Survival Models: Week 1

Introduction

- Course information and announcements on wattle
- 3 x 1 hour Lectures per week (Wed 10am-11am; Thu 1:00pm-2:00pm; Fri 9:00am-10:00am)
- 1 x 1 hour Tutorial per week (start in week 2). Tutorial Signup on the wattle page
- No prescribed textbook for the course - all course material provided through lecture notes, tutorial questions
- Consultation: 1pm-3pm Tuesday, Room 3.07 ANUCBE Bldg. 26C

Assessment

- Mid Semester Exam
  - Worth 20% (redeemable)
  - Mid Semester Exam will be 1.5 hours long and will cover the week sessions 1 to 6.
- Final Exam
  - Worth 70%
  - The final examination will be three hours long and will cover the entire syllabus.
- Both exams are **Closed book**. One A4-size paper with notes on both sizes and paper-based dictionary (no approval needed but all notes must be removed) are allowed for both examinations.

The Life Table

- A form of survival model - a model of random future lifetimes.
- Life table starts with a fixed number of persons at age 0, l_0.
- The life table reports the number of persons expected to survive to each age x, l_x.
- The next slide has an extract from the Australian Life Tables 2005-2007. (`http://www.aga.gov.au/publications/life_tables_2005-07/downloads/Australian_Life_Tables_2005-07.pdf`)

The Life Table

Table 1: Extract from Australian Life Tables 2005-07

number of persons alive at age x

arbitrary, but traditionally we use 100,000

Age	l_x	d_x	p_x	q_x
0	100,000	523	0.99477	0.00523
1	99,477	40	0.99960	0.00040
2	99,437	28	0.99972	0.00028
3	99,409	18	0.99982	0.00018
4	99,391	14	0.99986	0.00014

increases

decreases

Some Notation

d_0 : # dies between age 0 & 1.

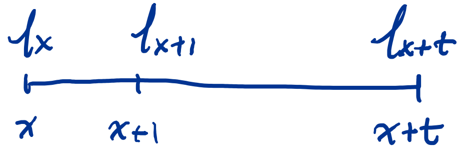
- l_x - number of persons alive at age x exact.
- d_x - number of persons dying between ages x and $x + 1$.
- p_x - probability a person aged x survives to age $x + 1$.
- q_x - probability a person aged x dies before age $x + 1$.
- $\boxed{{}_tp_x}$ - probability a person aged x survives to age $x + t$.
- $\boxed{{}_tq_x}$ - probability a person aged x dies before age $x + t$.

$p_x + q_x = 1$

${}_tp_x + {}_tq_x = 1$

Question: Express ${}_tp_x$ and d_x in terms of l_x. Express l_x in terms of d_x.

① $p_x = \frac{l_{x+1}}{l_x}$, $_tp_x = \frac{l_{x+t}}{l_x}$



② $d_x = l_x - l_{x+1}$

Note:
no such term as

$\cancel{_t d_x}$

③ l_x in terms of d_x

$$d_x = l_x - l_{x+1}$$
$$d_{x+1} = l_{x+1} - l_{x+2}$$
$$\vdots$$
$$d_{x+T} = l_{x+T} - l_{x+T+1} = 0$$

Finally it becomes zero b/c everyone dies eventually.

ADD UP TO GET

$$l_x = d_x + d_{x+1} + \cdots + d_{x+T} = \sum_{t=0}^{T} d_{x+T}$$ sum of deaths in all future years.

④ $q_x = 1 - p_x = \frac{l_x - l_{x+1}}{l_x} = \frac{d_x}{l_x}$

$_tq_x = 1 - {}_tp_x = \frac{l_x - l_{x+t}}{l_x}$

WE DO NOT have any closed form for this

Exercises

Compute the following using ALT 2005-07.

1. P(a male aged 1 survives to age 4). $= \frac{l_4}{l_1} = {}_3p_1 \neq {}_4p_1$ ①
2. E(number of new born males that will survive to age 2).
3. P(a male aged 1 dies aged 2). ②

$Y_i = \begin{cases} 1 & {}_2p_0 \\ 0 & 1-{}_2p_0 \quad ({}_2q_0) \end{cases}$

individual i

# of newborn survive to 2 : $\sum_{i=1}^{l_0} Y_i$

$$E\left(\sum_{i=1}^{l_0} Y_i\right) = \sum_{i=1}^{l_0} E(Y_i)$$
$$= \sum_{i=1}^{l_0} ({}_2p_0 \cdot 1 + {}_2q_0 \cdot 0)$$
$$= l_0 \times {}_2p_0$$
$$= l_0 \times \frac{l_2}{l_0}$$
$$= l_2$$

# of newborns

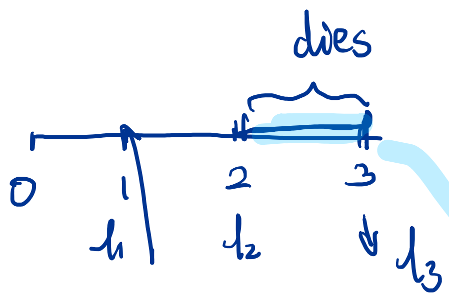
prob. new born survives to 2

Not surprising

Assumption: everyone has the same surviving prob.

