JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

JOHNS HOPKINS
UNIVERSITY

Life Table

Stan Becker, PhD
Bloomberg School of Public Health

JOHNS HOPKINS UNIVERSITY

Section A

Use, Types, Graphs, and Interpretation of the Columns

Introduction

- Although life tables were developed for the study of mortality, they can be applied to any other “failure” process so long as:
  - The process is measured in time, i.e., varies with age or some other measure of duration
  - The starting point of exposure can be defined unambiguously
  - The failure is an unambiguous event with a clear location in time

Continued

Introduction

- Life tables are used by demographers, public health workers, actuaries, and many others in studies of mortality, longevity, fertility, marriage, migration, and population growth, as well as in making projections of population, and in many other areas

Continued

Introduction

- Life tables are of particular value in the use of data that may be affected by censoring or loss to follow-up
- The life table approach can make efficient use of partial data by including observations up to the time of censoring

History

- John Graunt's *Natural and Political Observations upon the Bills of Mortality* (1662)
  - Implications of information obtained from christening and death lists in London
- Edmund Halley's presentation of the Breslau (Wroclaw, Poland) life table (1693)

Continued

History

- Antoine Deparcieux’s series of life tables for annuitants and monastic orders (1746)
- Lowell Reed and Margaret Merrell’s new technique to estimate the life table probability of dying within an interval from the age-specific death rates (1939)

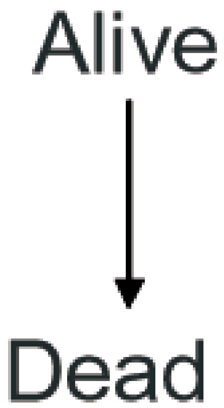
Continued

History

- T.N.E. Greville’s development of another technique to estimate the life table probability of dying within an interval from the age-specific death rate (1943)

Types of Life Tables

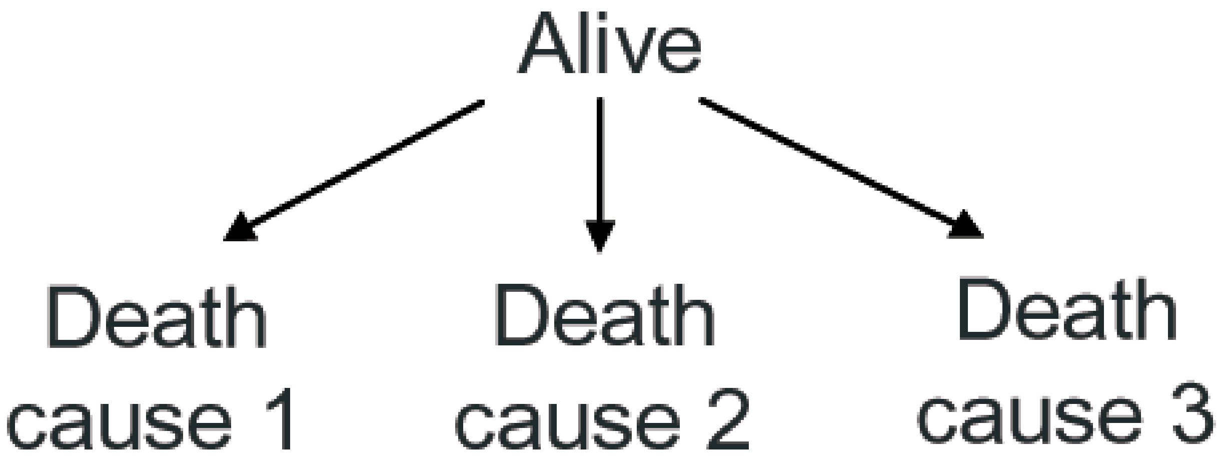
- Single decrement



Continued

Types of Life Tables

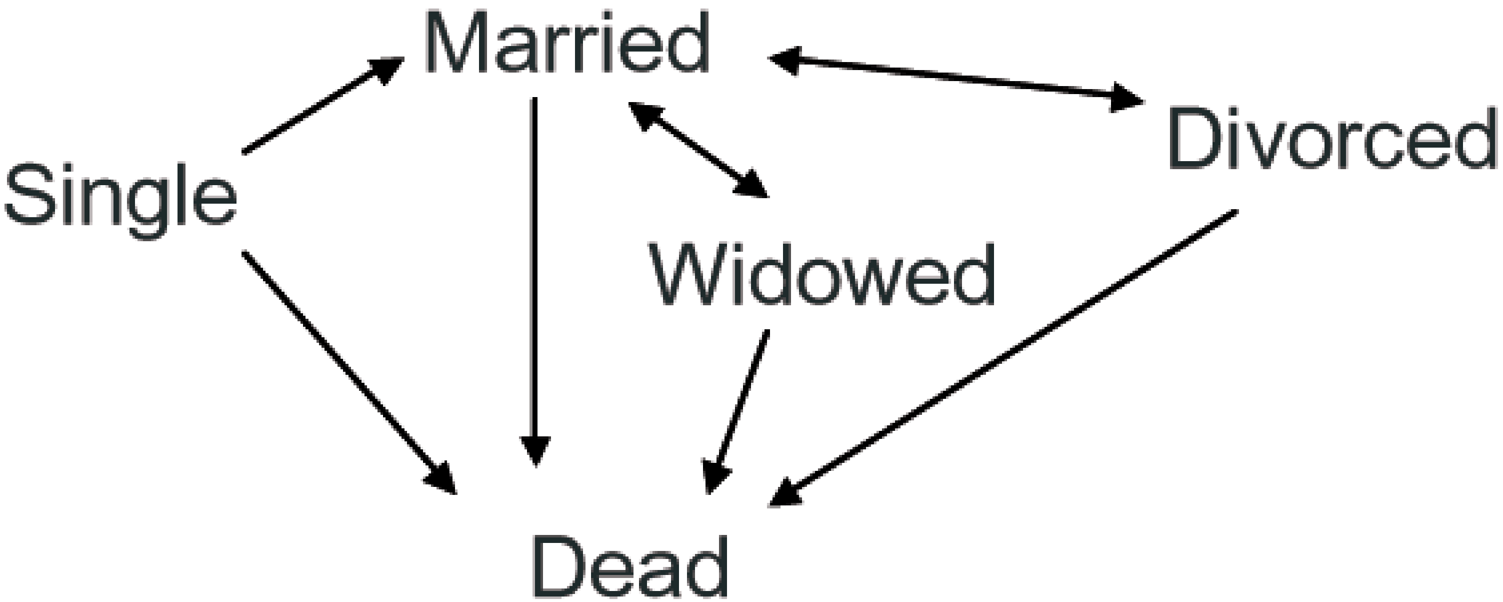
- Multiple decrement



Continued

Types of Life Tables

- Multi-state and increment-decrement



Conventional Life Table

- The life table conceptually traces a cohort of newborn babies through their entire life under the assumption that they are subject to the current observed schedule of age-specific mortality rates
- As age increases, the number of survivors of the original group declines, the decline being more rapid at ages where mortality rates are high

