Lecture 2:
Periods and cohorts

Ernesto F. L. Amaral

January 31–February 02, 2023
Demographic Methods (SOCI 633/320)

www.ernestoamaral.com

TEXAS A&M UNIVERSITY

Periods and cohorts

- Lexis diagram
- Person-years
- Rates, probabilities, ratios
- Crude rate model
- Infant mortality rate
- Person-years and areas
- Cohort person-years lived
- Stable and stationary populations

Exponential population growth model

- The exponential model treats all people as if they were alike
  - No mention to *age*
  - However, people are aging in the population

- Time enters demography in two ways
  - Chronological time: calendar dates, same for everyone
  - Personal time: age for each set of people who share same birthdate

Lexis diagram

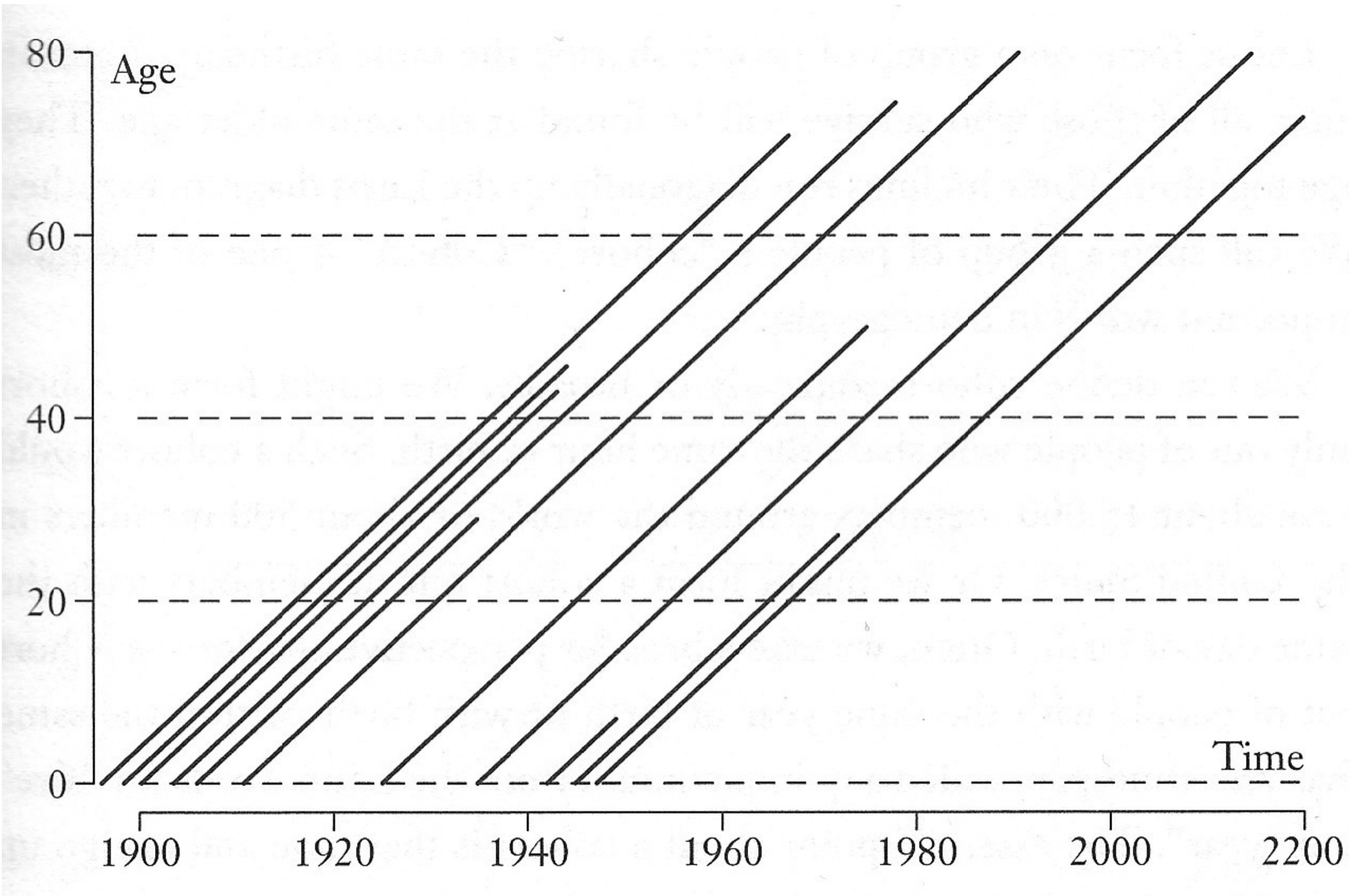
- Lexis diagram provides relationships between chronological time *t* (horizontal) and age *x* (vertical)

- Each person has a lifeline on a Lexis diagram
  - Starting at (t_b, 0), where t_b is the person's birthdate and 0 is the person's age at birth

- Line goes up to the right with a slope equal to 1
  - People age one year in one calendar year

- Lifeline goes up until time and age of the person's death

Lexis diagram



Source: Wachter 2014, p. 31.

Exploring Lexis diagram

- To find population size
  - Draw vertical line upward from the time point
  - Count how many lifelines cross vertical line

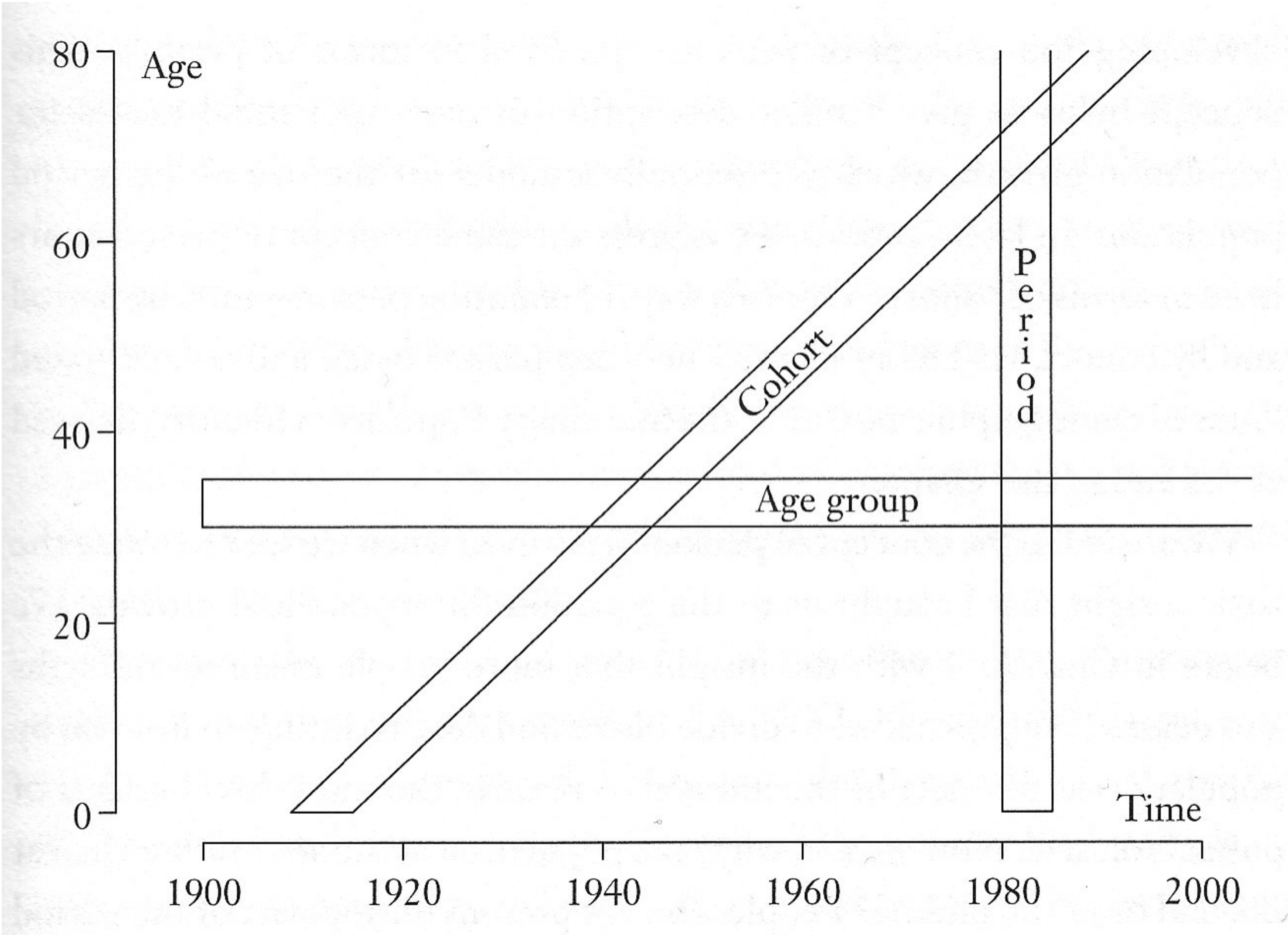
- To find how many people survive to some age
  - Draw horizontal line across at the height corresponding to that age
  - Count how many lifelines cross that horizontal line

- Immigrants start at age and time of immigration

Cohort

- Group of people sharing the same birthdate
- Group of individuals followed simultaneously through time and age
- Their lifelines run diagonally up the Lexis diagram together
- In a cohort, time and age go up together
- A cohort shares experiences