③ "dies aged 2" = survived at age 2 but DIES before age 3.

$P(\text{aged 1 dies aged 2}) \neq {}_{2}q_{1} \neq q_{1}$ $= \dfrac{d_2}{l_1}$



"deferred prob of die"

cannot express it with q directly,

need l_2 first, & l_3

$l_2 - l_3$: # dies in this period. $= d_2$

divide by l_1 → aged 1 (base)

The Force of Mortality

similar to "force of interest"

- The force of mortality (μ_x) is the ***instantaneous*** rate of mortality at age x.
- μ_x is expressed as an annualised rate (l_x continuous and differentiable):

derivative of l_x in terms of x

$$\mu_x = -\frac{1}{l_x}\frac{d}{dx}l_x = -\frac{d}{dx}log(l_x)$$

as ratio

"decrease of l_x" to the ratio of l_x

~~log x~~ $(\log x)' = \frac{1}{x}$

- The "negative" sign ensures that μ_x is non-negative.
- From the definition of the derivative we have: $\mu_x = \frac{{}_hq_x}{h}$.

$\approx \frac{{}_hq_x}{h}$

- How to derive the above result?

Show that $\mu_x = \dfrac{{}_hq_x}{h}$

$$\mu_x = -\frac{1}{l_x}\frac{d\,l_x}{dx}$$

$$= -\frac{1}{l_x}\lim_{h\to 0}\frac{l_{x+h}-l_x}{h}$$

$$\approx -\frac{1}{l_x}\cdot\frac{l_{x+h}-l_x}{h} \quad (\text{for small } h)$$

$$= \frac{1}{h}\cdot\frac{l_x-l_{x+h}}{l_x}$$

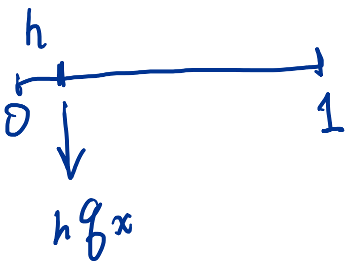
$$= \frac{1}{h}(1-{}_hp_x)$$

$$= \frac{1}{h}\cdot{}_hq_x$$

$$= \frac{{}_hq_x}{h}$$

why? As $h \to$ small, the mortality ~~rate~~ is not

annulised mortality at the precise moment of attaining age x.



The Force of Mortality

- From the approximation on the previous slide, it is clear that μ_x provides a measure of the instantaneous rate of mortality at age x.

- Knowledge of μ_x allows computation of survival times:

$$\int_0^n \mu_{x+s} ds \implies {}_n p_x = \exp(-\int_0^n \mu_{x+s} ds).$$

- How to derive the above result?

$${}_n p_x = \exp\left(-\int_0^n \mu_{x+s}\, ds\right)$$

$$\mu_x \rightarrow {}_n p_x$$

$$\mu_x = -\frac{d \ln l_x}{dx}$$

x is the starting age (fixed)

$\boxed{s}$ ~~is~~ changes

$$\underbrace{x \quad x+s}_{s}$$

$$\mu_{x+s} = -\frac{d \ln l_{x+s}}{ds}$$

TAKE INTEGRAL FOR both sides

$$\Rightarrow \int_0^n \mu_{x+s}\, ds = \int_0^n -\frac{d \ln l_{x+s}}{\cancel{ds}} \cancel{ds}$$

$$= \left[-\ln l_{x+s}\right]_0^n$$

$$= -\ln l_{x+n} + \ln l_x = -\ln \frac{l_{x+n}}{l_x} = -\ln({}_n p_x)$$

The Force of Mortality

- How about ${}_nq_x$?

  ${}_nq_x = \int_0^n {}_sp_x\mu_{x+s}ds$

- How to derive the above result?

  Note: Letting $n = 1$ in the above result gives us $d_x = \int_0^1 l_{x+t}\mu_{x+t}dt$.