Columns of a Life Table

- The basic columns of a life table are:

Age interval (years)	$_nq_x$	ℓ_x	$_nd_x$	$_nL_x$	T_x	e_x
(x, x+n)						

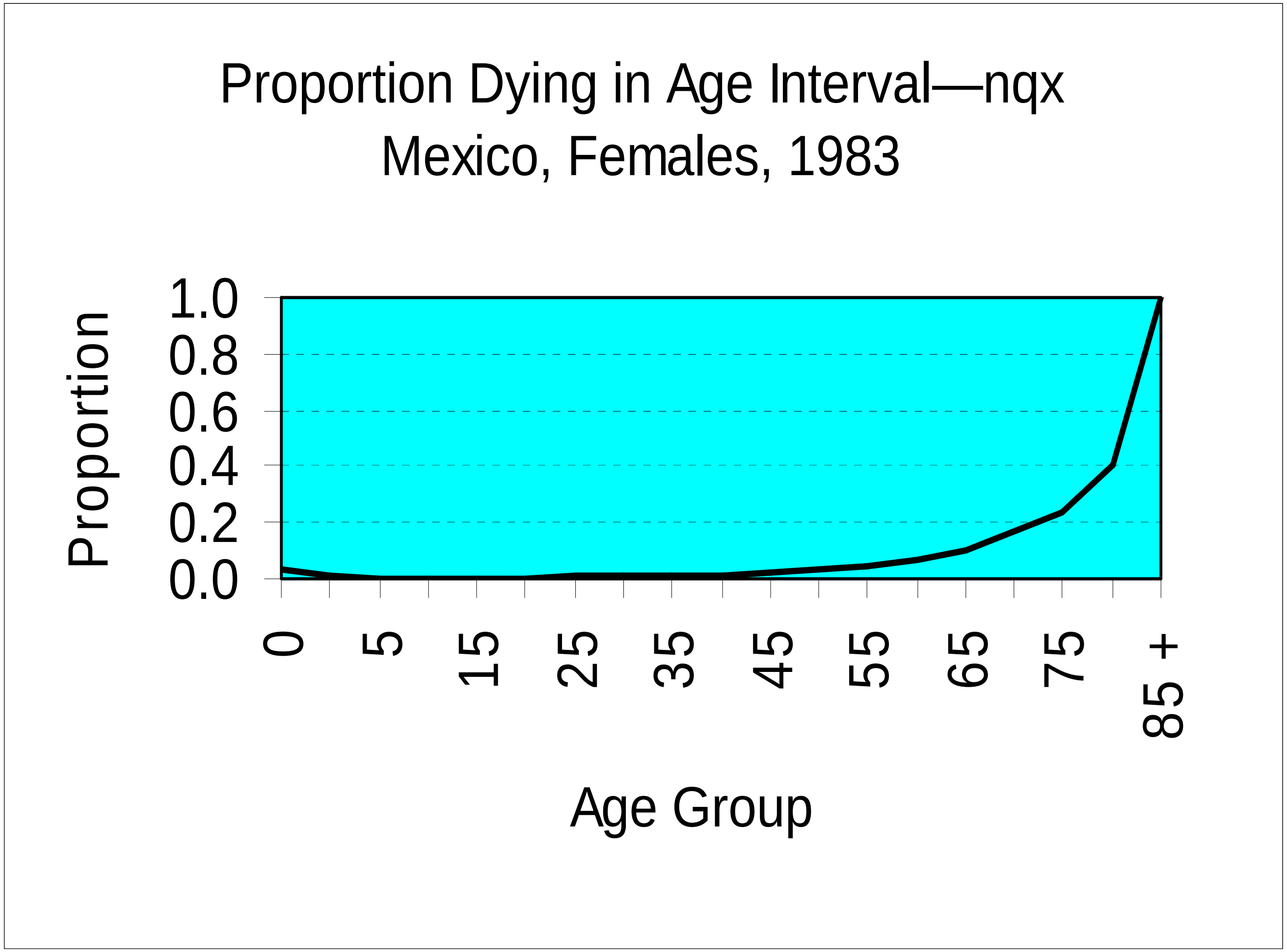
Conventional Life Table

- Note on life table notation:
  - The subscript before a symbol defines the width of the interval
  - The subscript after the symbol defines the starting point of the interval
  - ℓ_x, T_x and e_x never have a left subscript as they refer to exact age x

Description of Columns

- Age interval
  - Age interval from exact age x to age x+n, i.e., [x, x+n]
- ${}_{n}q_{x}$
  - (Conditional) probability of dying in the interval [x, x+n], given survival to age x

Conventional Life Table



Source: Adapted from Keyfitz and Fleiger, 1990

Conventional Life Table

- ℓ_x or $\ell(x)$
  - The first entry, for age 0, is called the radix
  - The ℓ_x column indicates the probability of survival to exact age x if the radix is set at 1, or if the value of ℓ_x is divided by the radix
  - i.e., $p(x) = \ell(x) / \ell(0)$

