Survival Models: Week 2

The Force of Mortality

- The force of mortality (also known as the *hazard function*), can be expressed in terms of the survival function $S(t)$ and the probability density function $f(t)$:

$-S'(t) = -(1-F(t))' = F'(t) = f(t)$

$$\mu(t) = \mu_T(t) = \lambda_T(t) = \frac{f(t)}{S(t)} = \frac{-S'(t)}{S(t)}.$$

- $S(t)$ is the probability alive at time t.
- In week 1 notes we saw that $f(t)$ is an unconditional failure rate.
- $\mu(t) \times \delta \approx P(\text{die in } (t, t+\delta) | \text{alive at } t)$.

Force of mortality
of exponential distribution

Exercise

$\mu(t) = \frac{-S'(t)}{S(t)}$

$\Rightarrow \frac{d}{dt}\log(S(t)) = -\mu(t)$

$[\log(S(r))]_0^t = -\int_0^t \mu(r)\,dr$

$\log \frac{S(t)}{1} = -\int_0^t \mu(r)\,dr$

$S(t) = e^{-\int_0^t \mu(r)dr}$

$F(0) = 0$
$S(0) = 1$

- How to express $S(t)$ in terms of $\mu(t)$?

- Suppose the random variable T, representing the future lifetime of a population, follows an exponential distribution, that is, $f(t) = \lambda \exp(-\lambda t)$. What is the force of mortality for this population?

  exponential

  $F(t) = 1 - \exp(-\lambda t)$
  $S(t) = \exp(-\lambda t)$

  Would you have any concerns using such a force of mortality for human populations?

  Not real.

  $\mu(t) = \frac{f(t)}{S(t)} = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda$

Exercise

Suppose the random variable T, representing the future lifetime of a population, is such that:

$$f(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1} \exp(-(\frac{t}{\beta})^\alpha) \quad t > 0, \alpha > 0, \beta > 0.$$

What is the force of mortality for this population?

What is the advantage of the above force of mortality compared to that from the exponential distribution?

$F(t) = 1 - \exp(-(\frac{t}{\beta})^\alpha)$

$F'(t) = -(-\frac{t^{\alpha-1}}{\beta^\alpha} \cdot \alpha) \cdot \exp(-(\frac{t}{\beta})^\alpha) = f(t)$

$S(t) = \exp(-(\frac{t}{\beta})^\alpha)$

$\mu(t) = \frac{f(t)}{S(t)} = \frac{\frac{\alpha}{\beta^\alpha} t^{\alpha-1} \cancel{\exp(-(\frac{t}{\beta})^\alpha)}}{\cancel{\exp(-(\frac{t}{\beta})^\alpha)}} = \frac{\alpha}{\beta^\alpha} t^{\alpha-1}$

constant

increasing

kind of better than the last one. (not constant)

but not perfect. Human mortality is complicated.

Some "Laws" of Mortality

It is sometimes convenient to assume that the force of mortality adheres to a simple parametric form (or "law"). Two commonly used parametric forms are:

- Gompertz Law: $\mu(t) = B \times C^t$.
- Makehams Law: $\mu(t) = A + B \times C^t$
- Gompertz law implies that $\frac{\mu(t+1)}{\mu(t)} = C$ - constant percentage increase.
- Makehams law adds the "fixed" component A - causes of death due to chance.

previously, it's t^α, but now we use C^t.
$\Rightarrow$ faster mortality rate.

Gompertz Law

Gompertz law states that:

$\mu(t) \propto t$: exponential rate

$$log(\mu(t)) = log(B) + tlog(C).$$

- For human populations, Gompertz law (and Makehams law) will only be suitable over relatively small age ranges. b/c infants are special with a high mortality rate.
- The above relationship says that a plot of $log(\mu(t))$ versus age (t) should look roughly linear, if Gompertz law is appropriate.
- The above point suggests a way to estimate the parameters B and C.

linear Regression.