Similarly, $\mu_x \longrightarrow {}_nq_x$

$$\mu_{x+s} = -\frac{d\,l_{x+s}}{ds}\frac{1}{l_{x+s}}$$

$$\frac{d\,l_{x+s}}{ds} = -l_{x+s}\mu_{x+s}$$

$\Downarrow$ take int. on both

$$\underbrace{\int_0^n \frac{d\,l_{x+s}}{ds}ds}_{\left[l_{x+s}\right]_0^n} = -\int_0^n l_{x+s}\cdot\mu_{x+s}\,ds$$

$$= l_{x+n} - l_x = -\int_0^n l_{x+s}\cdot\mu_{x+s}\,ds$$

divide l_x for both sides

$$-\frac{l_{x+n}-l_x}{l_x} = \int_0^n \frac{l_{x+s}}{l_x}\mu_{x+s}\,ds$$

$${}_nq_x = \int_0^n {}_sp_x\cdot\mu_{x+s}\,ds$$

Let $n=1$,

$$l_{x+1} - l_x = -\int_0^1 l_{x+s}\mu_{x+s}\,ds$$

$$l_x - l_{x+1} = d_x = \int_0^1 l_{x+s}\mu_{x+s}\,ds$$

Statistical Notation

- Revision of random variable theory from STAT2001.
- If a random variable X is continuous, the probability density function $f(x)$, is defined so that

$$P(a < X < b) = \int_a^b f_X(t)dt$$

- Other useful results from STAT2001 include:

CDF Expected value

$$\underbrace{P(X < x) = F_T(x)}_{} = \int_{-\infty}^{x} f_X(t)dt, \quad E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$$

- Conditional probability: consider two events A and B, we have

$$P(A|B) = \frac{P(AB)}{P(B)}$$

The Force of Mortality

Let T denote survival time. The distribution of T can be characterised by three equivalent quantities:

1. The survival function $S(t) = P(\text{a life survives longer than } t) = P(T > t)$. In practice, if there are no censored observations we could estimate $S(t)$ via:

$$\hat{S}(t) = \frac{\text{number of persons surviving longer than } t}{\text{total number of persons}}.$$

(estimated)

2. The probability density function (pdf) $f(t)$:

see next page

$$f(t) = \frac{\lim \Delta t \to 0 P(\text{life dies in } (t, t + \Delta t))}{\Delta t}.$$

$= F'(t)$ derivation ①

3. The hazard function $\mu(t)$ [Note: $\mu_t \equiv \mu(t)$]

$$\mu(t) = \frac{\lim \Delta t \to 0 P(\text{life dies in } (t, t + \Delta t), \text{given alive at } t)}{\Delta t}.$$

derivation ②

Actuarial Notation   Stat Notation

T_x = future lifetime for a person aged x.

↓

sometime omitted as T b/c starting age x is unimportant.

$_tq_x$ (prob of dying b/w x & $x+t$)

$$= \frac{l_x - l_{x+t}}{l_x} = P(T_x \le t) = F_{T_x}(t)$$

$$_tp_x = \frac{l_{x+t}}{l_x} = P(T_x \geq t) = S_{T_x}(t)$$

$$\mu_{x+t} = \boxed{-\frac{1}{l_{x+t}} \frac{d\, l_{x+t}}{dt}} = \mu(t) = \frac{f(t)}{S(t)}$$

(sometimes as $\lambda(t)$)

$$\boxed{_tq_x = 1 - \frac{l_{x+t}}{l_x}}$$

$$= -\frac{\frac{d\, l_{x+t}}{dt} / l_x}{l_{x+t} / l_x} = \frac{\frac{d_t\, q_x}{dt}}{_tp_x} = \frac{\frac{d\, F_{T_x}(t)}{dt}}{S_{T_x}(t)} = \frac{f(t)}{S(t)}$$

Note the difference between $f(t)$ and $\mu(t)$. The hazard function is a *conditional* rate.

Derivation ①

$$f(t) = \frac{dF(t)}{dt} = \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{\boxed{P(\text{life dies in } (t, t+\Delta t))}}{\Delta t}$$

① $F(t+\Delta t) = ?$
prob an individual's future life time $< t+\Delta t$

② $F(t) = ?$
prob ... $< t$

Derivation ②

$$\mu(t) = \frac{f(t)}{S(t)} = \lim_{\Delta t \to 0} \left(\frac{P(\text{life dies in } (t, \Delta t + t))}{\Delta t \cdot S(t)} \right)$$

$$= \lim_{\Delta t \to 0} \left(\frac{\dfrac{P(\text{life dies in } (t, t+\Delta t))}{P(\text{life surivives longer than } t)}}{\Delta t} \right)$$

def: A: life dies in $(t, t+\Delta t)$

B: life surives longer than t. = life dies in $(t, +\infty)$

$$\frac{P(AB)}{P(B)} = P(A|B)$$

$$\mu(t) = \lim_{\Delta t \to 0} \frac{P(\text{life dies in } (t, t+\Delta t) \mid \text{life surives longer than } t)}{\Delta t}$$

P_{T_x}(life dies in $(t, t+\Delta t)$)

→ starting time is x, ~~always~~ sometimes omitted.

$$\underbrace{l_x \quad\quad l_{x+t}}_{t} \quad \underbrace{\quad\quad l_{x+t+\Delta t}}_{\Delta t}$$

$$= \frac{l_{x+t} - l_{x+t+\Delta t}}{l_x}$$

P(life dies in $(t, t+\Delta t)$ | alive at t)

$$= \frac{l_{x+t} - l_{x+t+\Delta t}}{l_{x+t}}$$

→ given alive at t.