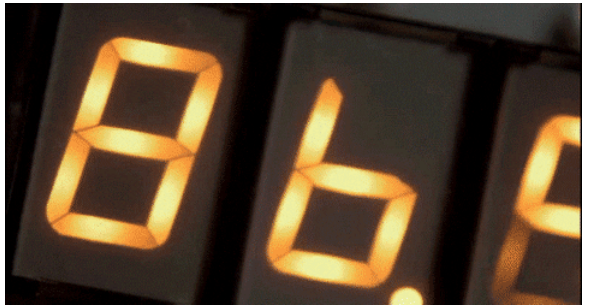
Demographic projection

Matt Bhagat-Conway

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

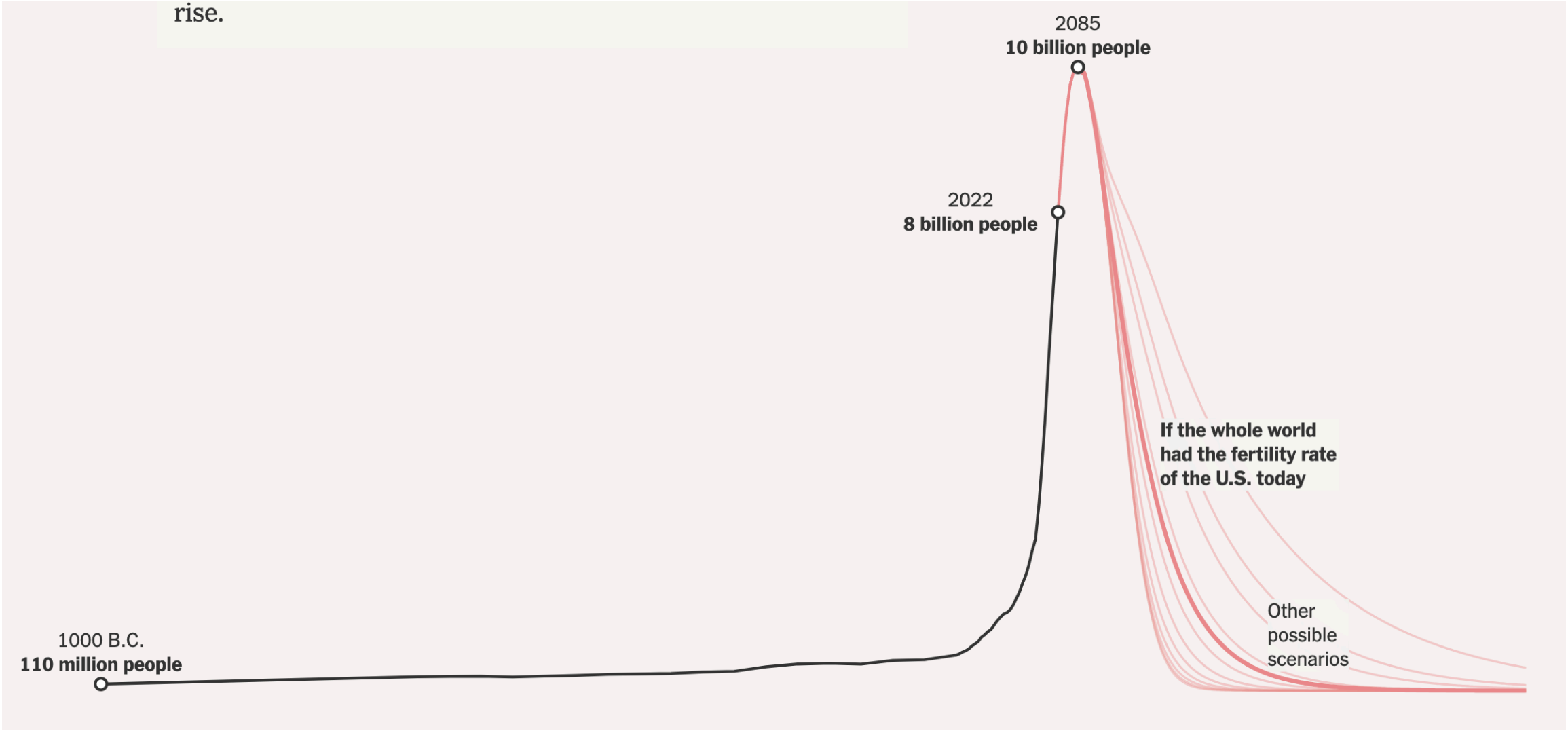
Demographic projection

- Census data is always about the (recent) past, not the present and certainly not the future
- In planning, though, we are by definition looking to the future
- This requires us to *project* what the future population will look like