R Example

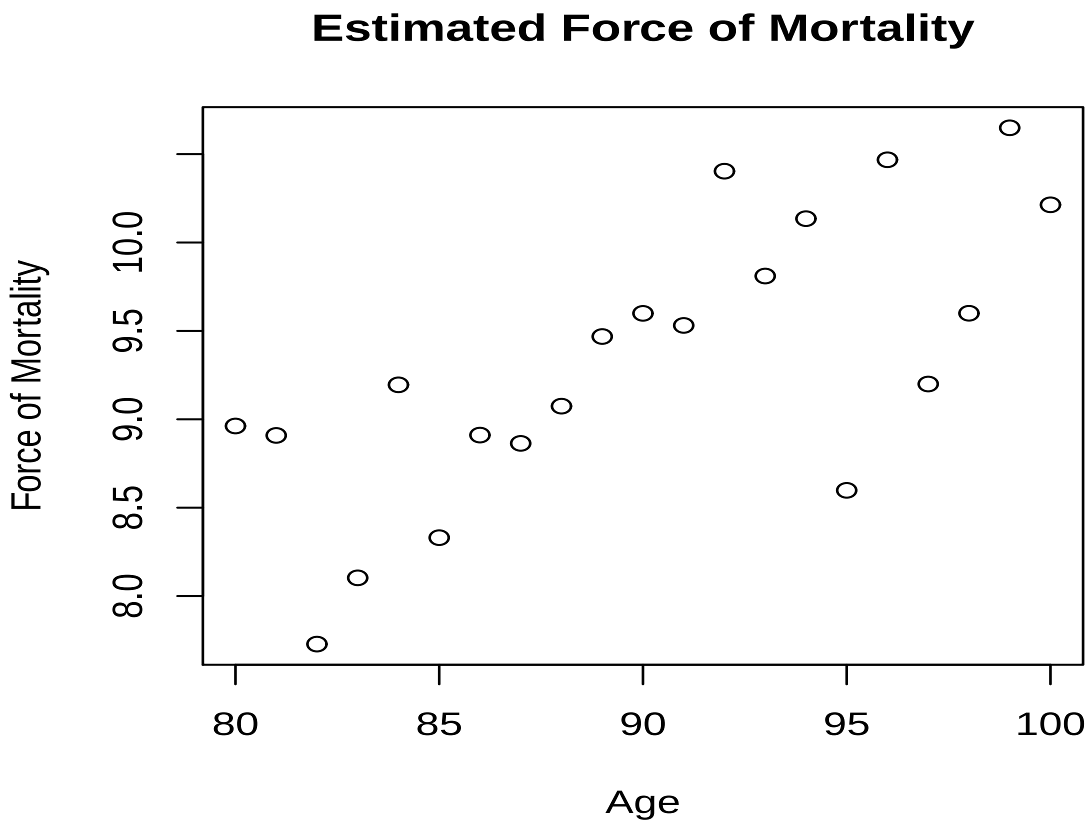


Figure 1: Plot of $log(\mu(t))$ versus age

R Example

We will fit a simple linear regression (SLR) model to estimate B and C.

```
> #fitting a SLR model to estimate B and C
> fit<-lm(logmu~age)
> fit$coef
(Intercept)         age
 1.10991281  0.09123931
> exp(fit$coef)
(Intercept)         age
   3.034094    1.095531
```

So, we have $\hat{B} = 3.03$ and $\hat{C} = 1.1$.

Complete Expected Future Lifetime

This is the "standard" expected future lifetime that most people envisage. The complete expected future lifetime is defined as:

$$e_x^o = E[T_x] = \int_0^\infty t f(t) dt,$$

where, T_x is the random variable representing the future lifetime of person aged x. Note, using the result ${}_nq_x = \int_0^n {}_sp_x \mu(x+s)ds$, we can express $f(t)$ as ${}_tp_x \mu(x+t)$. Then, we can derive $e_x^o = \int_0^\infty {}_tp_x dt$. How?

Why?

$E(X) = \int_0^\infty P(X \geq x)\,dx$ for non-negative X.