Age, period, cohort



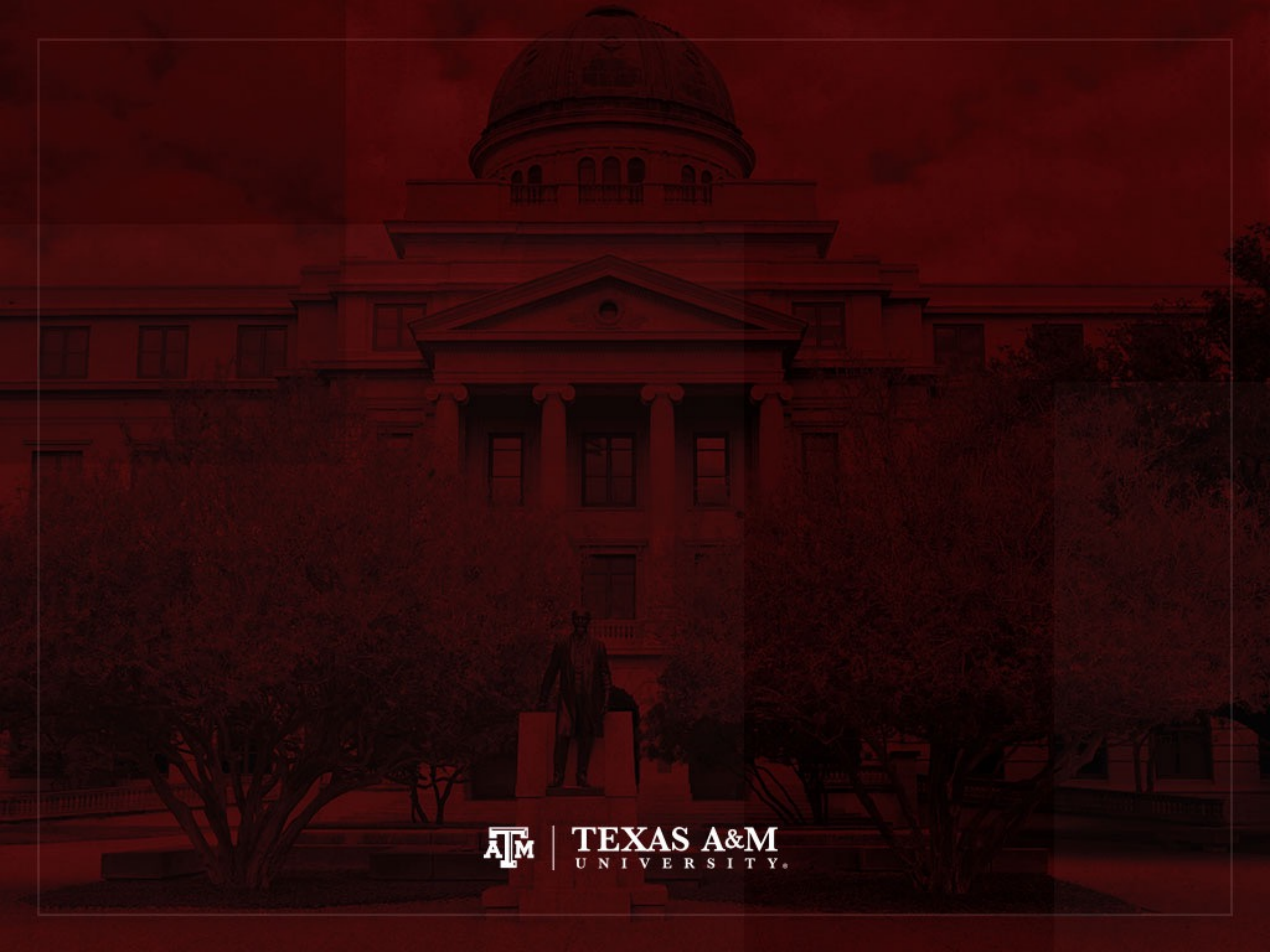
Source: Wachter 2014, p. 33.

Exponential growth

- For the equation for exponential growth
  - We divided births and deaths during an interval by population at the start of the interval

$$K(1) = K(0)\left(1 + \frac{B(0)}{K(0)} - \frac{D(0)}{K(0)}\right)$$

- Why not population at the end or in the middle?
  - People who are present during part of the period can also have babies or become corpses
  - More people present for more time in the denominator generate higher exposure (“risk”) to births and deaths


TEXAS A&M
UNIVERSITY

Person-years

- **Person-years** is the sum of each individual's time at risk of experiencing an event (e.g. birth, death, migration)
  - For those who do not experience event, person-years is the sum of time until end of period
  - For those who experience event, it is the time until the event

- **Period person-years lived** (PPYL) take into account that people are present during part of the period (fraction of years)
  - Each full year that a person is present in a period, he/she contributes one "person-year" to the total of PPYL
  - Each month a person is present in the population, he/she contributes 1 person-month, or 1/12 person-year, to PPYL

Example of person-years

Hypothetical population increasing at the rate of 0.001 per month

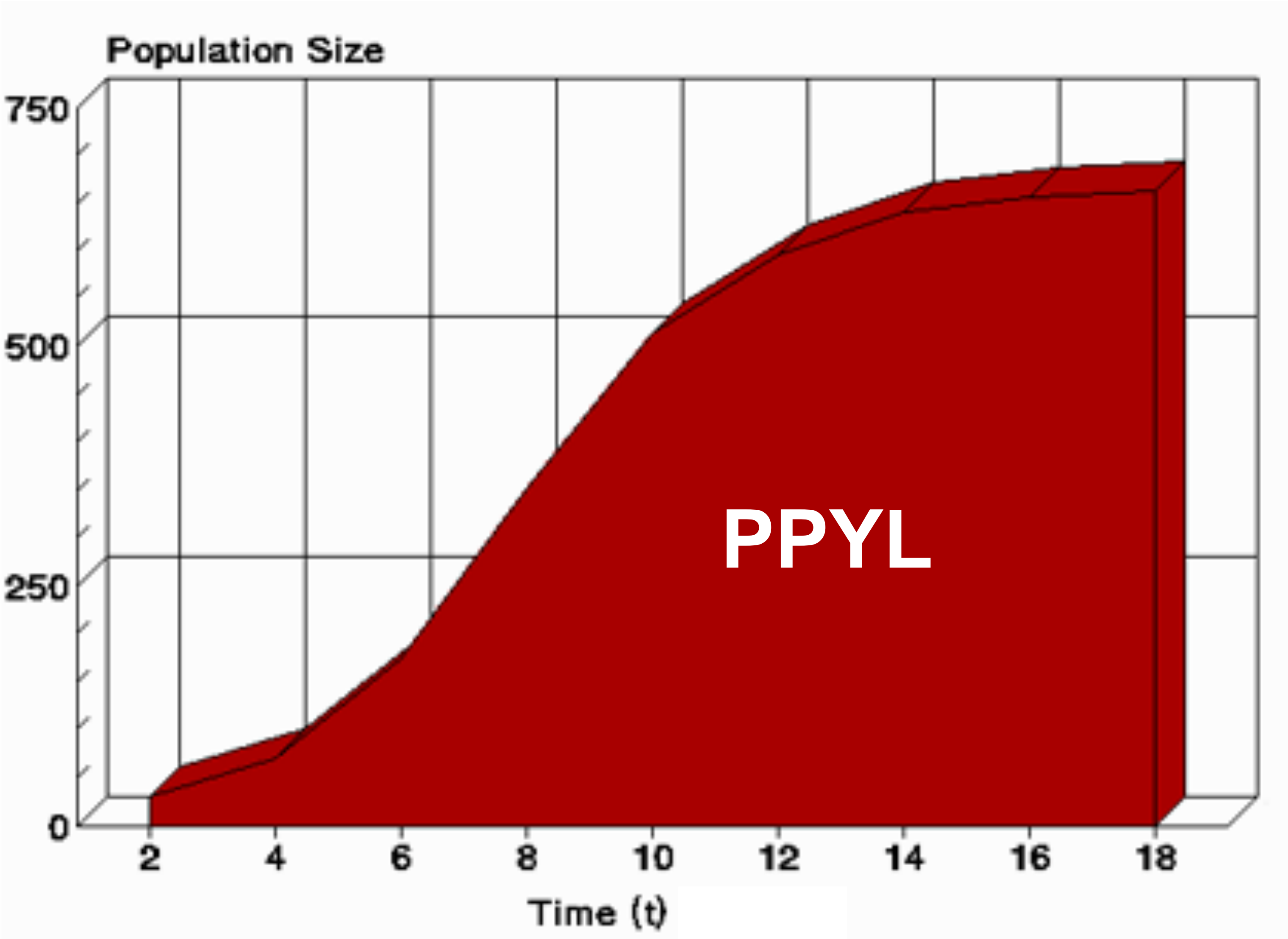
Month	Population	Person-years (population / 12)	Approximation for person-years: Mid-period	Approximation for person-years: Average of start and end
January	200.00	16.67		200.00
February	200.20	16.68		
March	200.40	16.70		
April	200.60	16.72		
May	200.80	16.73		
June	201.00	16.75		
July	201.20	16.77	201.20	
August	201.40	16.78		
September	201.61	16.80		
October	201.81	16.82		
November	202.01	16.83		
December	202.21	16.85		202.21
Period person-years lived (PPYL)		**201.10**	**201.20**	**201.11**

Calculating person-years

- Whenever we know the population sizes on each month over the period of a year
  - We can add up the person-years month by month
  - Take the number of people present on each month and divide by 12
  - Add up all monthly contributions

- When our subintervals are small enough
  - Our sum is virtually equal to the area under the curve of population as a function of time during the period...

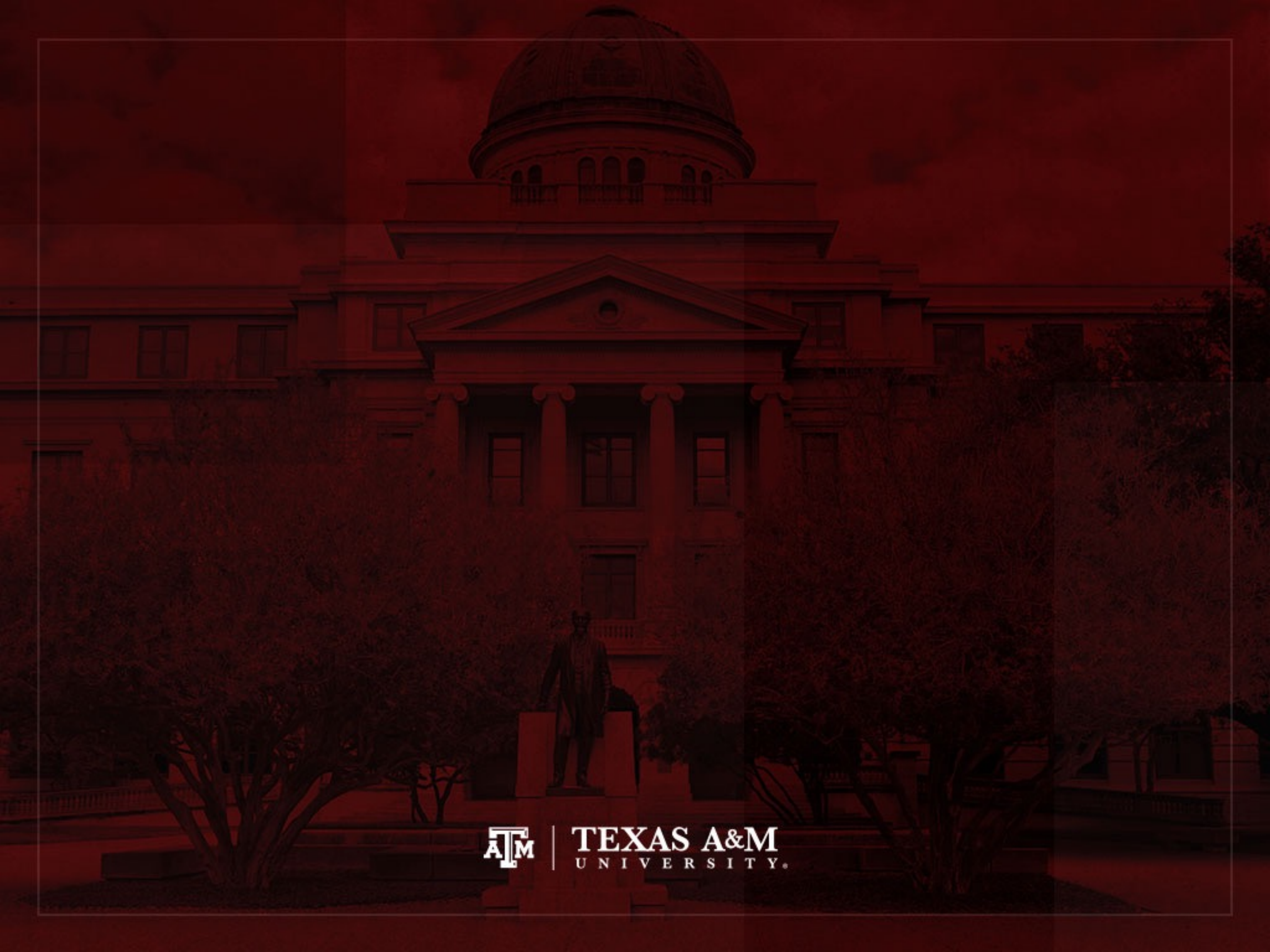
Person-years and areas



Source: https://www2.palomar.edu/users/warmstrong/lmexer9.htm.