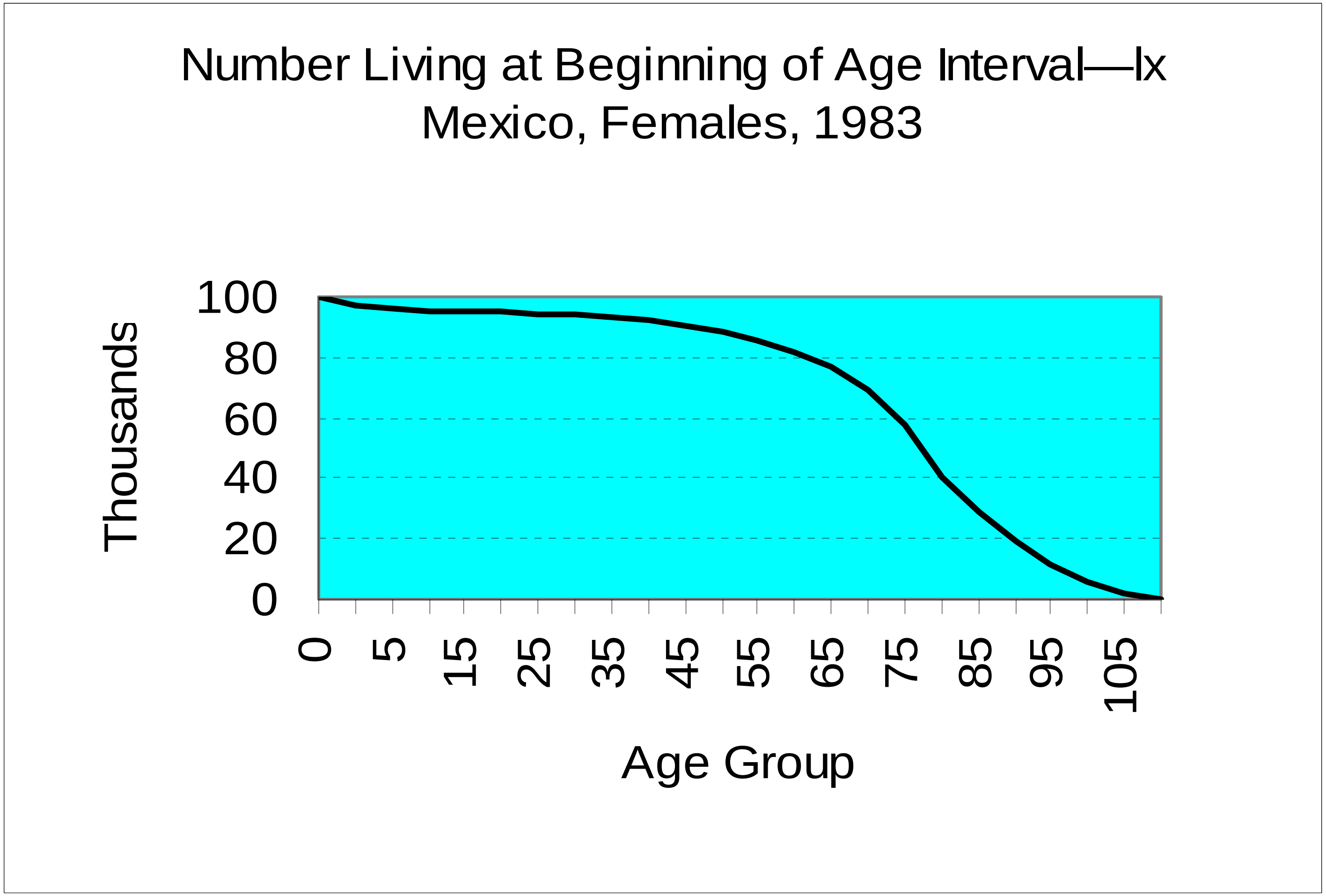
Continued

Conventional Life Table

- Alternatively, the ℓ_x column can also indicate survivorship if the radix is set to represent the number of initial births—the column then shows the number of the radix still alive by any exact age x
  - ℓ_{20} = Number of survivors at exact age 20 (their 20th birthday)
- In this case, the radix is usually set at 100,000

Continued

Conventional Life Table



Source: Adapted from Keyfitz and Fleiger, 1990

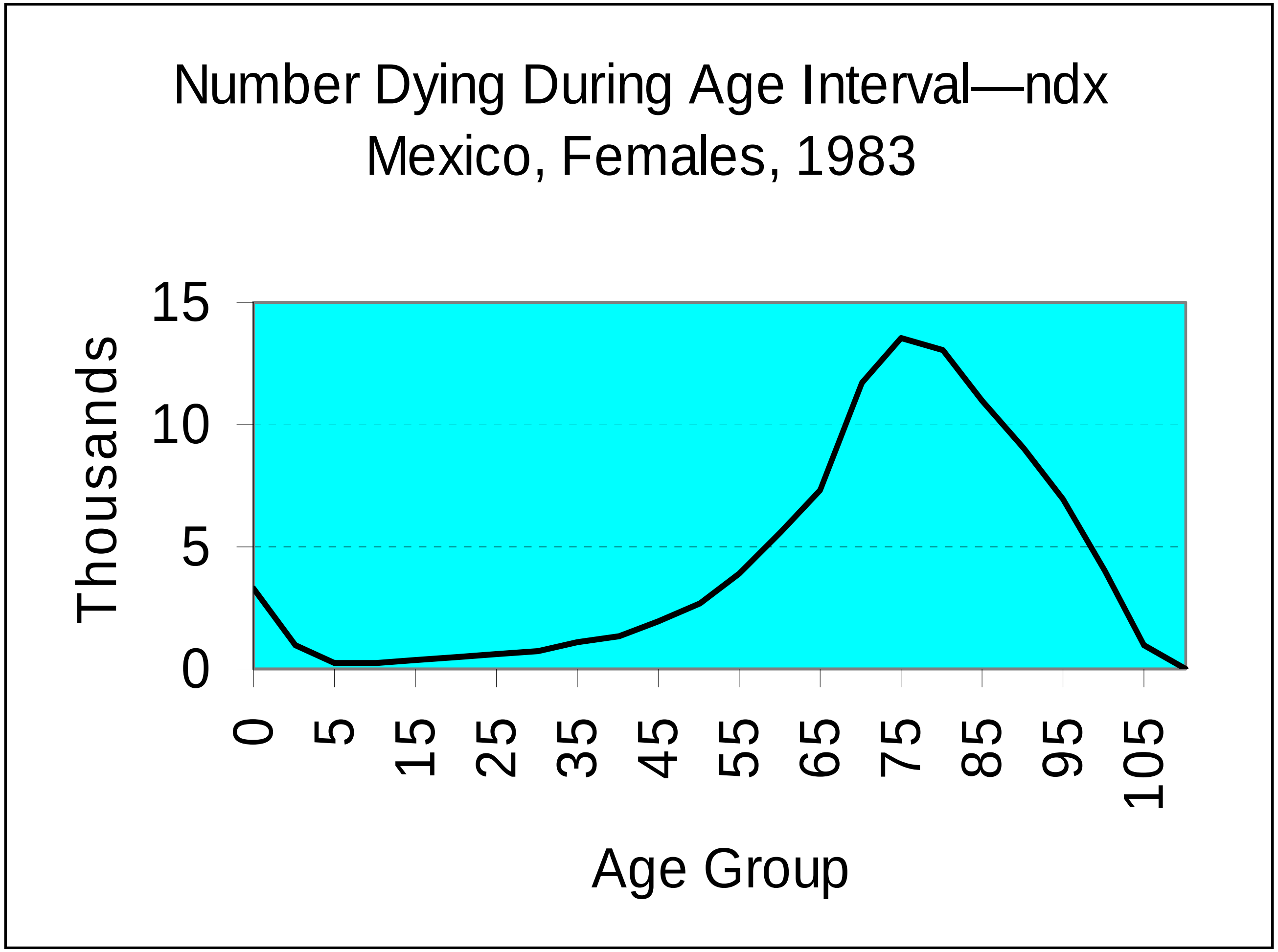
Continued

Conventional Life Table

- $_nd_x$ or $d(x,n)$
  - Number of deaths to the radix between exact ages x and x+n
  - Since all members of the radix die sooner or later, the sum of the $_nd_x$ over all ages is equal to the radix
  - If radix = 1, $_nd_x$ = (unconditional) probability of dying between ages x and x+n

Continued

Conventional Life Table



Source: Adapted from Keyfitz and Fleiger, 1990

Continued

Conventional Life Table

- ${}_nL_x$ or $L(x,n)$
  - Person-years lived between exact ages x and x+n
    - If radix = 1, ${}_nL_x$ = “expected” years lived between ages x and x+n
- T_x or $T(x)$
  - Person-years lived above age x
    - If radix = 1, T_x = “expected” years lived after age x

