variable-r Relations Give an Alternative Expression for a Common Fertility Measure

$$v(x) = \frac{Be^{-\int_0^x r(a)\,da} p(x)m(x)}{B}$$
$$= e^{-\int_0^x r(a)\,da} p(x)m(x)$$

Rearranging, we have:

$$v(x)e^{\int_0^x r(a)\,da} = p(x)m(x)$$

Integrating both sides of this last expression over the ages of childbearing, α to β, gives the expression for the net reproduction rate, NRR, on the right-hand side:

$$\int_\alpha^\beta v(x)e^{\int_0^x r(a)\,da}\,dx = NRR \qquad \textbf{(8.7)}$$

This rather odd expression shows that the net reproduction rate can be recaptured without any reference to underlying fertility or mortality rates. It is only necessary to observe age-specific growth rates and the proportionate age distribution of mothers at childbirth, two functions that are widely available from censuses and surveys in developing countries.

Cai (2008)

Table 1. Computing NRR Using the Variable-r Method, China 1990–2000

Age x	${}_5P_x$ (1990) million (1)	${}_5P_x$ (2000) million (2)	${}_5B_x$ million (3)	${}_5r_x$ (4)	e^{Sx} (5)	${}_5V_x$ (6)	${}_5V_x \cdot e^{Sx}$ (7)
0	55.4	31.3		–0.0551	0.8712		
5	47.7	41.8		–0.0127	0.7354		
10	47.0	60.1		0.0236	0.7558		
15	58.5	50.2	3.7	–0.0149	0.7724	0.0222	0.017
20	61.5	46.6	74.5	–0.0268	0.6959	0.4502	0.313
25	50.8	57.4	69.2	0.0119	0.6703	0.4185	0.281
30	40.2	62.0	14.4	0.0419	0.7668	0.0868	0.067
35	41.8	53.0	2.8	0.0230	0.9020	0.0171	0.015
40	30.4	39.0	0.6	0.0242	1.0151	0.0039	0.004
45	23.2	41.6	0.2	0.0563	1.2414	0.0012	0.001
Sum			165.4			1.0000	
NRR							0.698

Preston & Wang (2007) I: Immigration & Pop. Growth

$$N(a) = N(0)e^{-\int_0^a r(x)} p^*(a)p(a) \quad (3)$$

Multiplying both sides of (3) by $m(a)$, the rate of bearing female children by women aged a at time t, gives the number of births to women aged a:

$$B(a) = N(a)m(a) = N(0)e^{-\int_0^a r(x)dx} p^*(a)p(a)m(a) \quad (4)$$

The frequency distribution of mothers' ages at childbearing is

$$v(a) = \frac{N(a)m(a)}{\int_0^\infty N(a)m(a)da} = e^{-\int_0^a r(x)dx} p^*(a)p(a)m(a) \quad (5)$$

(substituting from (4) and observing that $\int_0^\infty N(a)m(a)da$, total births at time t, can equally be written as $N(0)$). Hence

$$p^*(a)p(a)m(a) = v(a)e^{\int_0^a r(x)dx} \quad (6)$$

The intrinsic growth rate of a population is the value of r^* that satisfies

$$\int_0^\infty e^{-r^*a} p^*(a)p(a)m(a)da = 1 \quad (7)$$

Substituting (6) into (7) gives

$$\int_0^\infty e^{-r^*a} v(a)e^{\int_0^a r(x)dx} da = 1 \quad (8)$$

Preston & Wang (2007) II: Immigration & Pop. Growth

Using (8), the value of the intrinsic growth rate of an open population can be calculated from observed age-specific growth rates and the age distribution of mothers at childbirth. Equation (8) is a new formula that should find wide application.