Approximation for PPYL

- When sequences of population sizes throughout a period are unknown
  - Take the population in the middle of the period and multiply by the length of the period
    - E.g., for 2005–2015, we take the mid-period count of 308,745,000 people in the U.S. from the 2010 Census and multiply by 10 years to obtain 3,087,450,000 person-years in the period
  - Or take the average of the starting and ending populations and multiply by the length of the period


TEXAS A&M
UNIVERSITY

Rates, probabilities, ratios

- Rates
  - Describe the number of occurrences of an event for a given number of individuals who had the chance to experience that event per unit of time
- Probabilities
  - Divide the number of events by the total number of people at risk in the relevant time frame
- Ratios
  - Compare the size of one group to the size of another group

Rates

(Fleurence, Hollenbeak 2007)

- Rates are an instantaneous measure that range from zero to infinity
  - Rates describe the number of occurrences of an event for a given number of individuals per unit of time
  - Rates consider the time spent at risk
- Numerator
  - Number of events (e.g. births, deaths, migrations) in a given time
- Denominator includes time
  - Sum of each individual's time at risk of experiencing an event for a specific population during a certain time period (person-years)
  - We can use approximations for the denominator
    - Population in the middle of the period or
    - Average of starting and ending populations for that period

Ideal way to estimate rates

- Crude Birth Rate (CBR or *b*)
  - Number of births to members of the population in the period divided by the total period person-years lived

- Crude Death Rate (CDR or *d*)
  - Number of deaths to members of the population in the period divided by the total period person-years lived

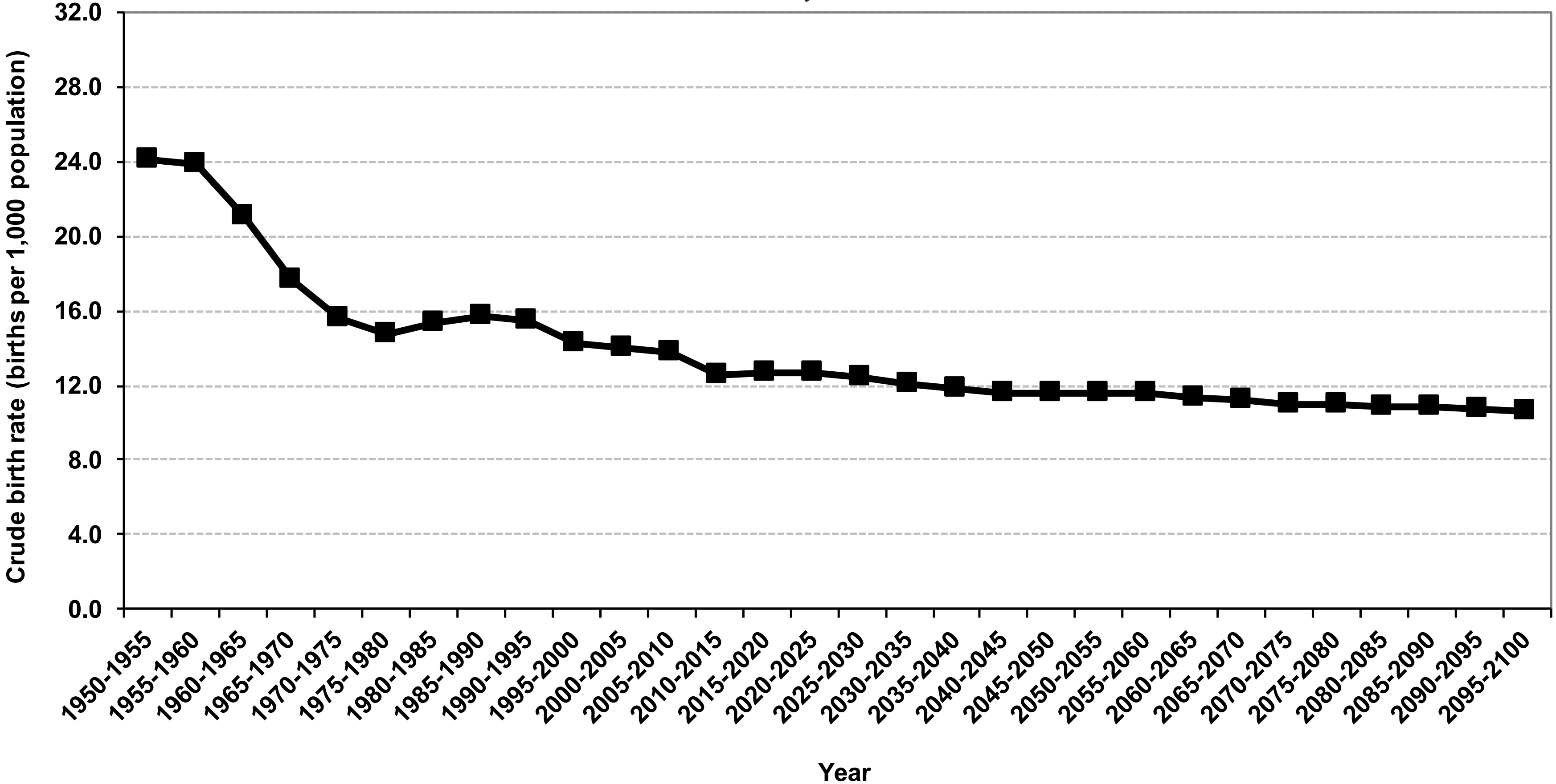
Usual way to estimate rates

- Express the number of actual occurrences of an event (e.g. births, deaths, homicides) vs. number of possible occurrences per some unit of time
  - Population in the middle of the period as denominator
- Examples

$$Crude\ birth\ rate = \frac{Number\ of\ births}{Total\ population} \times 1{,}000$$

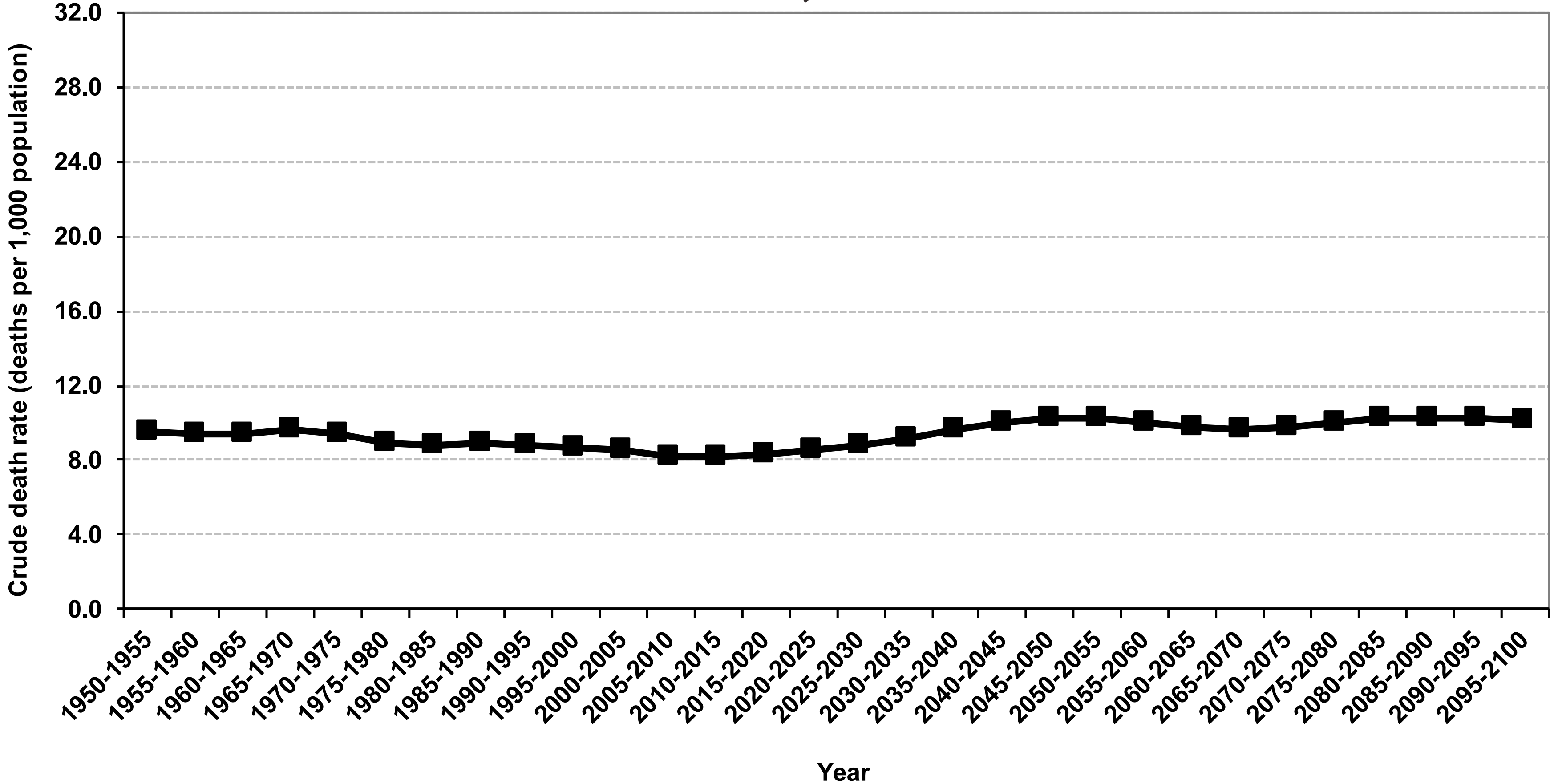
$$Crude\ death\ rate = \frac{Number\ of\ deaths}{Total\ population} \times 1{,}000$$