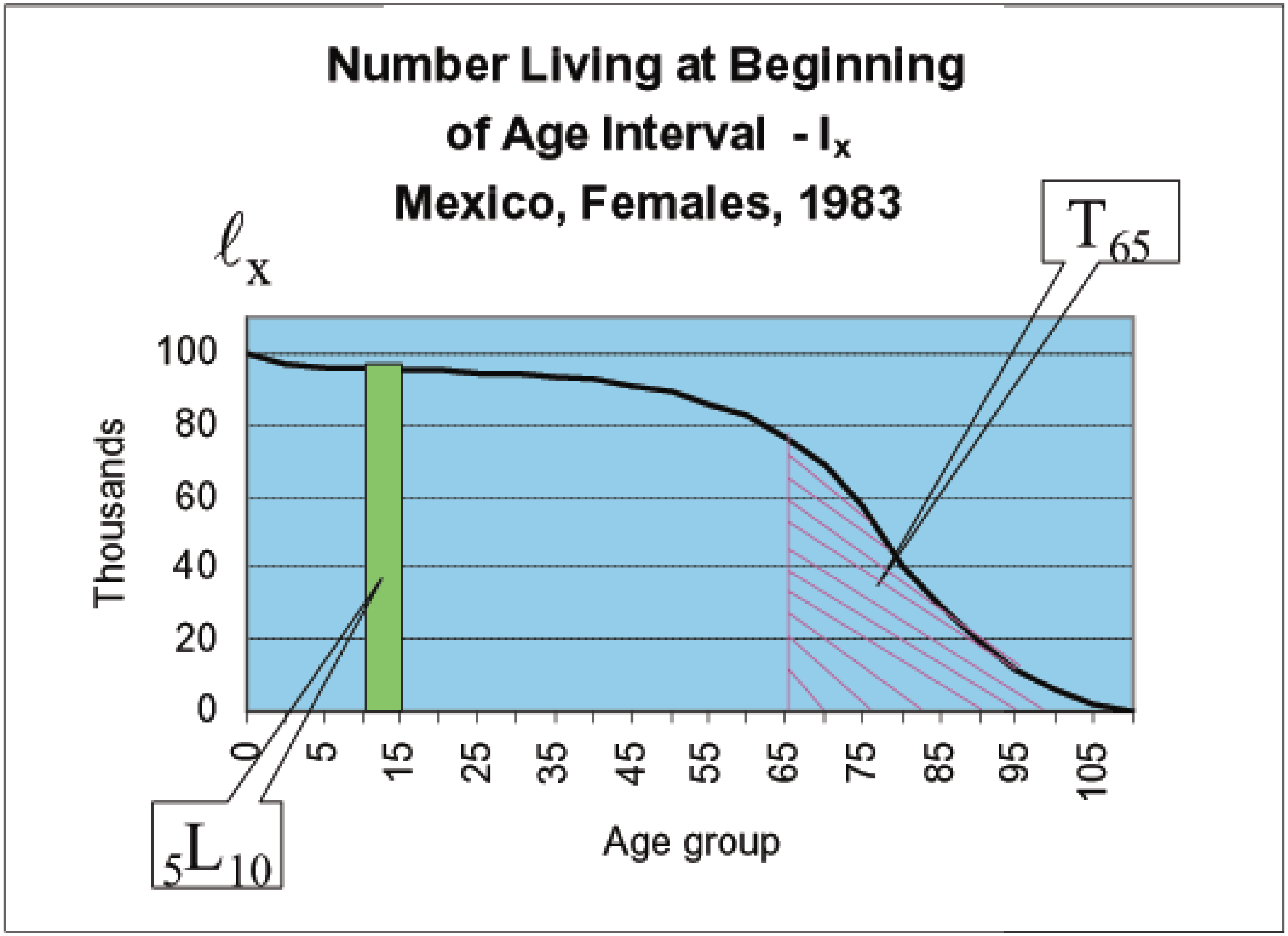
Continued

Conventional Life Table

- Note: ${}_nL_x$ and T_x are not expectations in the usual sense since deaths before age x are implicitly included

Continued

Conventional Life Table



Source: Adapted from Keyfitz and Fleiger, 1990

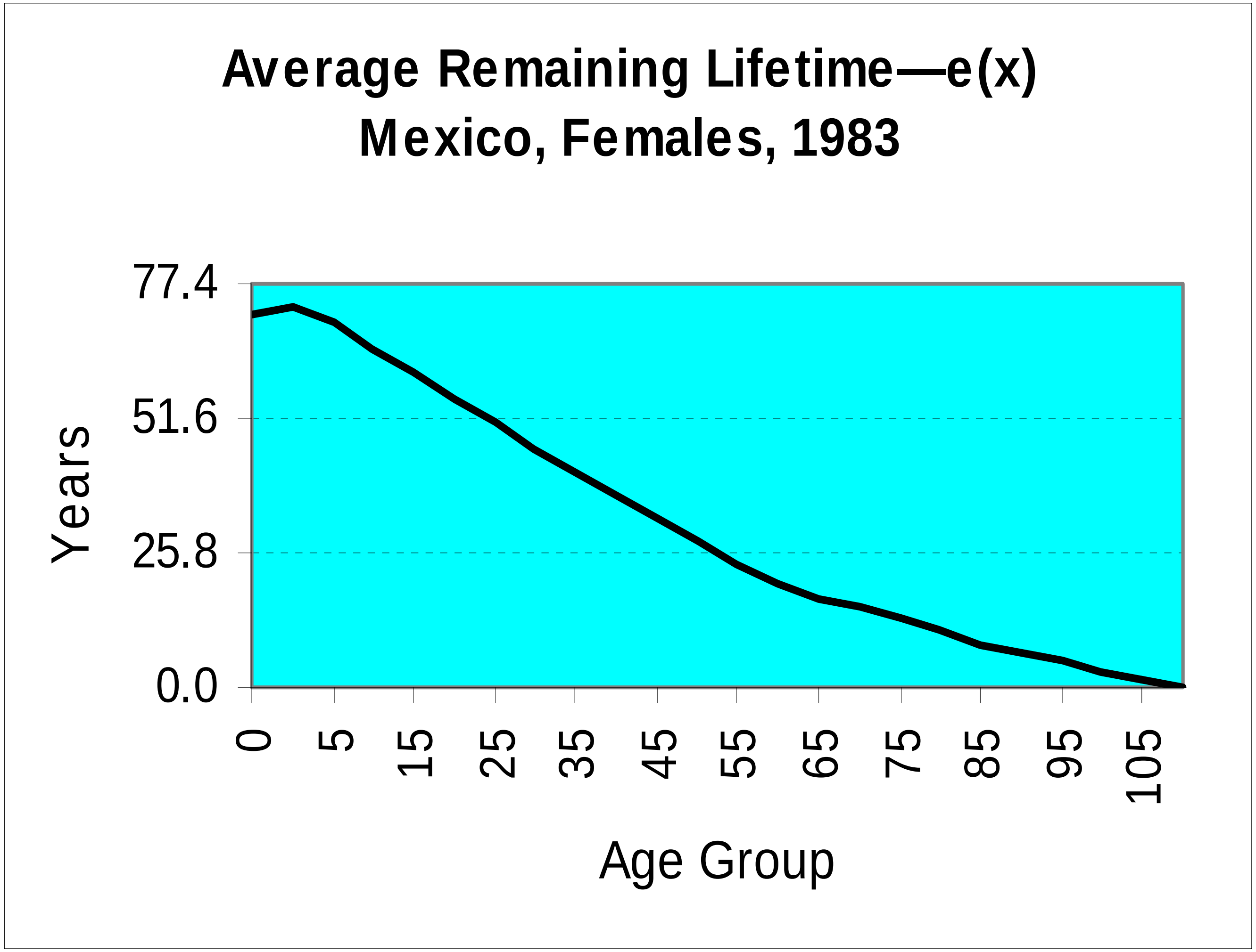
Continued

Conventional Life Table

- e_x or e(x)
  - Expectation of life at age x = Average remaining lifetime for a person who survives to age x
    - e.g., e(0) = Life expectancy at birth = number of years a person can expect to live in his/her life