© Universal Pictures

Uses for demographic projections



NYTimes

Uses for demographic projections

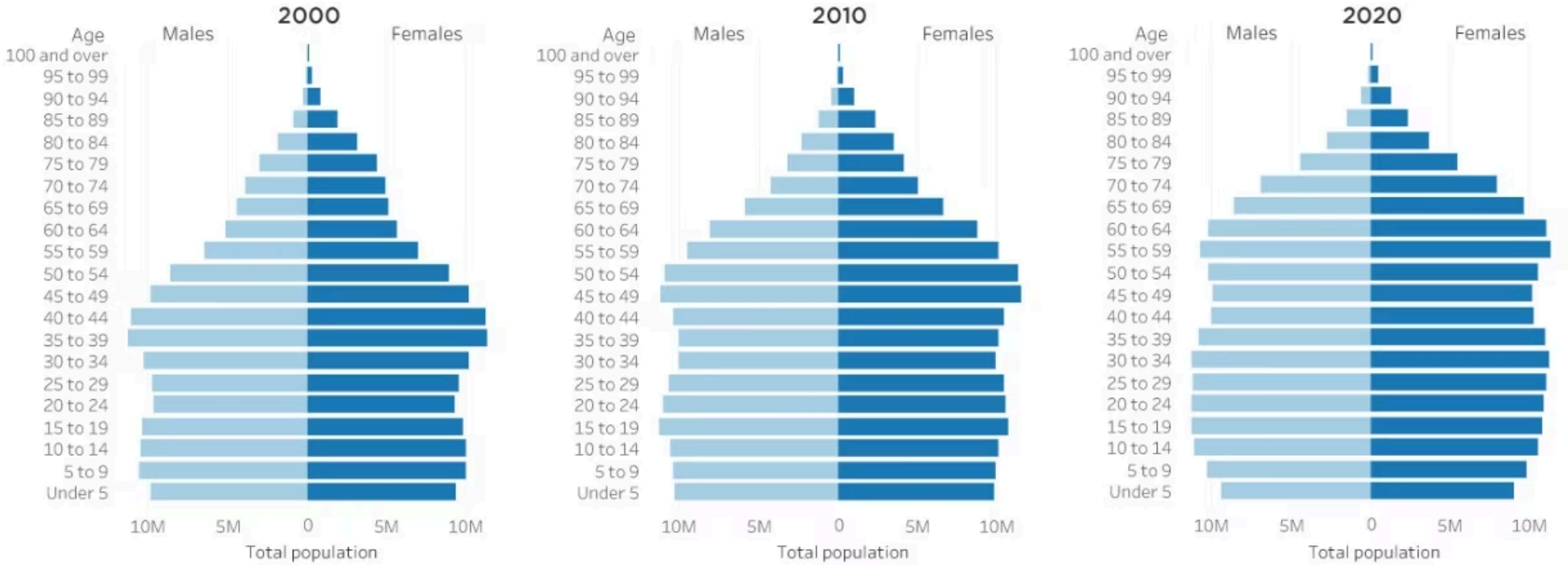
- Demographic projections form the basis of long-range planning
- Used to develop regional housing plans
- Critical in transportation modeling
- May be used to investigate scenarios (e.g. better maternal health care, increased immigration)

Describing populations: the population pyramid

Figure 2.