Crude birth rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

Crude death rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

Migration rates

- Crude or gross rate of out-migration

$$OMigR = OM / p * 1{,}000$$

- Crude or gross rate of in-migration

$$IMigR = IM / p * 1{,}000$$

- Crude net migration rate

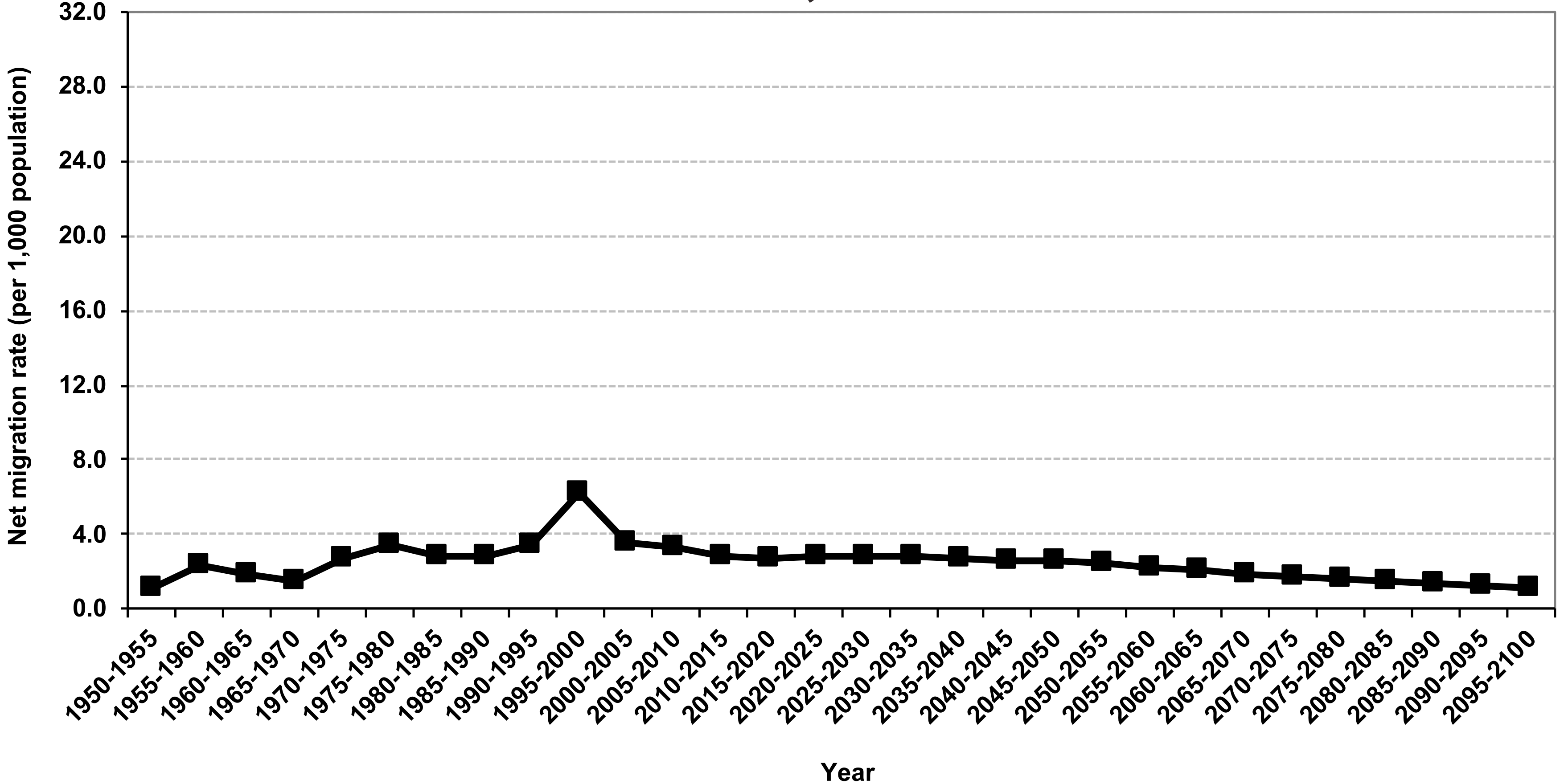
$$CNMigR = IMigR – OMigR$$

- Net migration rate

$$NMigR = IM – OM / \text{person-years lived} * 1{,}000$$

Source: Weeks 2015, Chapter 7, pp. 251–297.

Net migration rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

Probabilities

(Fleurence, Hollenbeak 2007)

- Probabilities describe the likelihood that an event will occur for a single individual in a given time period and range from 0 to 1
  - Do not include time in the denominator
  - Divide the number of events by the total number of people at risk in the relevant time frame
- Conversion between rates and probabilities:

$$\text{probability: } p = 1 - e^{-rt}$$

$$\text{rate: } r = -1/t * ln(1-p)$$

- An approximation for the denominator is the population at the beginning of the period

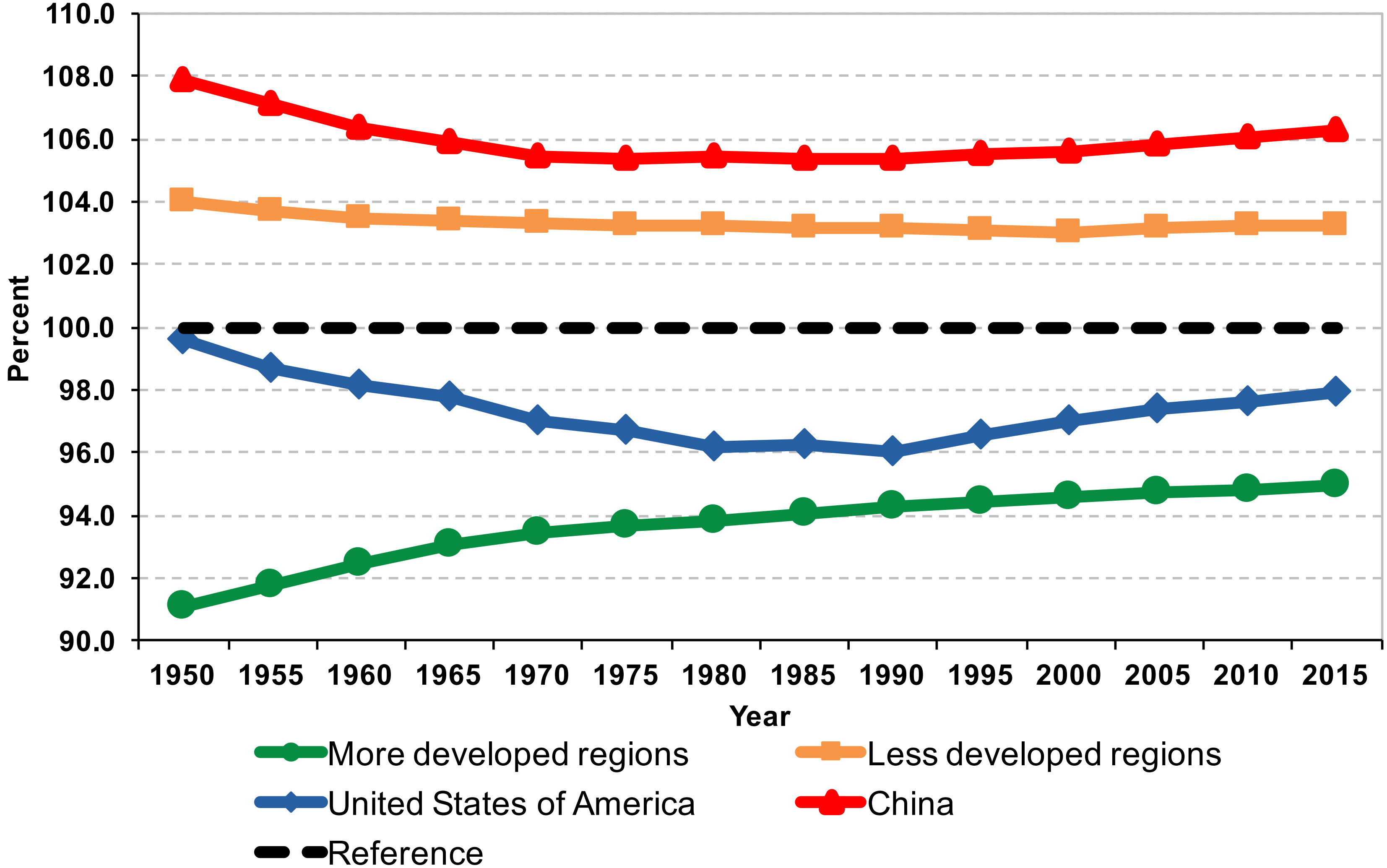
Ratios

- Describe a relationship between two numbers
  - Compare the size of one group to the size of another group
  - Compare the relative sizes of categories
  - Indicate how many times the first number contains the second
  - Denominator is not at "risk" of moving to numerator
  - Optional: multiply by 100 to get percentage

$$Sex\ ratio = \frac{Population\ of\ males}{Population\ of\ females}$$

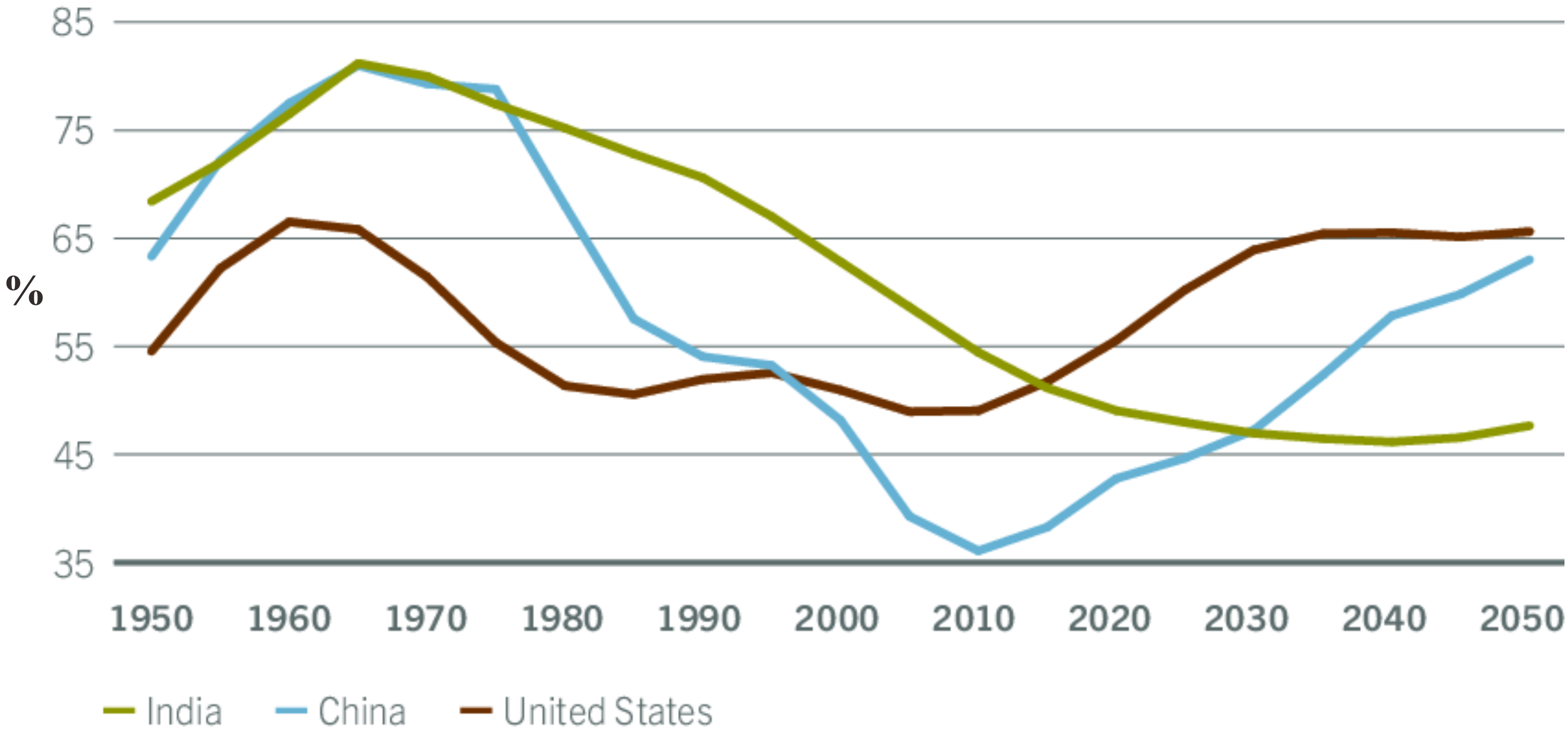
$$Total\ dependency\ ratio = \frac{Pop.\ children\ (0\ to\ 14)\ +\ Elderly\ pop.(65+)}{Working\ age\ population\ (15\ to\ 64)}$$