Continued

Conventional Life Table



Source: Adapted from Keyfitz and Fleiger, 1990

Continued

Conventional Life Table

- Notes:
  - The six functions are generally calculated and published for every life table; however, some columns may be omitted without a significant loss of information since the functions are interrelated and some can be directly calculated from the others

Continued

Conventional Life Table

- Notes:
  - In general, the conditional probability of death (${}_nq_x$) is the basic function in the life table

Mathematical Derivation and Relationships

- Mathematical derivation and relationships (cohort life table)
- Let X = Random variable for age at death

  f(X) = Probability density function of X

  F(X) = Cumulative density function of X

$$= \int_0^X f(y)\,dy$$

Continued

Mathematical Derivation and Relationships

- $P\{X > x\}$ $$= \int_{x}^{\infty} f(y)dy = 1 - F(x)$$ $$= \ell(x)$$
- $P\{x < X < x+n\}$ $$= \int_{x}^{x+n} f(y)dy = F(x+n) - F(x)$$ $$= \ell(x) - \ell(x+n)$$ $$= d(x,n)$$

Continued

Mathematical Derivation and Relationships

- $P\{x<X<x+n|X>x\}$ $= \dfrac{P\{x < X < x + n\}}{P\{X > x\}} = \dfrac{d(x,n)}{\ell(x)}$ $= q(x,n)$

- $P(x,n) = 1\text{-}q(x,n)$ $= \dfrac{\ell(x,n)}{\ell(x)}$

- Note: $P(x,0) = \ell(x)$

Continued

Mathematical Derivation and Relationships

$$T(x) = \int_{x}^{\infty} \ell(y)dy$$

$$L(x,n) = \int_{x}^{x+n} \ell(y)dy = T(x) - T(x+n)$$

$$T(x) = \sum_{y=x}^{\infty} L(y,n)$$

$$E(0) = \int_{0}^{\infty} xf(x)dx = \int_{0}^{\infty} \ell(y)dy = T_0 = e^0(0)$$

Continued

Mathematical Derivation and Relationships

$$E(X \mid X > x) = x + \frac{\int_x^\infty \ell(y)dy}{\ell(x)}$$

$$= x + \frac{T(x)}{\ell(x)}$$

$$E(X - x \mid X > x) = \frac{T(x)}{\ell(x)} = e^0(x)$$

Continued

Mathematical Derivation and Relationships

$$E(X-x \mid x < X > x+n) = \frac{\int_{t=x}^{x+n} tf(t)dt}{\int_{t=x}^{x+n} f(t)dt} - x$$

$$= \frac{L(x,n) - n * \ell(x+n)}{d(x,n)}$$

$$= a(x,n)$$

Average Life Lived in Interval

by Those Dying in Age Interval [x,x+n]

$$a(x,n) = {}_n a_x = \frac{\sum_{i=1}^{{}_n d_x} (X_i - x)}{{}_n d_x}$$

- Where X_i = Exact age at which a person dies

  x = Age at beginning of interval

Continued

Average Life Lived in Interval

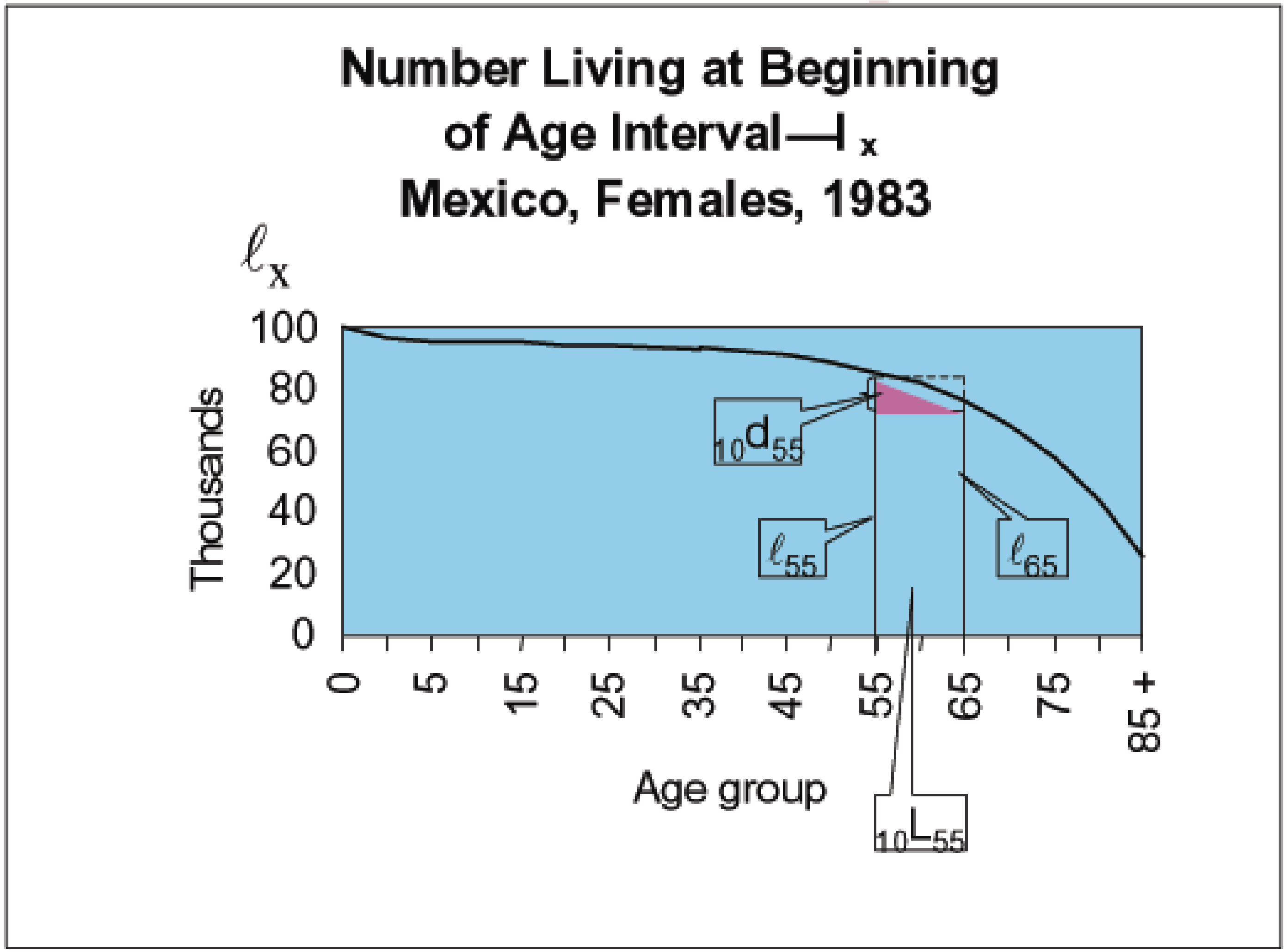
by Those Dying in Age Interval [x,x+n]

