Introduction to Demographic Heterogeneity
Dem260 Math Demog
Spring 2020
Lecture 1

Joshua R. Goldstein

January 23, 2020

Welcome

- Quick overview of course
- Options for last part of semester
- Resources (lecture, notes, readings, videos)
- Practice (some standard math problems, some open ended questions, some simulations, and two applied investigations)

Agenda for today

1. What demographic heterogeneity is (and isn't)
2. Dynamics of population growth with two sub-groups

* Break

3. Keyfitz's result $\bar{r}'(t) = \sigma_r^2(t)$.
4. Ken's model of Poisson heterogeneity

Part I. Conceptual Introduction

1. What is Demographic Heterogeneity?

- If we see different outcomes (e.g., people dying at different ages), is this Demographic Heterogeneity?

1. What is Demographic Heterogeneity?

- If we see different outcomes (e.g., people dying at different ages), is this Demographic Heterogeneity? NO.

1. What is Demographic Heterogeneity?

- If we see different outcomes (e.g., people dying at different ages), is this Demographic Heterogeneity? NO.
- Demographic heterogeneity = different rates for different folks.

  > *In a* demographically heterogeneous *population, people are of different types, with different type-specific rates.*

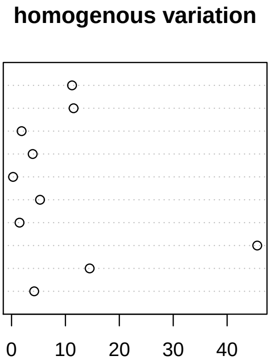
  (These types can be discrete, with individuals being homogeneous within their type, or they can be continuous with possibly no individual having exactly the same risk as another.)

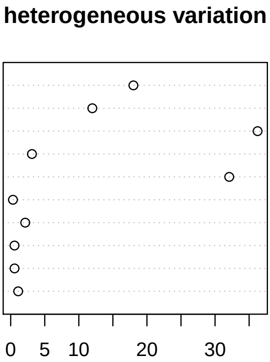
1. An example

Let's draw 10 individuals from a homogeneous population and heterogeneous population.

```
## Homogeneous hazard of 1/10
set.seed(13)
x.homo <- rexp(10, rate = 1/10)
## Heterogeneous hazard (half 1/6 and half 1/13)
## Note: I didn't pick these particular numbers for any sp
x.hetero <- c(rexp(5, rate = 1/6),
              rexp(5, rate = 1/13))
par(mfrow = c(2,1))
dotchart(x.homo, main = "homogenous variation")
dotchart(x.hetero, main = "heterogeneous variation")
```

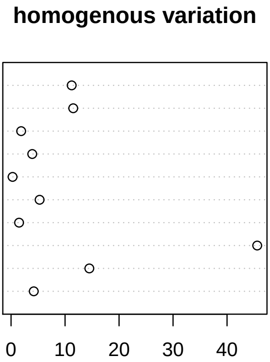
1. An example, continued

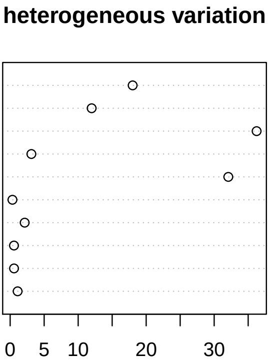




- Can you tell which is which?

1. An example, continued





- Can you tell which is which?
- Would we expect to see a difference if we increased sample?

1. Components of variation

- **Homogeneous:** Chance only
- **Heterogeneous:** Chance + group variation in risk

1. Analogies

Social inequality: equal opportunity vs. equal outcomes

Analysis of variance

$$\text{total variance} = \text{within group} + \text{between group}$$

Statistical models $y = a + bx + \epsilon$

What's new? Dynamics.

Heterogeneous populations evolve differently. Aggregates $\neq$ Individuals

- Rates of growth (or decline)
- Changes over time or age or duration
- The trajectory of even the average individual differs from population average
- Relative positions, change of groups, may be misleading.

1. Terminology

- Heterogeneity
- Unobserved heterogeneity
- Selection
- Selective survival
- Other terms?

1. Heterogeneity at work? Black-White Crossover

1. Heterogeneity at work? Mortality Plateus

1. Big Caveat: Fundamental Unidentifiability

Same data of N observations

- N draws from 1 distribution
- 1 draw from N distributions
- Something in-between

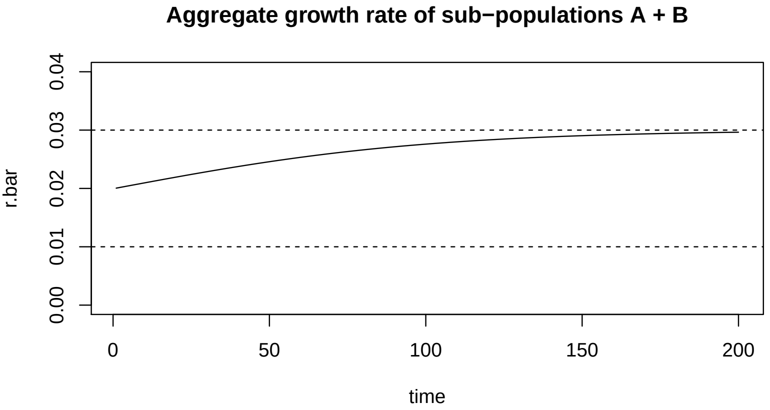
Abel (66) and Beth (76) example.