Sex ratios, 1950–2015



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/

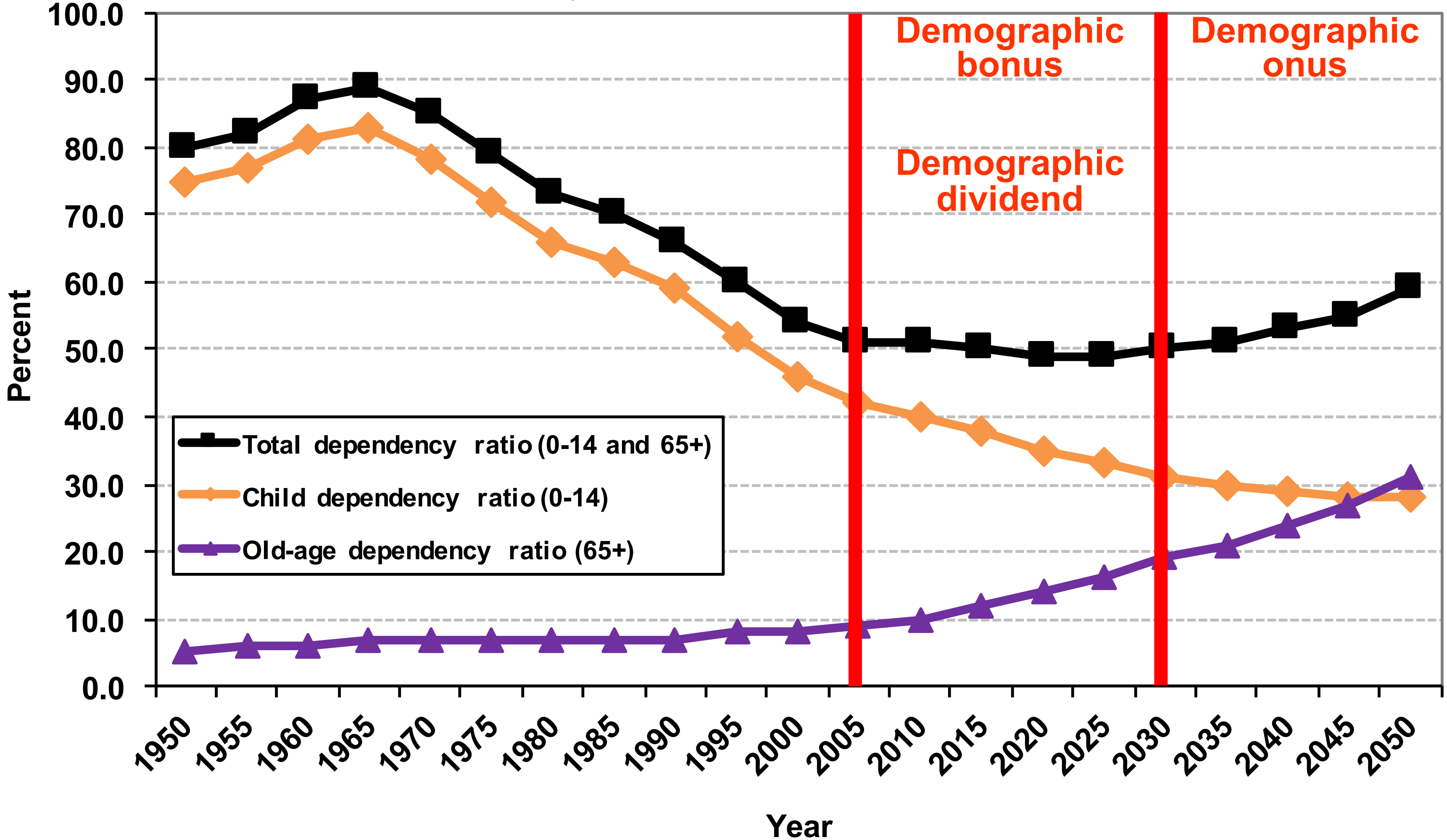
Total dependency ratios, India, China, United States



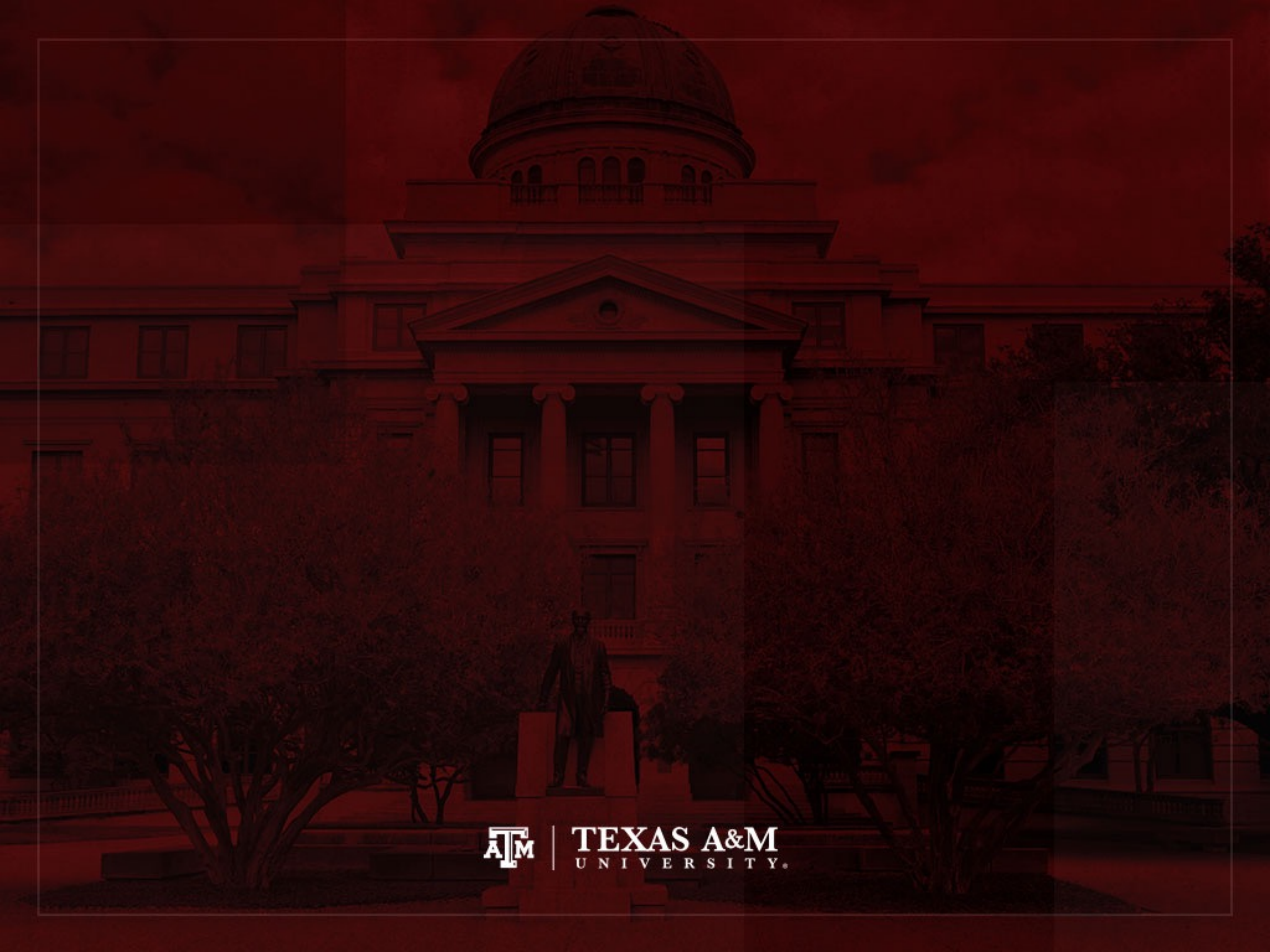
Source: United Nations Population Division

Source: https://www.researchgate.net/figure/Total-dependency-ratios-of-China-India-and-the-US-In_fig3_275071394.

Dependency ratios, Brazil, 1950–2050



Source: United Nations - http://esa.un.org/unpp (medium variant).