$$ {}_nL_x = n * \ell_{x+n} + {}_na_x * {}_nd_x $$

- Assuming that deaths are linearly distributed in age interval, then:

$$ {}_nL_x = n * \ell_{x+n} + \frac{n}{2} * {}_nd_x $$

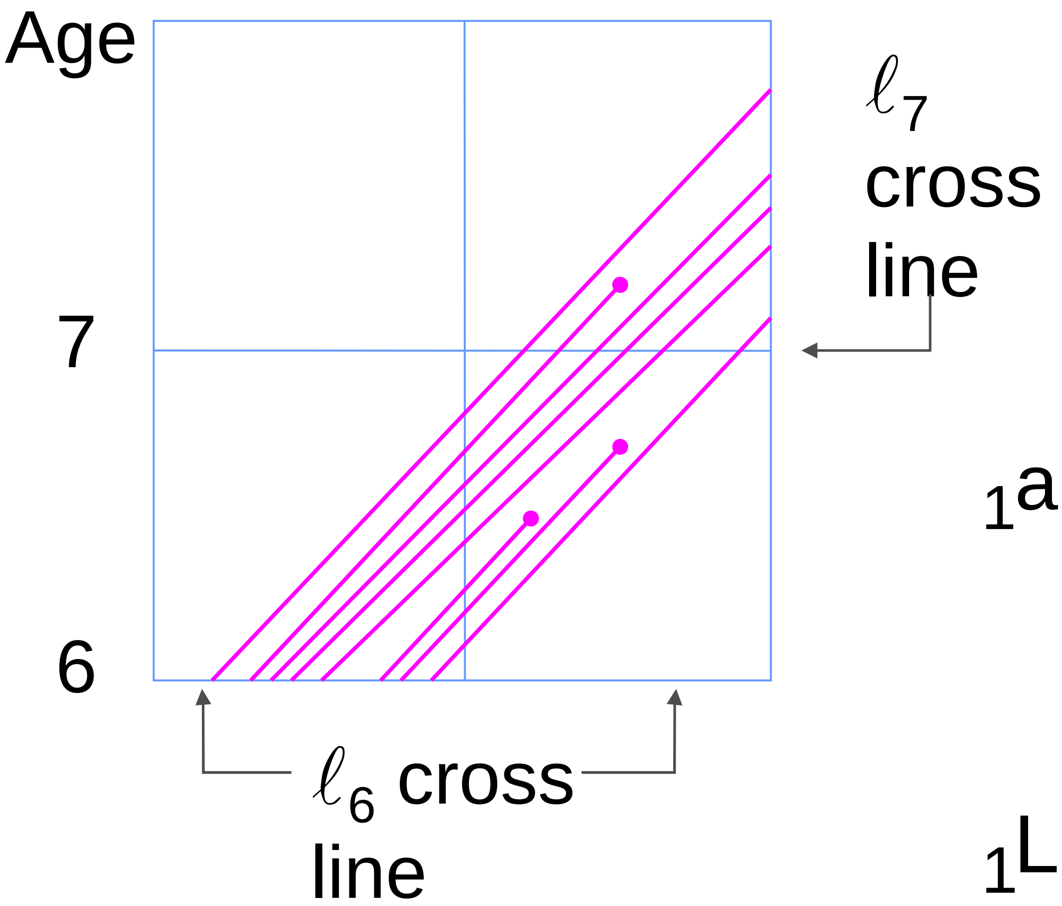
Mathematical Derivation and Relationships



Source: Adapted from Keyfitz and Fleiger, 1990

Continued

Mathematical Derivation and Relationships



$\ell_0 = 10 \qquad {}_1d_6 = 2$

$\ell_6 = 8 \qquad \ell_7 = 6$

$$ {}_1a_6 = \frac{(6.5-6)+(6.7-6)}{2} $$

$$ = 0.6 $$

$$ {}_1L_6 = 6 + 0.6 * 2 $$

$$ = 7.2 $$

Summary

- There are three main categories of life tables: Decrement only vs. increment-decrement, single decrement vs. multiple decrement, and single state vs. multi-state
- The conventional life table traces a cohort of newborn babies as they go through life assuming that they are subject to the current observed schedule of age-specific mortality rates

Continued

Summary

- There are seven basic columns in a life table
- The six functions are generally calculated and published for every life table, however, some columns may be omitted without significant loss of information since the functions are interrelated and some can be directly calculated from the others

JOHNS HOPKINS
UNIVERSITY

Section B

Construction of a Life Table and Use of Survivor Ratios

Construction

Conventional Life Table

- Conventional life table
  - Assumption: Data on births, deaths, and population are accurate
  - However, one of the most important aspects of the preparation of a life table is the testing of the data for possible biases and other errors; the level of accuracy that can be tolerated depends mostly on the intended use of the life table

Continued

Construction

Conventional Life Table

- Excessive smoothing of the data by mathematical methods would eliminate or reduce true variations in the age pattern of mortality rates
- Therefore, it must be decided whether the emphasis in the life table should be on its closeness to the actual data or on its presentation of the underlying mortality picture after fluctuation has been removed

Continued

Construction

Conventional Life Table

- Life table functions, presumed for cohorts, are derived from period mortality data

Steps
Age-Specific Death Rates

- Steps
  - Derive the probability of dying ($_nq_x$) from the age-specific death rates by one of various methods
  - Calculate each entry of the $\ell(x)$ and d(x,n) columns
  - Derive $_nL_x$, T_x and e_x columns