Population Pyramids for the United States: 2000, 2010 and 2020

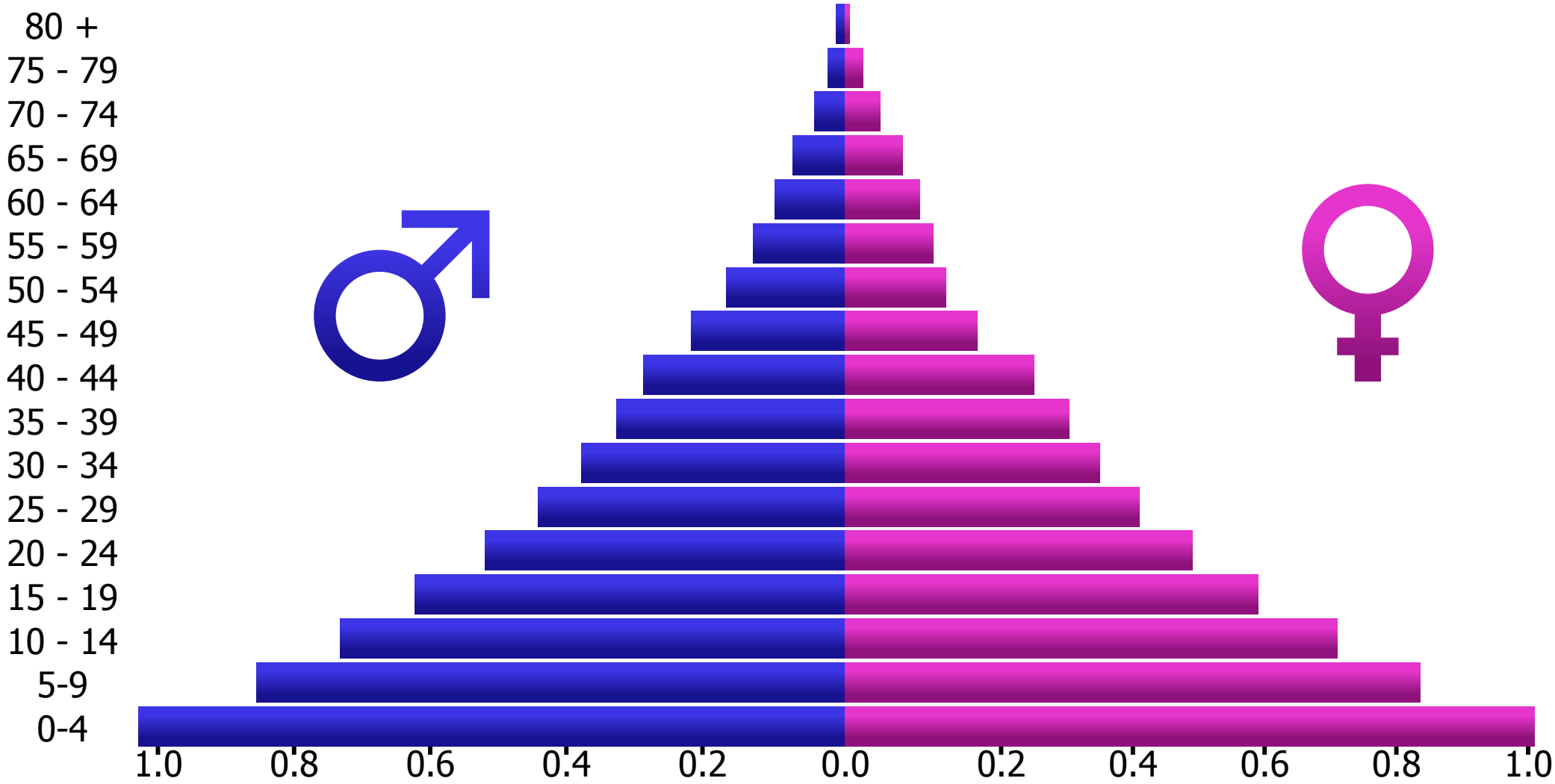
(In millions)



Source: U.S. Census Bureau, Census 2000 Summary File 1 (SF1), 2010 Census Summary File 1 (SF1) and 2020 Census Demographic and Housing Characteristics File (DHC).

US Census Bureau

Population pyramids in developing countries



Population pyramid for Angola, Wikipedia

The demographic transition

- Historically, human populations had high birth and death rates
- In the last century or so, death rates have fallen rapidly around the world
- Birth rates have as well, but lagged behid death rates

How are populations forecasted?

The cohort-component model

- By far the most common demographic forecasting method
- We divide the population into *cohorts*—groups of people sharing some characteristic (e.g. men aged 30-35)
- We then simulate how each cohort changes over time—how many people die, how many people give birth, how many people leave the region, and how many people migrate into the region

Many more details available in Smith, Tayman, and Swanson (2013), which this lecture is based on

The cohort-component model

- You start with the population from a base or *launch year*, divided into cohorts
  - Generally divided by age and biological sex/child-bearing potential, possibly also by race and ethnicity
- You then adjust that to make a prediction for the next *period* (often 1, 5, or 10 years later)
  - Use survival rates by age to estimate how many in each cohort survive to the next period
  - The survivors "graduate" to the next cohort—e.g. women 20–25 in a 2020 launch year will be 26–30 in 2025
  - Use fertility rates to predict how many births there will be; these become the new youngest cohort
  - Use migration estimates to adjust how the population changes due to migration
- Repeat the process to forecast the next period, and so on