The formula for the net reproduction rate in the presence of migration does not require any derivation, just a recognition that the effect of migration at a particular age, expressed through $p^*(a)$, is directly analogous to the effect of mortality expressed through $p(a)$. The net reproduction rate in the presence of migration, NRR*, is thus

$$\mathrm{NRR}^* = \int_0^\infty p(a)p^*(a)m(a)da \tag{9}$$

Substituting (6) into (9) gives

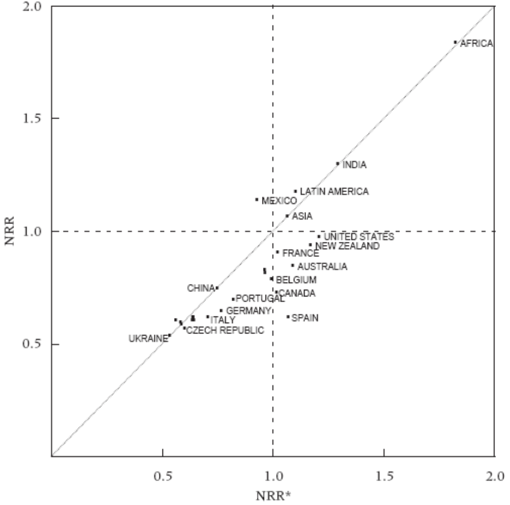
$$\mathrm{NRR}^* = \int_0^\infty v(a)e^{\int_0^a r(x)dx} da \tag{10}$$

As is the case when intrinsic growth rates and net reproduction rates refer to populations closed to migration, r^* must be positive when NRR* is greater than one, zero when NRR* = 1.0, and negative when NRR* is less than one.

To implement (8) and (10), we use discrete approximations in 5-year intervals and "locate" the observation at the midpoint of an age interval.

Preston & Wang (2007) III: Immigration & Pop. Growth

FIGURE 1 Net reproduction rate in the presence of migration (NRR*) and in the absence of migration (NRR), 2000–05



Applications: Labor Demography

Basic Idea: Jobs as a Population

The tools use for thinking about human populations can be applied to the population of employer-employee relations.

- Target: first transition from current job.
- "birth" $\rightarrow$ hire; "death" $\rightarrow$ separation;
- "Life expectancy" $\rightarrow$ job tenure expectancy
- "Birth cohort" $\rightarrow$ job vintage/hiring year

Basic Idea: Jobs as a Population

The tools use for thinking about human populations can be applied to the population of employer-employee relations.

- Target: first transition from current job.
- "birth" $\rightarrow$ hire; "death" $\rightarrow$ separation;
- "Life expectancy" $\rightarrow$ job tenure expectancy
- "Birth cohort" $\rightarrow$ job vintage/hiring year
- "Causes of death" $\rightarrow$ reason for separation: layoff, firing, quitting, mortality
- "Age-specific mortality rate" $\rightarrow$ tenure-specific separation rate
- Demographic methods yield information about synthetic cohorts: what **would** happen to jobs if all sources of decrement were held constant.
- Demographic standardization and decomposition can be used to study sex, race, and other labor market inequalities in the "life" of a job.

Live Demo - Go to R

U.S. Marriage Survival in the 1970s - Again

Example: United States, females, life table for first marriages corresponding to divorce and mortality rates in 1975–80; $l_0 = 1{,}534$