Relation of ${}_nM_x$ and ${}_nm_x$

- Let D = Observed number of deaths in age group x, x+n

  P = Mid-point population in age group x, x+n

  ${}_nd_x$ = Number of deaths in life table

  ${}_nL_x$ = Life table population in age group x, x+n

Observed Age-Specific Death Rate

- ${}_nM_x$ = Observed age-specific death rate in age group x, x+n

$$= \frac{D}{P}$$

- ${}_nm_x$ = Life table death rate in age group x, x+n

$$= \frac{{}_nd_x}{{}_nL_x}$$

Finding ${}_nq_x$ from ${}_nm_x$

- Finding ${}_nq_x$ from ${}_nm_x$

$$ {}_nq_x = \frac{{}_nd_x}{\ell_x} $$

$$ {}_nm_x = \frac{{}_nd_x}{{}_nL_x} \qquad \Rightarrow {}_nd_x = {}_nm_x * {}_nL_x $$

$$ \text{so } {}_nq_x = \frac{{}_nm_x * {}_nL_x}{\ell_x} $$

Construction

$$\text{since } {}_nL_x = {}_na_x * \ell_x + (n - {}_na_x) * \ell_{x+n}$$

$$ {}_nq_x = \frac{{}_nm_x * [{}_na_x * \ell_x + (n - {}_na_x) * \ell_{x+n}]}{\ell_x}$$

$$= {}_nm_x * {}_na_x + {}_nm_x * (n - {}_na_x) * \frac{\ell_{x+n}}{\ell_x}$$

$$\text{since } {}_nq_x = 1 - \frac{\ell_{x+n}}{\ell_x}$$

Continued

Construction

$$ {}_nq_x = {}_nm_x * {}_na_x + {}_nm_x * (n - {}_na_x) * (1 - {}_nq_x) $$

Rearranging and bringing q to other side,

$$ {}_nq_x = \frac{{}_nm_x * {}_na_x + n * {}_nm_x - {}_nm_x * {}_na_x}{[1 + {}_nm_x * (n - {}_na_x)]} $$

$$ = \frac{n * {}_nm_x}{1 + (n - {}_na_x) * {}_nm_x} $$

Continued

Construction

- When n=1 and ${}_n a_x = 0.5$,

$$q_x = \frac{2 * m_x}{2 + m_x} \Rightarrow q_x < {}_1 m_x$$

- When n=5 and ${}_n a_x = 2.5$

$${}_5 q_x = \frac{5 * m_x}{1 + 2.5 * m_x} \Rightarrow {}_5 q_x > {}_5 m_x$$

Abridged Life Table

- The abridged life table is less burdensome to prepare and more convenient to use
- Contains data by intervals of five to ten years of age
  - Except the zero to one and one to four age groups, which are retained separately
- Values for five or ten year intervals are sufficiently accurate for most purposes

Constant Risk Method

- Finding ${}_nq_x$ from M_x
  - Assumption: The force of mortality is constant in (x, x+n) and equal to ${}_nM_x$

$$ {}_nq_x = 1 - e^{-n*({}_nM_x)} $$

  - Note: Constant risk of deaths $\Rightarrow$ more deaths at beginning of interval

Actuarial Method

- Assumption: Population distribution is the same as life table distribution and life table distribution is linear in (x, x+n)

$$ {}_nq_x = \frac{n * {}_nM_x}{1 + \frac{n}{2} * {}_nM_x} $$

- Note: Same number of deaths through the interval

Reed-Merell's Method

- Assumption:

$${}_nM_x = {}_nm_x$$

empirical fit to find "a"

$${}_nq_x = 1 - e^{-n * {}_nm_x - a * n^3 * {}_nm_x^2}$$

- Note: a = 0.008 takes lack of fitness into account

Greville's Method

- Assumption: ${}_nM_x = {}_nm_x$

  $\ell(x)$ is fit by cubic equation

$$ {}_nq_x = \frac{{}_nm_x}{\frac{1}{n} + {}_nm_x * \left[\frac{1}{2} + \frac{n}{12} * \left({}_nm_x - \log_e c \right) \right]} $$

- Note: $\log_e c$ could be assumed to be about 0.095

Sirken's Method

- Assumption: ${}_ng_x$ is derived from a complete life table

$$ {}_nq_x = \frac{n * {}_nM_x}{1 + {}_ng_x * {}_nM_x} $$

- Note: Since this method obtains the new table by reference to a standard table, it should only be used when mortality in both tables is of a comparable level

Relation of ${}_n g_x$ to ${}_n a_x$

$$ {}_n g_x = \frac{n}{{}_n q_x} - \frac{1}{{}_n m_x} = \frac{n}{\frac{{}_n d_x}{\ell_x}} - \frac{1}{\frac{{}_n d_x}{{}_n L_x}} $$

$$ = \frac{n * \ell_x}{{}_n d_x} - \frac{{}_n L_x}{{}_n d_x} = \frac{n * \ell_x - {}_n L_x}{{}_n d_x} $$

$$ = n - {}_n a_x $$

$$ \text{Note}: n - {}_n g_x = {}_n a_x $$

Chiang's Method

- Assumption: Population distribution is the same as life table distribution
- Need ${}_na_x$

$$ {}_nq_x = \frac{n * {}_nm_x}{1 + (1 - {}_na_x) * n * {}_nm_x} $$

- Note: Chiang defines ${}_na_x$ as the fraction of an interval lived by those dying

Survival Ratios

- One can use the life table to either "project" or "reverse survive" populations

Continued

Survival Ratios

- Project or look forward
- Proportion surviving from birth to an age interval [x, x+n]:

$$\frac{{}_{n}L_{x}}{n\ell_{0}} \quad \text{or} \quad \frac{\ell_{x+n/2}}{\ell_{0}}$$

- e.g., Proportion surviving from birth to age 20–24

$$\frac{{}_{5}L_{20}}{5\ell_{0}} \quad \text{or} \quad \frac{\ell_{22.5}}{\ell_{0}}$$

Continued

Survival Ratios

- Proportion surviving from an age interval [x, x+n] to a fixed age "a":

$$\frac{n * \ell_a}{{}_nL_x}$$

- e.g., Proportion of 60-year-olds who survive to their 65th birthday

$$\frac{\ell_{65}}{{}_1L_{60}}$$

Continued

Survival Ratios

- Proportion surviving from an age interval [x, x+n] to an age interval [y, y+n]:

$$\frac{{}_nL_y}{{}_nL_x}$$

- e.g., Proportion surviving from age group 30–34 to age group 35–39

$$\frac{{}_5L_{35}}{{}_5L_{30}}$$

Continued

Survival Ratios

- Reverse survive to recover a previous population from a present one:
  - i.e., Persons age y now (P_y) represent how many persons age x, y–x years ago

$$\frac{{}_nL_x}{{}_nL_y} * P_y$$

Continued

Survival Ratios

- Births represented by children now age two

$$\frac{\ell_0}{{}_1L_2} * P_2$$

- Births in one year represented by children age zero to four

$$\frac{\ell_0}{{}_5L_0} * P_{0-4}$$

Continued

Survival Ratios

- Persons age 60–64, 10 years ago represented by persons now 70–74:

$$\frac{{}_5L_{60}}{{}_5L_{70}} * P_{70-74}$$

Summary

- In general, the conditional probability of death, ${}_nq_x$, is the basic function from which all other functions are derived
- ${}_nq_x$ is derived from the age-specific death rates (there are various methods for performing this derivation)
- Life tables can be used to project or reverse survive population (in a projection, the survival ratio is <1, whereas in reverse survival, it is >1)

JOHNS HOPKINS
UNIVERSITY

Section C

Stationary Population and Model Life Tables

Stationary Population

- Characteristics
  - A stationary population is defined as a population whose total number and distribution by age do not change with time

Continued

Stationary Population

- Assuming no migration, such a hypothetical population can be obtained if the number of births per year remained constant (usually assumed at 100,000) for a long period of time and each cohort of births experienced the current observed mortality rates throughout life
- The annual number of deaths would thus equal 100,000 also, and there would be no change in size of the population