$E(T_x) = \int_0^\infty P(T_x \geq t)\,dt = \int_0^\infty t f(t)\,dt = \mathring{e}_x$

~~$_nq_x = \int_0^n {_sp_x}\,\mu(x+s)\,ds$~~

$f(t) = {_tp_x}\,\mu(x+t)$

$\mathring{e}_x = \int_0^\infty {_tp_x}\,dt$? => To show this:

$_nq_x = \int_0^n {_tp_x}\,\mu(x+t)\,dt$: $P(T_x \leq n)$ die in n years

$= F_{T_x}(n)$

$= \int_0^n f(t)\,dt$

or $f(t) = -S'(t) = \frac{-d\,{_tp_x}}{dt}$

${_tp_x}\,\underline{\mu(x+t)} = f(t)$

$\mu(x+t) = \frac{1}{l_{x+t}} \cdot \frac{d\,l_{x+t}}{dt}$

$\frac{\text{rate of decline}}{\text{\# survived}}$

$= \frac{-\frac{d\,l_{x+t}}{l_x}/dt}{\frac{l_{x+t}}{l_x}}$

$= \frac{-d\,{_tp_x}/dt}{{_tp_x}}$

SO : $\mathring{e}_x = \int_0^\infty t f(t)\,dt = \int_0^\infty t \cdot \cancel{{_tp_x}}\,\frac{-d\,{_tp_x}/dt}{\cancel{{_tp_x}}}\,dt = \int_0^\infty t\left(\frac{-d\,{_tp_x}}{dt}\right)dt$

let $u = t$

$dv = \frac{-d\,{_tp_x}}{dt}\,dt$

$v = -{_tp_x}$

$du = 1$

$\int u\,dv = uv - \int v\,du$

$= \underbrace{\left[t \cdot -{_tp_x}\right]_0^\infty}_{=?\ \text{zero}\ ?} + \int_0^\infty {_tp_x}\,dt$

How to derive this easily?

e.g. Exp Dist.

${_tp_x} = e^{-\lambda t}$

$\frac{t}{{_tp_x}} \to 0$ as $t \to \infty$

or "people dies", ${_tp_x} \to 0$

$= \int_0^\infty {_tp_x}\,dt$

Curtate Expected Future Lifetime

This is a "non-standard" expected future lifetime. The curtate expected future lifetime is defined as:

$$e_x = E[K_x] = \sum_{k=0}^{\infty} kP(K_x = k),$$

K_x is discrete

where, K_x is the random variable representing the *whole number* of future years lived by a person aged x. For example, if a person aged 20 lives to age 25.8, $T_{20} = 5.8$ and $K_{20} = 5$. We can also show that (how?):

$K_{20} = \lfloor T_{20} \rfloor$

survives ~~to~~ to $x+k$ years

but dies before $x+k+1$

$$P(K_x = k) = {}_k p_x \times q_{x+k} \quad \text{and} \quad e_x = \sum_{k=1}^{\infty} {}_k p_x .$$

$= P(k \leq T_x \leq k+1)$

$= {}_k P_x \cdot q_{x+k}$

$$e_x = \sum_{k=0}^{\infty} k \cdot P(K_x = k)$$

$$= \sum_{k=0}^{\infty} k \cdot {}_kP_x \cdot q_{x+k}$$

$$= 1 \cdot P_x \cdot q_{x+1} + 2\, {}_2P_x\, q_{x+2} + 3\, {}_3P_x\, q_{x+3} + \cdots$$

$$= 1 \cdot P_x\, q_{x+1} + {}_2P_x\, q_{x+2} + {}_3P_x\, q_{x+3} + \cdots$$

$$+ \qquad {}_2P_x\, q_{x+2} + {}_3P_x\, q_{x+3} + \cdots$$

$$+ \qquad {}_3P_x\, q_{x+3} + \cdots$$

$$= \sum_{j=1}^{\infty} {}_jP_x\, q_{x+j} + \sum_{j=2}^{\infty} {}_jP_x\, q_{x+j} + \cdots$$

$$= \sum_{k=1}^{\infty} \boxed{\sum_{j=k}^{\infty} {}_jP_x\, q_{x+j}}$$

1 term $\quad {}_kP_x\, q_{x+k} = \dfrac{l_{x+k}}{l_x} \cdot \dfrac{l_{x+k} - l_{x+k+1}}{l_{x+k}}$

2 term $\quad \dfrac{l_{x+k+1}}{l_x} \cdot \dfrac{l_{x+k+1} - l_{x+k+2}}{l_{x+k+1}}$