TEXAS A&M
UNIVERSITY

Crude rate model

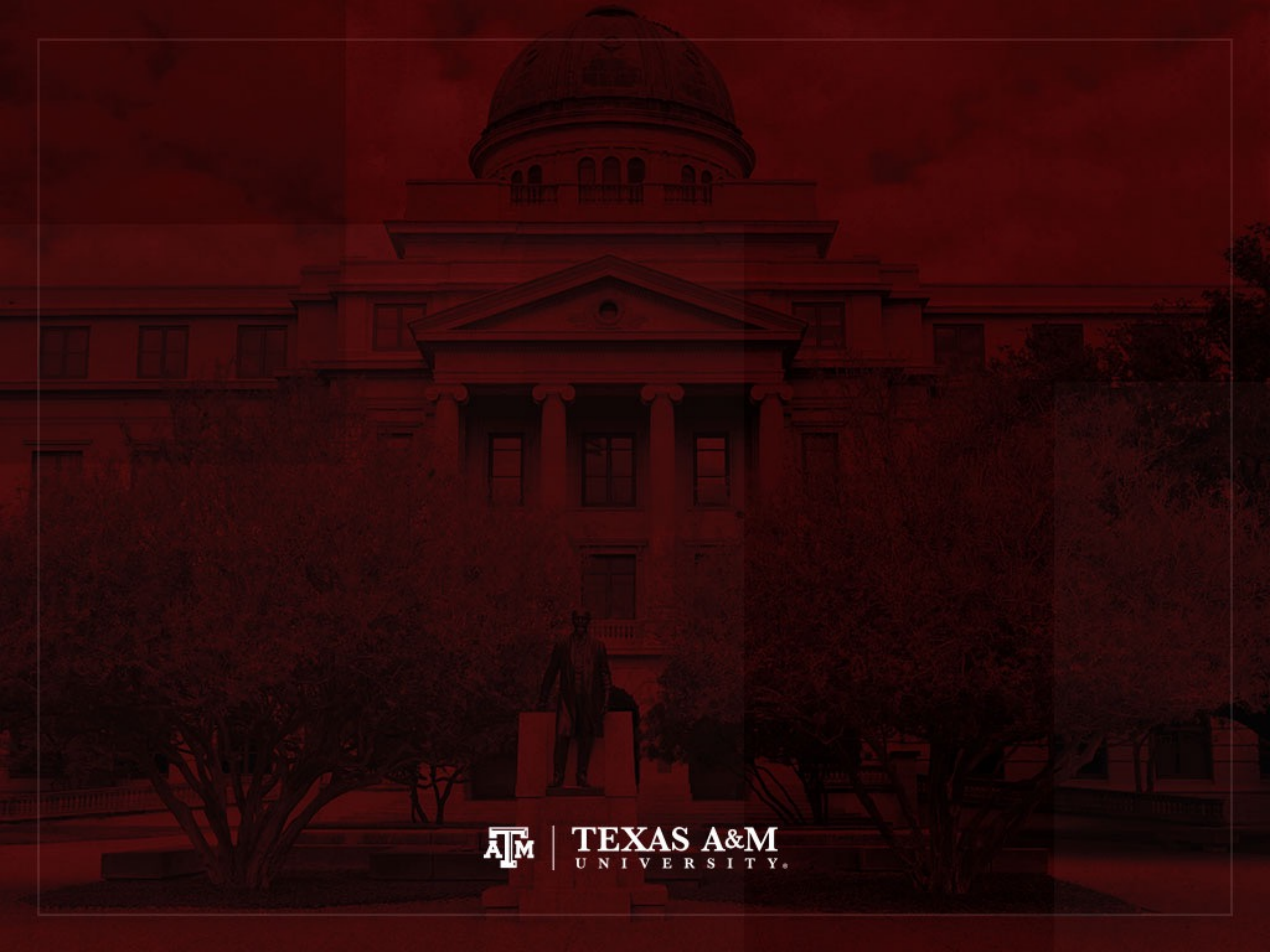
- Imagine a population
  - In which each person, each instant, is subject to constant independent risks of dying and having a baby
  - b: expected numbers of births per person per year
  - d: expected number of deaths per person per year
- Assumptions
  - Closed population
  - Homogeneous risks among people
  - No measurement of change over time inside the period

Growth rate

- Expected size of population has exponential growth
  - Growth rate = R = *b* – *d*
- Most actual populations are not closed and risks are not homogeneous over time
  - Need a measure of Crude Net Migration Rate (MIG)
  - Crude Growth Rate (CGR) = CBR – CDR + MIG

Most populous countries, 2012

Rank	Country	Pop. (million)	CBR (‰)	CDR (‰)	MIG (‰)	R (‰)	IMR (‰)	e_0
1	China	1,350	12	7	–0	5	17	73
2	India	1,260	22	7	–0	16	47	65
3	USA	314	13	8	+3	9	6	78
4	Indonesia	245	19	6	–1	12	29	71
5	Brazil	194	16	6	–0	11	20	73
6	Pakistan	188	28	8	–2	21	64	63
7	Nigeria	170	40	14	0	24	77	47
8	Bangladesh	153	23	6	–3	14	43	65
9	Russia	143	12	15	+2	–1	8	68
10	Japan	128	9	9	0	0	3	83
	World	**7,017**	**20**	**8**	**0**	**12**	**46**	**69**

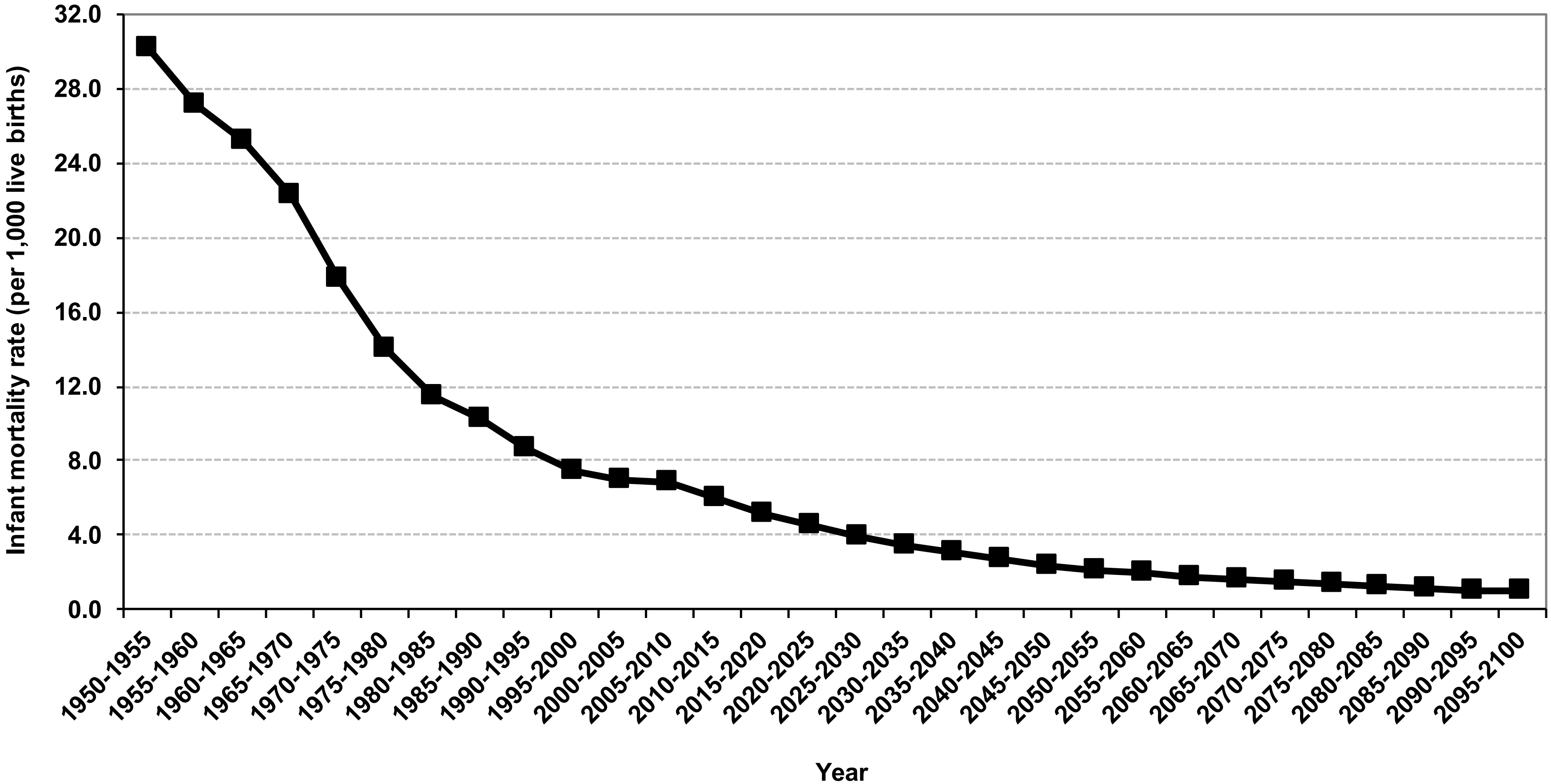

TEXAS A&M
UNIVERSITY

Infant mortality rate (IMR)

$$IMR = \frac{\textit{the number of deaths under age } 1 \textit{ in the period}}{\textit{the number of live births in the period}}$$

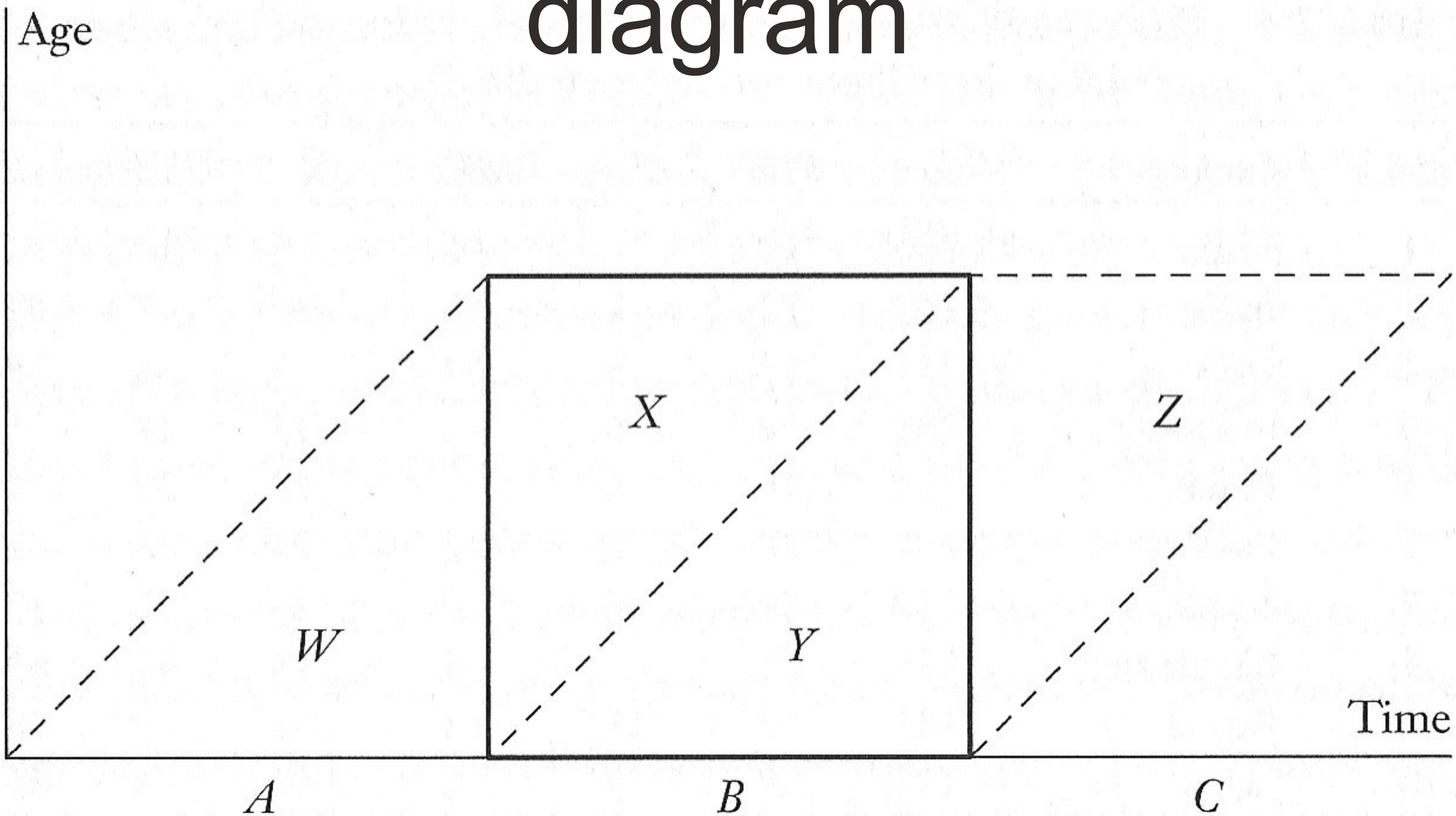
- IMR is a period measure
- It uses current information from vital registration
- It can be computed for countries without reliable census or other source for a count of the population at risk by age
- Infants born by teenagers and by older mothers are at higher risk

Infant mortality rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

IMR contributions on a Lexis diagram



Source: Wachter 2014, p. 38.