Continued

Stationary Population

- The characteristics of a stationary population are given by the corresponding life table

Interpretation of Columns

- The columns "age interval," "$_nq_x$," and "e_x" have the same interpretation as in the regular life table
- ℓ_x

  = Number of persons who reach the beginning of the age interval each year

Continued

Interpretation of Columns

- ${}_nd_x$
  = Number of persons who die each year within the age interval
- ${}_nL_x$
  = Number of persons in the population who at any moment are living within the age interval

Continued

Interpretation of Columns

- T_x
  = Number of persons who at any moment are living within the age interval and all higher age intervals
- T_0
  = Total population

Examples of Calculation

- Crude birth rate = crude death rate

$$= \frac{\ell_0}{T_0}$$

- Since ℓ_0= Number of births, but also number of deaths

Proportion of Population in Age Group

- Proportion of population in age group [x, x+n] among those above age "a"
  - Single-year life table

$$= \frac{\ell_{x+0.5}}{T_a} \qquad = \frac{{}_1L_x}{T_a}$$

  - Abridged life table

$$= \frac{{}_5L_x}{T_a}$$

Dependency Ratio

$$= \frac{P_{<15} + P_{65+}}{P_{15-64}}$$

$$= \frac{{}_{15}L_0 + {}_wL_{65}}{{}_{50}L_{15}}$$

- Where w = last age group

Mean and Median Age of Population

- Mean age of population

$$= \frac{\sum_{x=0}^{w} (x+0.5) * \ell_{x+0.5}}{T_0}$$

- *Median age of population*—age x such that $T_x = T_0 / 2$

Mean and Median Age at Death

- Mean age at death

$$= \sum_{x=0}^{w} (x + 0.5) * d_x = \frac{T_0}{\ell_0} = e_0$$

- *Median age at death*—age x such that ℓ_x = 0.5

Model Life Tables

- Model life tables were first developed in the 1950s by the United Nations
- These first models were single parameter models with no variation other than mortality level

Continued

Model Life Tables

- A number of more flexible model systems have since been developed
- Among them are the following:
  - Coale-Demeny Model Life Tables
  - Brass Logit Life Table System
  - United Nations Model Life Tables for Developing Countries

Coale-Demeny Model Life Tables

- Based on 192 observed and evaluated life tables from the 19th and 20th centuries, largely for European or European-origin populations
- A one-parameter system with 25 mortality levels from expectations of life from 20 years to 80 years by 2.5 year intervals for four distinct mortality age patterns called North, South, East, and West

Continued

Coale-Demeny Model Life Tables

- Each level of mortality has tables for males and females
- Thus, a given life table for one sex implies a corresponding table for the other sex, the difference in e(0) being an average of the differences observed in the observed tables
- The female advantage in e(0) ranges from about two years for the heaviest mortality (lowest levels) to about 3.5 years for the lowest mortality (highest levels)

Continued

Coale-Demeny Model Life Tables

- *Regional family*—North
  - Based on northern European tables
  - For given level of expectation of life at birth, have low infant and high child mortality, below expected mortality above age 50
  - Based on only nine life tables

Continued

Coale-Demeny Model Life Tables

- *Regional family*—South
  - Based on southern European tables
  - High mortality under age five (and high child mortality relative to infant mortality), low mortality from 40 to 65, high mortality above 65
  - Based on 22 observed life tables

Continued

Coale-Demeny Model Life Tables

- *Regional family*—East
  - Based on eastern European tables
  - Very high mortality in infancy relative to child and adolescent mortality, high mortality again after age 50
  - Based on 31 observed tables

Continued

Coale-Demeny Model Life Tables

- *Regional family*—West
  - Based on all other tables i.e., 130 tables, including some from non European populations (e.g. East Asia)
  - Standard against which other families are compared
  - As the residual pattern, ‘West” is sometimes recommended as the family to choose if no other basis for choice exists

Brass Logit Life Table System

- Based on the expectation that, on a suitably transformed age scale, the survivorship function p(a) of any life table should be linearly related to the survivorship function of any other life table, or of a standard, p*(a), by a level parameter α and a shape parameter β
- The transformation used is the logit

  $0.5*\ln\{[1-p(a)]/p(a)\}$

Continued

Brass Logit Life Table System

- Thus:

$$0.5 * \ln\left\{\frac{1-p(a)}{p(a)}\right\} = \alpha + \beta * 0.5 * \ln\left\{\frac{1-p^{*}(a)}{p^{*}(a)}\right\}$$

- Where α effectively determines the level of mortality in the observed life table relative to the standard
- β determines the relative importance of child versus adult mortality in the observed relative to the standard tables

Continued

Brass Logit Life Table System

- The standard has $\alpha=0$ and $\beta=1$
- An overall mortality that is lower than the standard gives a negative α
- Heavier adult relative to child mortality gives $\beta>1.0$
- And vice versa in both cases

Continued

Brass Logit Life Table System

- Brass provides two standards:
  - A general standard
  - “African” standard with heavier child mortality (and particularly heavier child mortality relative to infant mortality)

Continued

Brass Logit Life Table System

- Any reliable life table can be used as a standard
- The model has two parameters (given a particular standard) and can be fitted to any two points of an observed life table

Continued

Brass Logit Life Table System

- The logit system is very easy to use, and very flexible given the range of standards that can be adopted
- However, in most developing countries, one has no basis on which to select a standard, and the choice of an inappropriate one can give poor results
- Further, old age mortality is sensitive to the value of β, and values very different from one can produce implausible levels of old age mortality

United Nations Model Life Tables for Developing Countries

- Developed in the early 1980s
- Five regional groupings of mortality patterns:
  - Latin America
  - Chile
  - South Asia
  - Far East
  - General

Continued

United Nations Model Life Tables for Developing Countries

- Tabulated values of the models for each regional grouping are available for males and females separately for e(0) ranging from 35 to 75 years

Continued

United Nations Model Life Tables for Developing Countries

- The models were developed using principal components analysis, and the mortality level e(0) is merely the first component
- Thus the models can be used as two- or three-parameter models (the second parameter being the relation of mortality under age five to over age five, and the third being a childbearing parameter for females and diverse factors for males)

Continued

United Nations Model Life Tables for Developing Countries

- The main advantages of the U.N. models are that they are based on recent experience in developing countries and that they are flexible with the second and third parameters

Continued

United Nations Model Life Tables for Developing Countries

- The main disadvantages are that the regional patterns may reflect residual data problems such as age misreporting, that they provide a limited range of e(0) (though adequate for most current applications) and that the second and third parameters may not be very useful since their estimation requires heavy reliance on uncertain data

Continued

United Nations Model Life Tables for Developing Countries

- It should also be noted that the regional groupings are not very precise

Summary

- A stationary population has the same number of births and deaths, so its size does not change; the interpretation of the columns is slightly different than in the conventional life table
- Model life tables have been developed, the most renown of which are those of the Coale-Demeny, Brass, and the United Nations