∞ term $\quad \dfrac{l_{x+k} - \cancel{l_{x+k+1}} + \cancel{l_{x+k+1}} - \cancel{l_{x+k+2}} + \cdots}{l_x}$

So the whole first Σ is just ${}_kP_x$

$$= \sum_{k=1}^{\infty} {}_kP_x$$

Another way to do so:

$$\sum_{j=k}^{\infty} {}_jp_x\, q_{x+j} = \sum_{j=k}^{\infty} P(j \le T_x < j+1)$$

$$= P(k \le T_x < k+1) + P(k+1 \le T_x < k+2) + \cdots$$

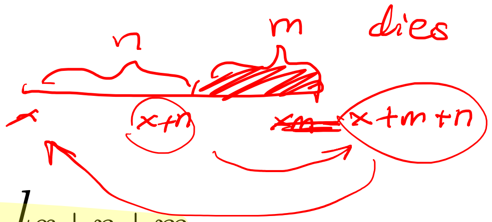
$$= P(T_x \ge k)$$

$$= {}_kp_x$$

So $e_x = \sum_{k=1}^{\infty} {}_kp_x$

Deferred Probabilities

We now introduce some more actuarial notation.



survived n years

now aged x

$$_{n|m}q_x = P(n < T_x < n+m) = \frac{l_{x+n} - l_{x+n+m}}{l_x}$$

dies within next m years

In words, $_{n|m}q_x$ is the probability that a life aged x will survive for n years but die during the subsequent m years.

When the probability involves dying within one year we omit the 1, interpret $_{n|}q_x$.

$_{n|}q_x = P(n < T_x < n+1)$

recall: ?

Non integer Ages

Life table functions are generally only recorded for integer ages. In order to get a value for say ${}_{0.5}q_{55}$ we must make an assumption about the trend in mortality rate between ages 55 and 56. We shall describe two methods.

- Method One: Uniform Distribution of Deaths (UDD)

  Assumption: ${}_tp_x\mu_{x+t}$ ($= f_t$) is a constant for $0 < t < 1$; $= f_t$: pdf of future lifetime

  Equivalent to a uniform distribution of the time to death (T_x), conditional on death falling between these two ages. only within that one year interval.

- Method Two: Constant Force of Mortality

  Assumption: $\mu_{x+t} = \mu$ =constant for $0 < t < 1$. Further discussion in Week 7.

Exercise

1. Show that UDD implies that $_sq_x = s \cdot q_x$, where $0 < s < 1$.

$$_sq_x = \int_0^s {}_tp_x \mu_{x+t}\, dt = s \cdot {}_tp_x \mu_{x+t}$$

UDD

$$q_x = \int_0^1 {}_tp_x \mu_{x+t}\, dt = {}_tp_x \mu_{x+t}$$

${}_tp_x \mu_{x+t}$ constant

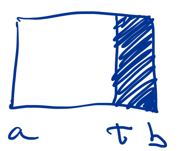
$\Rightarrow {}_sq_x = s \cdot q_x$

2. Assuming the future lifetime of an individual is uniformly distributed over a given age (e.g. $T \sim U(0, 1)$), show that the hazard function for this individual is a monotonically increasing function.

$f(t) = {}_tp_x \mu_{x+t}$ constant, when $t \sim U(0,1)$

$$\mu(t) = \mu_{x+t} = \frac{f(t)}{S(t)} = \frac{\frac{1}{b-a}}{\frac{b-t}{b-a}} = \frac{1}{1-t}$$

$t \uparrow$, $1-t \downarrow$, whole $\nearrow$



$b-t$: survived more $\geq t$

$b-a$: original

Central Rate of Mortality

aka central exposed to risk.

The central rate of mortality (m_x) is an alternative to the initial rate of mortality (q_x). The central rate of mortality is often used instead of the initial rate in demography because collected data is more amenable to computing m_x.

When asked,
people ~~most~~ more
likely to say:
I'm x years old,
not $x.5$ years old.

$$m_x = \frac{d_x}{\int_0^1 l_{x+t}dt}.$$

The term $\int_0^1 l_{x+t}dt$ is more appropriate when we have data on the number of people aged between x and $x+1$. We will come back to m_x when we talk about how life tables are constructed (Week 12).