Understanding previous figure

- Any lifeline which ends within the square
  - Contributes a death to the numerator of the IMR

- Any lifeline that starts on the base of the square
  - Contributes a birth to the denominator of the IMR

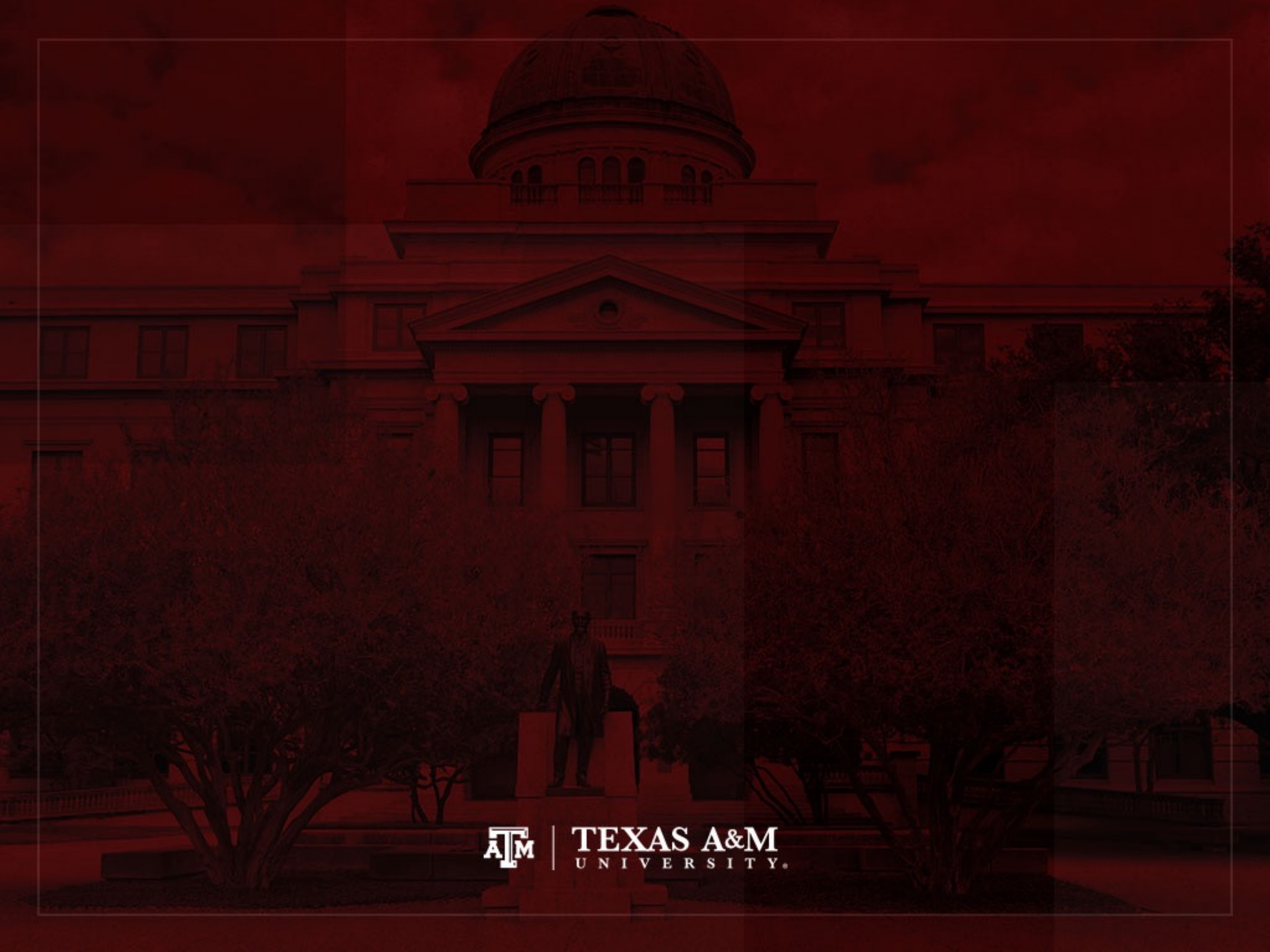
Still on previous figure

- Babies born outside the period in the preceding year (A) may die as infants during the period (X)
  - Counted in the numerator, but not in denominator
- Babies born during the period (B) may die after the end of the period (Z)
  - Counted in the denominator, but not in numerator
- Usually mismatched terms balance each other
  - IMR is close to the probability of dying before age 1

Period ≠ Cohort

- Period deaths and period person-years lived
  - Come from deaths and lifelines in the square (*X, Y*)
  - Dividing these deaths by person-years gives a period age-specific mortality rate (*M*)

- Cohort deaths and cohort person-years lived
  - Come from deaths and lifelines in parallelogram (*Y, Z*)
  - Dividing these deaths by person-years gives a cohort age-specific mortality rate (*m*)


TEXAS A&M
UNIVERSITY

Person-years and areas

- *PPYL* in the period between time 0 and time *T* is the area under the curve *K*(*t*) between 0 and *T*

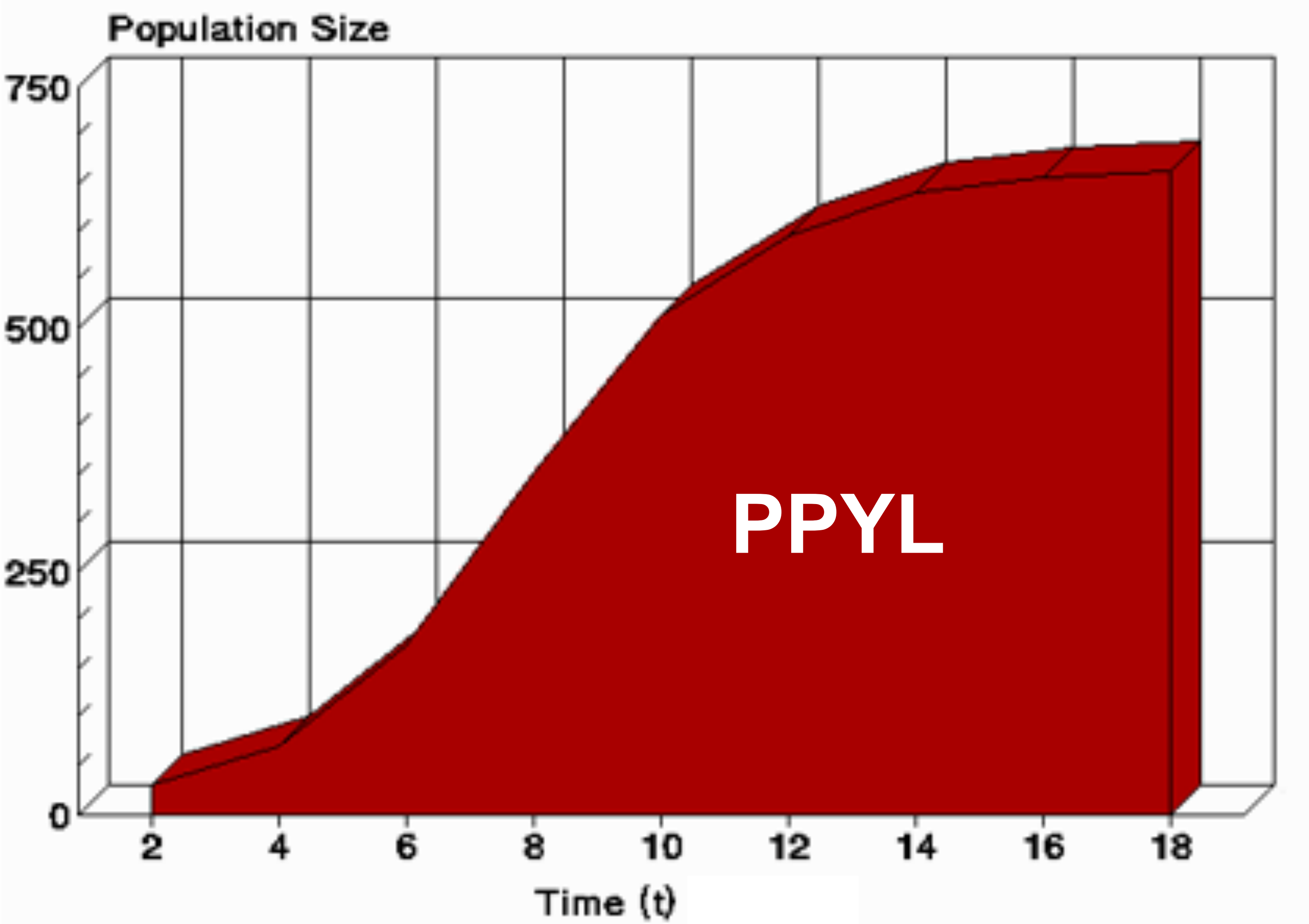
$$PPYL = \int_0^T K(t)dt$$

- *PPYL* between 0 and *T* when the exponential growth rate is constant

$$PPYL = K(0)(e^{RT} - 1) / R$$

$$PPYL = (K(T) - K(0)) / R = (K_T - K_0) / R$$

Person-years and areas



Source: https://www2.palomar.edu/users/warmstrong/lmexer9.htm.