Mortality

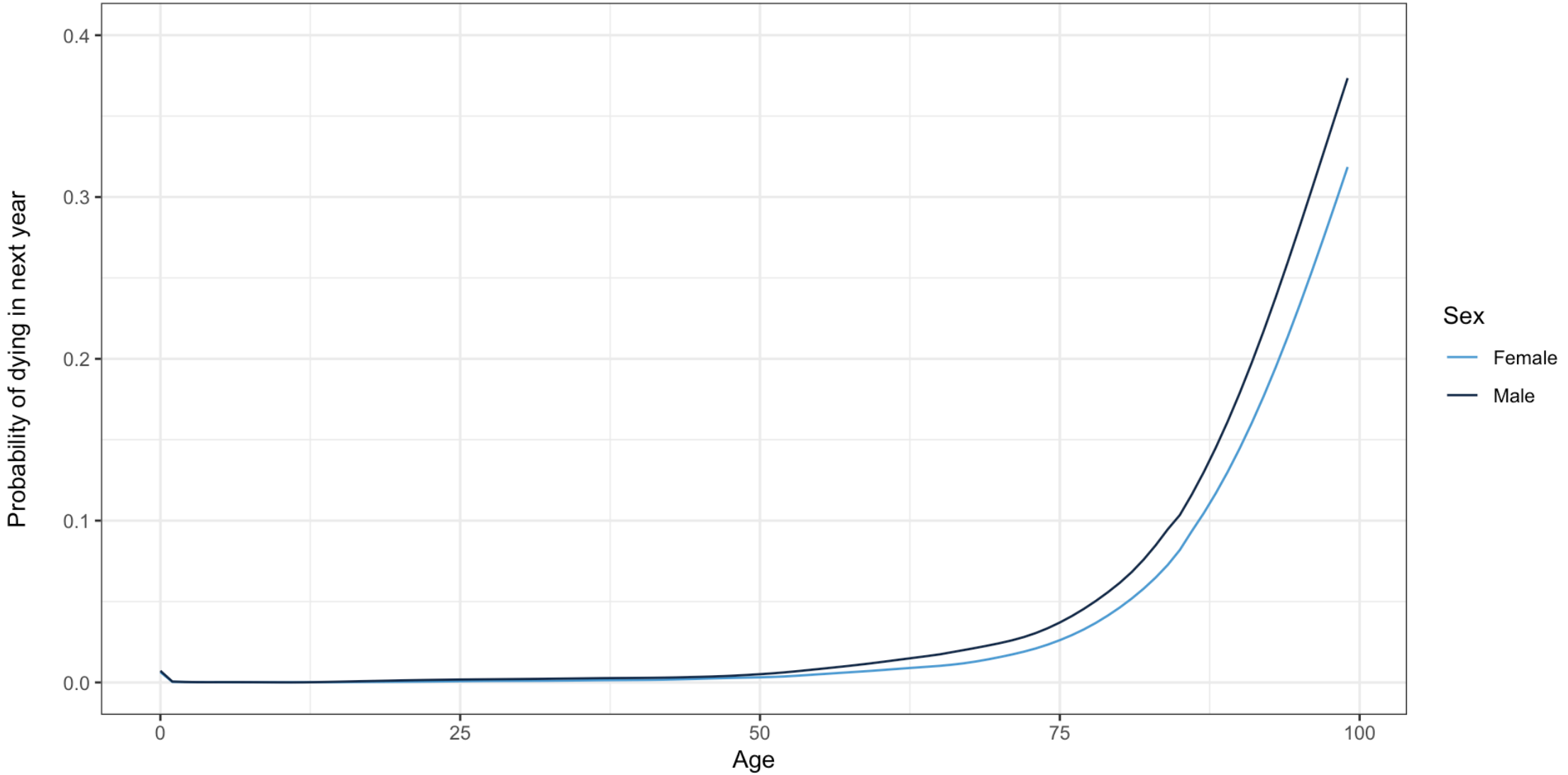
- Mortality data is generally presented in *life tables*
- Life tables show the probability of death in the next period for people of a particular age ::::