Duration x	${}_nN_x(1975)$	${}_nN_x(1980)$	${}_nN_x^*$	${}_nr_x$	S_x	${}_nL_x$	l_x	T_x	e_x^o
0	1,428	1,395	1,411	−0.00468	−0.00234	1,408	1,534	43,216	28.17
1	1,448	1,432	1,440	−0.00222	−0.00579	1,432	1,420	41,808	29.44
2	1,457	1,411	1,434	−0.00642	−0.01011	1,419	1,426	40,377	28.32
3	1,436	1,246	1,338	−0.02838	−0.02751	1,301	1,360	38,957	28.64
4	1,315	1,288	1,301	−0.00415	−0.04377	1,246	1,274	37,656	29.57
5	1,462	1,242	1,348	−0.03262	−0.06216	1,266	1,256	36,410	28.99
6	1,383	1,180	1,277	−0.03175	−0.09434	1,162	1,214	35,144	28.94
7	1,258	1,317	1,287	0.00917	−0.10563	1,158	1,160	33,981	29.29
8	1,104	1,217	1,159	0.01949	−0.09130	1,058	1,108	32,823	29.62
9	1,020	1,161	1,088	0.02590	−0.06861	1,016	1,037	31,765	30.63
10	1,039	1,143	1,090	0.01908	−0.04612	1,041	1,028	30,749	29.90
11	975	1,118	1,044	0.02737	−0.02290	1,020	1,031	29,709	28.83
12	943	1,071	1,005	0.02546	0.00352	1,009	1,014	28,688	28.28
13	943	908	925	−0.00756	0.01246	937	973	27,680	28.46
14	954	1,022	987	0.01377	0.01557	1,003	970	26,743	27.57
15	4,477	4,477	4,477	0.00000	0.02245	4,579	988	25,740	26.04
20	4,424	4,200	4,311	−0.01039	−0.00353	4,295	887	21,161	23.85
25	4,712	3,984	4,333	−0.03357	−0.11342	3,868	816	16,866	20.66
30	6,475	7,553	6,993	0.03080	−0.04334	6,697	767	12,998	16.95
40	3,277	3,621	3,445	0.01996	0.21048	4,252	547	6,301	11.51
50	1,351	1,343	1,347	−0.00119	0.30436	1,826	304	2,049	6.74
60	170	167	168	−0.00356	0.28062	223	102	223	2.18

Space for Theory

- Preston et al. (2001: 253) variable-r not the most accurate algorithm for constructing life tables from two enumerations

Space for Theory

- Preston et al. (2001: 253) variable-r not the most accurate algorithm for constructing life tables from two enumerations
  - More accurate "intercensal intracohort interpolation" methods discovered by Coale (1984) and refined in Stupp (1988)
  - Turns out to yield MLE for the life table
  - Stupp (1988) only used around three times in thirty years according to Google Scholar, why?
    - Complicated, computational methods with no convergence guarantee (Stupp 1988: 220; 1995: 234)
    - Same weakness as variable-r: errors in variables, e.g. age heaping, often yield incoherent life tables or no convergence (Coale 1984: 203; Stupp 1995: 234)
  - Working with Tim Riffe and Iván Williams to add methods to `DemoTools` by January
  - Hopefully will stimulate theoretical work on indirect estimation, especially connections with formal statistical approaches

Extra Slides

Part I: Current Births & the Population Reaching Age X

$$N(x,t) = N(0,t) \cdot \underbrace{\frac{\ell_x}{\ell_0}}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

Part I: Current Births & the Population Reaching Age X

$$N(x,t) = N(0,t) \cdot \underbrace{\frac{\ell_x}{\ell_0}}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

$$N(x,t) = B(t) \cdot \underbrace{\frac{\ell_x}{\ell_0}}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

Part I: Current Births & the Population Reaching Age X

$$N(x,t) = N(0,t) \cdot \underbrace{\frac{\ell_x}{\ell_0}}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

$$N(x,t) = B(t) \cdot \underbrace{\frac{\ell_x}{\ell_0}}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

$$N(x,t) = B(t) \cdot \underbrace{p(x)}_{\substack{\text{Probability of} \\ \text{surviving from 0 to } x}} \cdot \underbrace{e^{-\int_0^x r(a,t)da}}_{\substack{\text{age-a-specific} \\ \text{growth from 0 to x}}}$$

Part II: Constructing Midpoint Pseudo-Population

$$N(x, t_1) = N(y, t_1) e^{-\int_y^x r(a,t_1) da} {}_{x-y}p_y(t_1)$$

$$N(x, t_2) = N(y, t_1) e^{-\int_y^x r(a,t_2) da} {}_{x-y}p_y(t_2)$$

Part II: Constructing Midpoint Pseudo-Population

$$N(x, t_1) = N(y, t_1) e^{-\int_y^x r(a,t_1) da} {}_{x-y}p_y(t_1)$$

$$N(x, t_2) = N(y, t_1) e^{-\int_y^x r(a,t_2) da} {}_{x-y}p_y(t_2)$$

Take geometric mean:

$$\sqrt{N(x, t_2) \cdot N(x, t_1)} = N^*(x)$$

$$N^*(x) = N^*(y) e^{-\int_y^x \frac{r(a,t_1)+r(a,t_2)}{2} da} {}_{x-y}p_y^*(t_2)$$