2. A 2nd example: Exponential growth, two countries

Two countries start equal size, but grow at different rates. What happens to aggregate growth rate?

```r
rA = .03 ## growth rate of A
rB = .01 ## growth rate of A
KA = 100 ## starting pop size
KB = 100
t = 0:200
KA.t = KA*exp(rA*t) ## exp growth of A
KB.t = KB*exp(rB*t) ## exp growth of B
K <- KA.t + KB.t ## combined pop
r.bar = diff(log(K)) ## growth rate
plot(t[-1], r.bar, type = "l", ylim = c(0, 0.04),
     ylab = "r.bar", xlab = "time")
abline(h = c(rA, rB), lty = 2)
title("Aggregate growth rate of sub-populations A + B")
```

2. A 2nd example: Exponential growth, two countries



- What determines growth rate?
- How does it change over time?
- Does the process converge?

2. Some more examples

1. Differential, constant mortality ($\mu_A = .03$; $\mu_B = .01$)
2. Differential, *time-varying* mortality or growth.
3. "Movers and Stayers" (Migration)
4. "Movers and Stayers" (Marriage)

2. Even more examples

1. Fecundity: aging or heterogeneity?
2. Divorce: duration or heterogeneity?
3. Duration of unemployment: duration or heterogeneity?
4. Recidivism by time out of prison
5. What else?

2. Fun with App

https://shiny.demog.berkeley.edu/josh/het_ruse/

- Can you create a plateau?
- Can you create a crossover?
- Can you get aggregate rate to decline?
- Anything else?

Cookie break here?

Part II. Formal Analysis

1. Keyfitz result
2. Ken's Poisson-Exponential Model
3. Keyfitz USA-Mexico example

3. Keyfitz result

$$\frac{d}{dt}\bar{r}(t) = \sigma_r^2(t)$$

When group-specific growth rates are constant, the rate of change of the aggregate growth rate equals the variance of the growth rates.

3. Derivation

By definition,

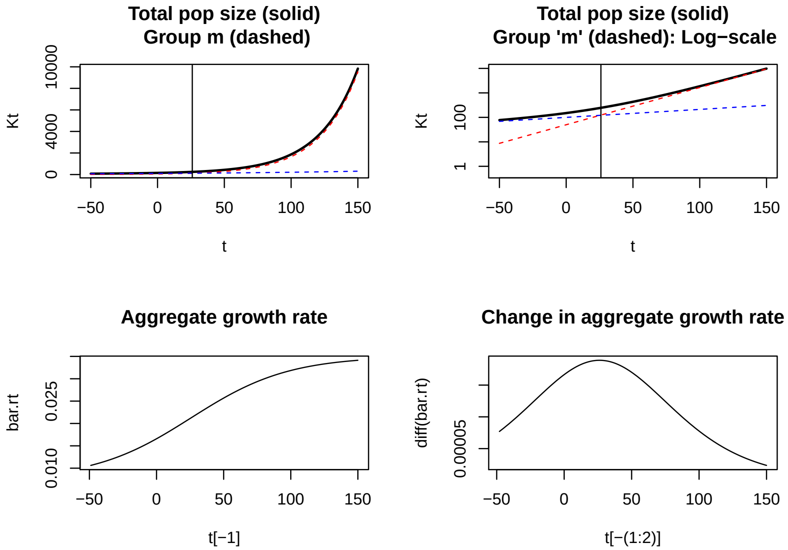
$$\bar{K}(t) = \sum_i K_i(t) = \sum_i K_i e^{r_i t}$$

and

$$\bar{r}(t) = \frac{\frac{d}{dt}\bar{K}(t)}{\bar{K}(t)}.$$

Let's take derivatives and simplify, recalling definition of variance.

3. US-Mexico Example



Appears that pop sizes maximizes variance. (Check at home?)

US-Mexico Example

```r
rm = 3.5/100
ru = .75/100
Km = 50
Ku = 100
t <- -50:150 ## go back in time to see rise and fall of variance
Kt = Km * exp(t*rm) + Ku * exp(t*ru)
bar.rt <- diff(log(Kt))
par(mfrow = c(2,2))
plot(t, Kt, lwd = 2, type = 'l')
title('Total pop size (solid)\n Group m (dashed)')
lines(t, Km * exp(t*rm), lty = 2, col = "red")
lines(t, Ku * exp(t*ru), lty = 2, col = "blue")
my.v = 26
abline(v = my.v)
plot(t, Kt, lwd = 2, type = 'l', log = 'y', ylim = c(.5, max(Kt)))
lines(t, Km * exp(t*rm), lty = 2, col = "red")
lines(t, Ku * exp(t*ru), lty = 2, col = "blue")
abline(v = my.v)
title("Total pop size (solid)\n Group 'm' (dashed): Log-scale")
plot(t[-1], bar.rt, type = 'l', main = 'Aggregate growth rate')
plot(t[-(1:2)], diff(bar.rt), type = 'l',
main = 'Change in aggregate growth rate')
```

Appears that pop sizes maximizes variance. (Check at home?)