Life tables

age	Probability of dying in interval	Number surviving to start of interval	Person-years lived in interval
0–5	0.0072	100,000	496,629
5–10	0.0004	99,259	496,149
10–15	0.0004	99,207	495,916
15–20	0.0011	99,151	495,437
20–25	0.0019	99,007	494,455
25–30	0.0034	98,749	492,732
30–35	0.0040	98,325	490,398
35–40	0.0053	97,817	487,501
40–45	0.0067	97,160	483,817
45–50	0.0103	96,315	478,552
50–55	0.0143	95,039	471,052
55–60	0.0231	93,252	459,695
60–65	0.0331	90,461	443,102
65–70	0.0452	86,613	421,051
70–75	0.0710	81,518	389,908
75–100	0.9610	73,955	911,918

Selected rows from the 2019 life table for females in North Carolina, US CDC

Mortality



Probability of death in next year, North Carolina, 2019 (CDC)

Calculating survival rates from life tables

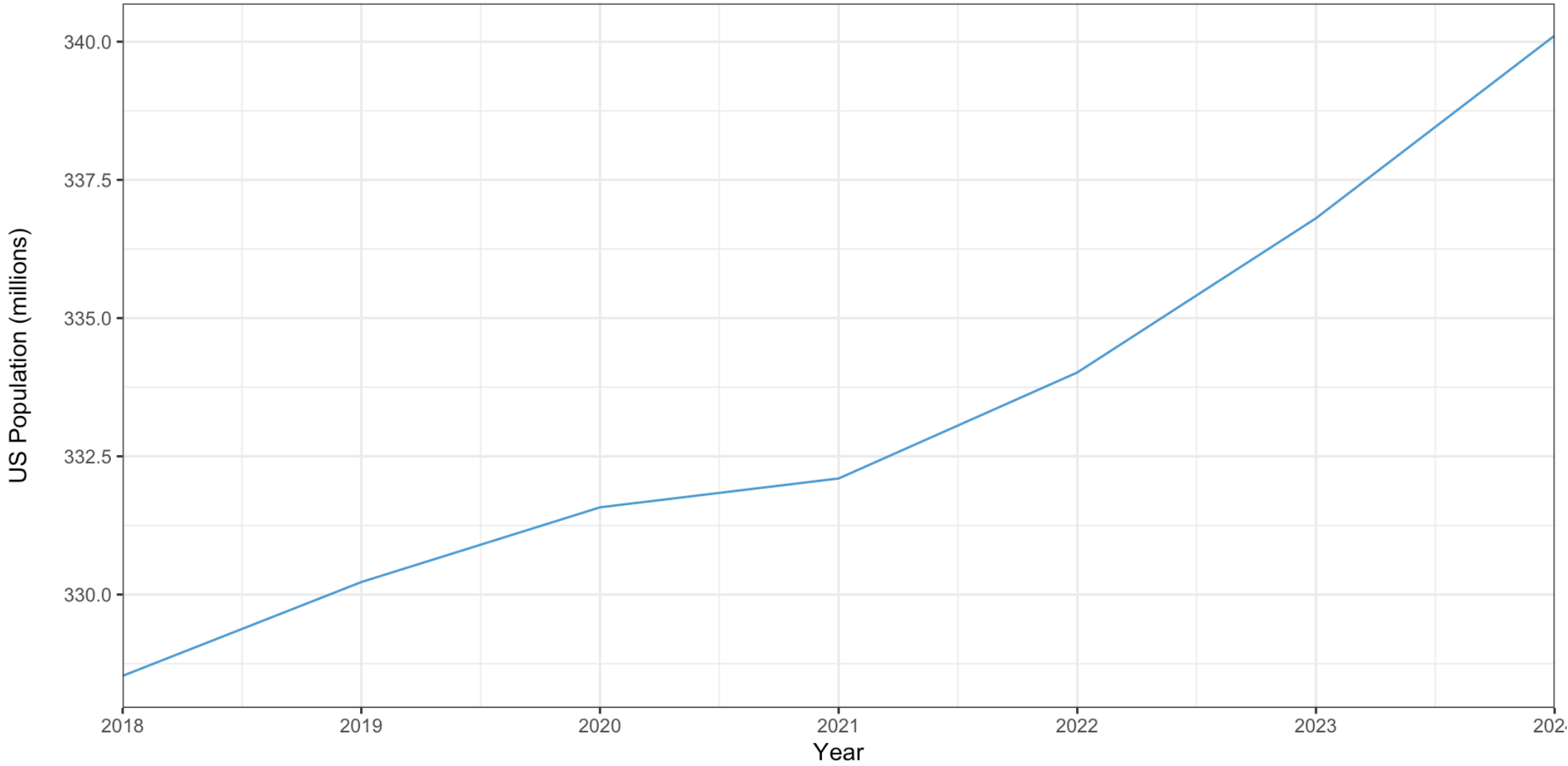
Calculating survival rates from life tables

Applying survival rates in the cohort-component model

How do you forecast survival rates?