Part III: Making Discretization Assumptions

Assumption 1: growth rate changes linearly during the time interval

$$\frac{r(a,t_1) + r(a,t_2)}{2} = \frac{\ln\left(\frac{N(a,t_2)}{N(a,t_1)}\right)}{t_2 - t_1} = \bar{r}(a)$$

Part III: Making Discretization Assumptions

Assumption 1: growth rate changes linearly during the time interval

$$\frac{r(a,t_1)+r(a,t_2)}{2} = \frac{\ln\left(\frac{N(a,t_2)}{N(a,t_1)}\right)}{t_2 - t_1} = \bar{r}(a)$$

Return to the birth equation and combine it with the midpoint population:

$$N(x,t) = B(t)\cdot p(x)\cdot e^{-\int_0^x r(a,t)da}$$

$$N^*(x) = \underbrace{B^*}_{\substack{\text{Average of} \\ \text{Births from 0 to } x}} e^{-\int_0^x \bar{r}(a)da}\cdot p^*(x)$$

Part III: Making Discretization Assumptions

Assumption 2: Approximate midpoint quantities by summing across intervals and dividing by the length of the interval. For the halfway point, sum to the halfway point.

$$ {}_5N_x^* = B^* \underbrace{\exp\left[2.5\,{}_5r_x + 5\cdot\sum_{a=0,5}^{x-5} {}_5r_a\right]}_{\substack{\text{CumulationFunction} \\ S_x}} {}_5L_x^* $$

Part III: Making Discretization Assumptions

Assumption 2: Approximate midpoint quantities by summing across intervals and dividing by the length of the interval. For the halfway point, sum to the halfway point.

$$ {}_5N_x^* = B^* \underbrace{\exp\left[2.5\,{}_5r_x + 5 \cdot \sum_{a=0,5}^{x-5} {}_5r_a\right]}_{\substack{\color{blue}{CumulationFunction} \\ S_x}} {}_5L_x^* $$

$$ {}_5N_x^* = B^* \cdot S_x \cdot {}_5L_x^* $$

Part III: Making Discretization Assumptions

Assumption 2: Approximate midpoint quantities by summing across intervals and dividing by the length of the interval. For the halfway point, sum to the halfway point.

$$ {}_5N_x^* = B^* \underbrace{\exp\left[2.5\,{}_5r_x + 5\cdot\sum_{a=0,5}^{x-5}{}_5r_a\right]}_{\substack{\color{blue}{CumulationFunction} \\ S_x}} {}_5L_x^* $$

$$ {}_5N_x^* = B^* \cdot S_x \cdot {}_5L_x^* $$

Part III: Making Discretization Assumptions

Assumption 3: Assuming that l_x is linear in the interval surrounding x.

$$\ell_x^* = \frac{\left({}_5L_x^* + {}_5L_{x-5}^*\right)}{10}$$

Preston et al. (2001) also uses this to characterize the last interval.

Part IV: Algorithm

The basic idea is to estimate:

$$p^*(x) = \frac{N^*(x) \exp[S_x]}{N^*(0)}$$

- Estimate ${}_5L^*_x$ Assumptions 1 and 2
- Estimate ℓ^* Assumption 3
- If $B(t)$ ("hires" for each group) sequence is available, all quantities estimable.
- Otherwise, we must start $N^*(0)$ at first calculable average, and construct "conditional" life ("job tenure" expectancies)
- Cook up some solution for dealing with the open-ended interval.

Acknowledgement & Gratitude

I would like to thank the following for support:

- Penn Population Studies Center
- National Institutes of Health (NIH T32 HD007242)
- Institute for Education Sciences (ERIC T32 HD007242)
- Samuel Preston, Michel Guillot, Herb Smith, Xi Song, Irma Elo
- and many others...