First application

- How many person-years would have been lived altogether by members of the human race if growth rate had always been equal to 0.012 per year?

$$PPYL = (K(T) - K(0)) / R = (K_T - K_0) / R$$

$$PPYL = (8 - 0) / 0.012$$

$$PPYL \approx 667 \text{ billion person-years}$$

Second application

- Consider the *CBR* and *CDR* under the assumption that population size is growing exactly exponentially over the course of a year
- Population increases by

$$K(1) - K(0) = K(0)(e^R - 1) = B - D$$

- Difference between *CBR* and *CDR*

$$CBR - CDR = \frac{B - D}{PPYL} = \frac{B - D}{K(0)(e^{RT} - 1)/R} = \frac{B - D}{K(0)(e^{R} - 1)(1/R)} = \frac{1}{1/R} = R$$

- The growth rate (*R*) equals the difference between the crude birth rate and the crude death rate in a closed population subject to truly exponential growth

Third application

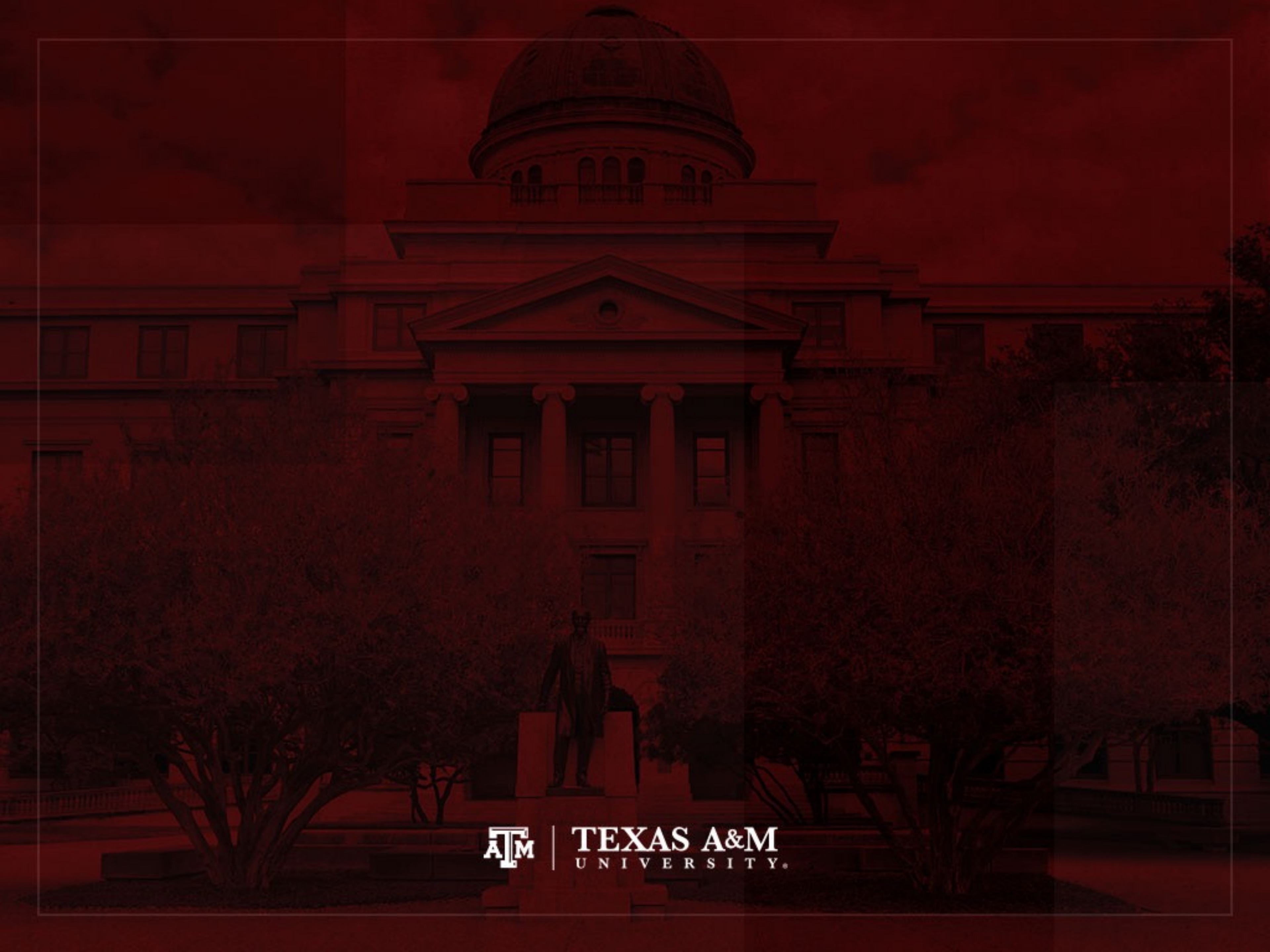
- Consider
  - Mid-period population: $K(T/2) = K_{T/2}$
  - Average population: $(K(0) + K(T))\ /\ 2 = (K_0 + K_T)\ /\ 2$
- *PPYL* can be approximated in terms of the mid-period population

$$PPYL = K(0)(e^{RT} - 1)\ /\ R \approx K_{T/2}T$$

- Or as average between initial and ending populations

$$PPYL \approx (K_0 + K_T)\ (T\ /\ 2)$$

- These expansions tell us the differences between the area formula, the mid-period approximation, and the average approximation as estimates of *PPYL*


TEXAS A&M
UNIVERSITY

Cohort person-years lived

- We get cohort person-years lived (*CPYL*) by adding up all person-years lived by all members of the cohort
  - Instead of counting people from a rectangle of the Lexis diagram, we consider a parallelogram
- If we divide *CPYL* by the total number of members of the cohort (counted at birth)
  - We obtain the expectation of life at birth (e_0)
  - Average number of person-years lived in their whole lifetimes by members of the cohort

Number of people in a period

- Calculation of number of people who lived over a specific period
  - Divide period person-years lived (*PPYL*)
    - Area under the curve of total population versus time over a specific period
  - By average lifespan (e_0) over the whole period

Example

- Number of people who lived between origins of farming (around 8000 B.C.) and birth of Christ (1 A.D.)?
  - Assumption: smooth exponential growth
  - Average lifespans (e_0): around 25 years
  - Data from Table 1.4 (page 25, Wachter, 2014)
    - Population in 8000 B.C.: 5 million = 0.005 billion
    - Population in 1 A.D.: 250 million = 0.250 billion
    - Growth rate: 0.000489
- Period person-years lived

$$PPYL = (K_T - K_0) / \mathrm{R} = (0.250 - 0.005) / 0.000489$$

$$PPYL = 501 \text{ billion person-years}$$

- Number of people who lived over this period

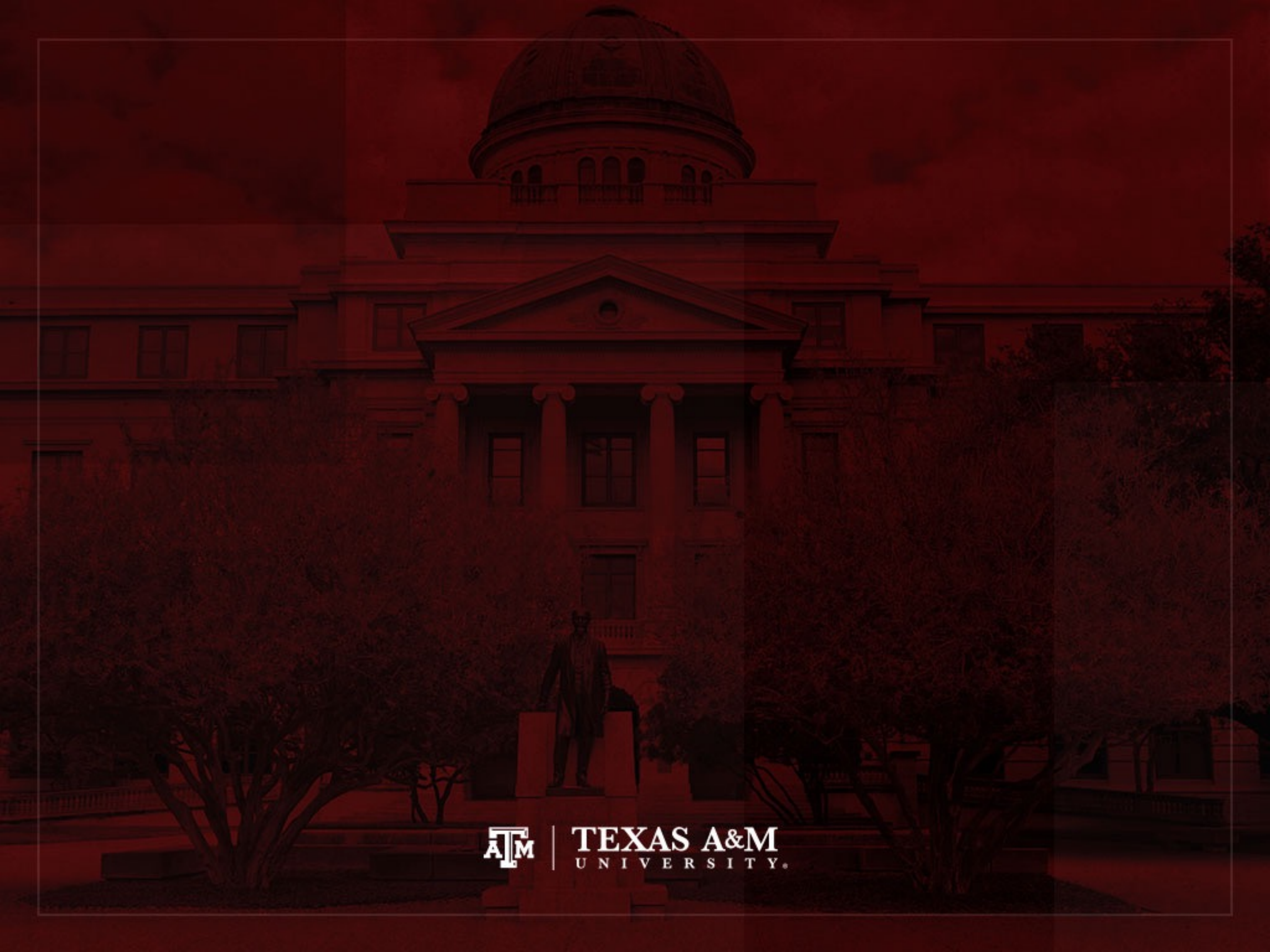
$$PPYL / e_0 = 501 / 25 \approx 20 \text{ billion people}$$

References

Fleurence RL, Hollenbeak CS. 2007. “Rates and probabilities in economic modelling: Transformation, translation and appropriate application.” Pharmacoeconomics, 25(1): 3–6.

Wachter KW. 2014. Essential Demographic Methods. Cambridge: Harvard University Press. Chapter 2 (pp. 30–47).

Weeks JR. 2015. Population: An Introduction to Concepts and Issues. Boston: Cengage Learning. 12th edition. Chapter 7 (pp. 251–297).


TEXAS A&M
UNIVERSITY