- This is where demographic forecasting becomes an art
- We know what survival rates are from past periods to now, but we don't know what scientific breakthroughs/disasters/policy changes might occur in the next 50 years
- Since you repeat the cohort-component model many times, any errors in survival rate will compound

Survival rates and major events

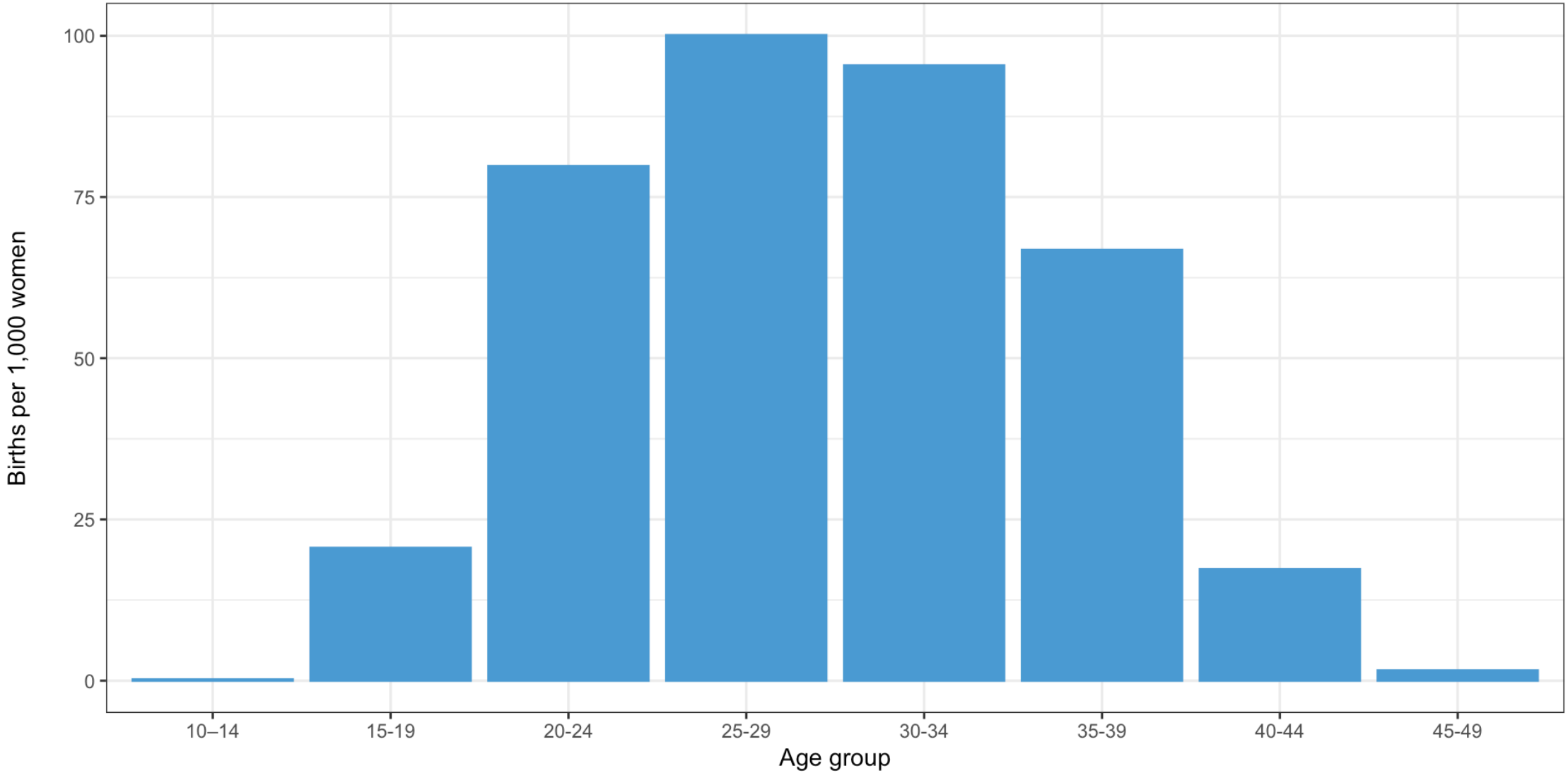


US population, 2018–2023. Source: FRED

Fertility

- Fertility is one of two main ways populations grow
- We measure the *age-specific birth rate* and multiply that by the number of women in each age group to get an estimate of births
- These births are added directly to the first cohort in our next period