3. Commentary on Keyfitz result

- Growth rates in heterogeneous populations start at pop average and then increase.
- Heterogeneity pop growth
- We will extend to cover non-constant growth

But

- Doesn't tell us how much bigger $\bar{K}(t)$ is projection using constant aggregate rate $\bar{r}(0)$.
- Doesn't give us a formula for time path of aggregate $\bar{K}(t)$ or $\bar{r}(t)$

Note: our homework will try to address some of this using Taylor approximation.

4. The Origin of Ken's Poisson-Exponential Model

Given a world with many sub-populations, each growing expontentially at their own rate, what can we say about the time-path of world population growth?

4. Ken's Poisson-Exponential Model

From an email:

Josh asks: Suppose we have a discrete mix of subpopulations growing at different intrinsic rates r whose maximum is r0. Is there a handy approximation for the growth path of the aggregate populations?

The assumption of a discrete mix is essential here. Otherwise Tauberian theorems apply and, with a vanishingly small portion of the population close to the maximum growth rate, we do not obtain long-run exponential growth.

I recommend modeling the discrete distribution of growth rates as a mixture of Poisson distributions.

4. Poisson-based model

We are considering

$$\bar{K}(t) = \sum_i e^{r_i t} K_i(0). \tag{1}$$

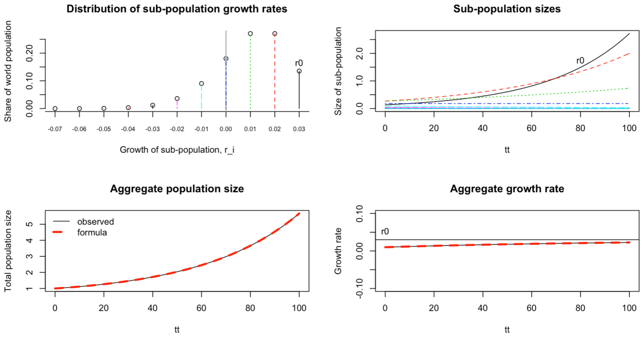
Ken suggests

$$r_i = r_0 - s(\lambda) \cdot a, \tag{2}$$

- r_0 growth rate of the fastest growing sub-population
- s a non-negative Poisson distributed integer
- λ the parameter of the Poisson distribution (also it's mean and variance)
- a gap between adjacent growth rates.

Example: sub-populations have growth rates 3, 2, 1, 0, -1, ... percent, then $r_0 = 0.03$ and $a = 0.01$. Sizes of sub-pops determined by Poisson dis'n

Figure: A simulation of heterogeneous growth



4. Closed-form result

$$K(t) = K(0)e^{r_0 t}e^{-\lambda(1-e^{-at})}.$$

To derive:

- Write out mixture to get

$$K(t) = K(0)e^{r_0 t}\sum_i e^{-sat}f(s)$$

- Substitute for $f(s)$: *Pois* $\frac{\lambda^s e^{-\lambda}}{s!}$
- Recognize that our mixture contains the series representation of e^{-at}
- Done

4. Interpretation

$$K(t) = K(0)e^{r_0 t} e^{-\lambda(1-e^{-at})}.$$

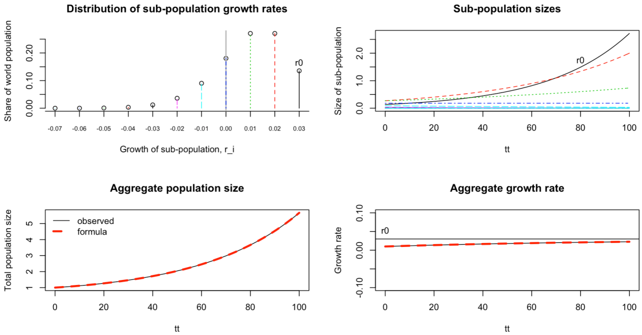
- Dominant term contains the maximum population growth rate r_0,
- Second term gives the diminishing effect of the sub-populations with smaller population growth rates over time.

4. Some further analysis

What is the closed-form expression for $\bar{r}(t)$?

4. And the formula works!

Figure: A simulation of heterogeneous growth



4. Some commentary

- Poisson and Exponential “fit”
- We’ll this complementarity again (e.g., with Gamma)
- Tractable models are super powerful for enhancing our understanding.
- But be careful. Avoid extremes: the model is right/wrong.

A BIG caveat

Are disaggregated models necessarily better?

A BIG caveat

Are disaggregated models necessarily better? Some potential problems:

- Aggregate constraints?
- Interacting sub-populations?
- Illusion of precision?

Next time . . .

- Mortality
- Frailty (Multiplicative and fixed)
- Gamma