Age-specific birth rates in NC



Age-specific birth rates, North Carolina, 2017 (CDC)

Estimating fertility

Forecasting fertility

- Again, this is difficult, with many factors
  - Economic conditions
  - Support for parents
  - Access to contraception, abortion, and other reproductive care
  - etc etc

Variations in fertility

Migration

- Migration from/to outside the region is a critical driver of population change
  - especially in NC
- Data on migration is somewhat hard to come by, relative to fertility and mortality data

Migration

- There are two methods for accounting for migration in population projection: *net migration* and *gross migration*
- *Gross migration* involves calculating both rates of inmigration and outmigration
- *Net migration* involves just calculating the difference, which is easier with some datasets

Gross migration rates

- The *outmigration* rate is the number of people who left the area divided by the number of people in the area at the start
- The *inmigration* rate is the number of people who entered the area, *generally* divided by the population of the nation less the population of the area
  - i.e. the probability someone somewhere else in the US moves to the Triangle
- Foreign immigration often handled separately

Migration

- There are few data sources on gross migration
- The ACS does ask about migration, but only on a one-year time horizon
- Most population forecasting models use a longer time horizon
- Because people can move more than once, aggregation is difficult

Migration

Migration: denominator

- Smith, Tayman, and Swanson (2013) recommend for slow-growing regions to use a migration rate based on the regional population, and for fast-growing regions to base it on the adjusted national population (e.g. US population minus Triangle population)
- Because more folks are moving in than out, so the population of the origin is more relevant
- You can apply migration rates to launch year population or to survived population, you just need to calculate the migration rate based on the same population you apply it to
- Even though the Triangle is fast-growing, we're going to use the regional population for simplicity
- Otherwise we'd need to estimate the national population as well

Forecasting migration

- Migration depends on a variety of economic factors domestically and internationally
- For instance, high housing prices in coastal cities are likely partly driving growth in the Triangle
- Forecasting this can be difficult
- We'll just assume constant migration

Migration

Putting it all together

Putting it all together

- Download Excel population forecasting workbook from https://go.unc.edu/population

Calculating survivors

- Multiply each cohort by its corresponding survival rate from the life tables sheet
- You can enter one formula and then just drag down

Calculating migration

- Multiply each cohort's starting value by 9.29%

Calculating fertility

- Multiply each child-bearing cohort by the fertility rates in the fertility sheet
  - Remember to divide by 1000 and multiply by five to get a five-year birth rate per capita
  - Use the average of the launch-year population and the survived population (assume that women who died lived through half the interval)
  - Add the migration as well
  - Use the average of this cohort and the next cohort's birth rates, because women will get older during the projection period

Finishing the projection

- Add up survival and migration, move them to the next cohort
- For the oldest cohorts, add survival and migration from the oldest and second-oldest in the launch year
- For the youngest cohorts, sum up fertility and multiply by 0.5 (we're assuming births are completely balanced by sex)
- Ideally, we would also apply a survival rate to these births

Finishing the projection

Special events and populations

- In some areas, you may need to account for certain special events or populations
  - e.g. New Orleans post-Katrina, or the pandemic across the US - extrapolating from these events could give incorrect answers
- In Fayetteville, population dynamics might be different due to the large military presence (folks are assigned there)
- College towns may have similar dynamics
- Often these populations are modeled separately

Sensitivities

- Forecasts can be sensitive to the values used, so think carefully about this
- Not only what your projected rates are, but what group you're applying to
  - e.g. recall how we didn't just use the fertility rate for each cohort, but averaged them, because some people will age out of the cohort during the projection period

References

Smith, Stanley K., Jeff Tayman, and David A. Swanson. 2013. *A Practitioner's Guide to State and Local Population Projections*. Vol. 37. The Springer Series on Demographic Methods and Population Analysis. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-7551-0.

This work by Matthew Bhagat-Conway is licensed under a Creative Commons Attribution 4.0 International License.