LECTURE 10

The Expectation of Life

Recall:

X is the time-until-death for newborn.

$T(x)$ is the time-until-death for (x) (or, equivalently, the future lifetime at age x)

If the probability distribution of X is known (or, equivalently, the survival function) then the probability distribution of $T(x)$ can be obtained by the conditioning of the distribution of $X-x$ by the event $X>x$:

$$f_{T(x)}(t) = \frac{f_X(x+t)}{s(x)} = \frac{-\frac{d}{dt}s(x+t)}{s(x)} \qquad \text{[see Lecture 7, formula (k)]}$$

Alternatively,

$$f_{T(x)}(t) = {}_tp_x\mu(x+t).$$

This formula contains no reference to conditioning and can be used when the survival probabilities and the force of mortality are known from age x onwards.

Aside:-

${}_tp_x$	is the probability that (x) survives further t years,
$\mu(x+t)dt$	is the probability that $(x+t)$ dies within a fraction of time dt.

Hence,

${}_tp_x\mu(x+t)dt$	can be interpreted in words as “the probability that (x) survives further t years, i.e. to age $(x+t)$, and then dies instantly”.

The Complete Expectation of Life

The expected value of the time-until-death for (x) is called the *complete-expectation-of-life at age x* and is denoted by $\mathring{e}_x$:

$$\mathring{e}_x \overset{\text{def.}}{=} E(T(x)) = \int_0^\infty t\, f_{T(x)}(t)\,dt = \int_0^\infty t\,{}_tp_x\mu(x+t)\,dt$$

There is a simpler expression for $\mathring{e}_x$:

$$\mathring{e}_x = \int_0^\infty {}_tp_x\,dt. \tag{1}$$

To derive it notice that

$$\frac{d}{dt}({}_tp_x) = \frac{d}{dt}\frac{s(x+t)}{s(x)} = \frac{s(x+t)}{s(x)}\frac{\frac{d}{dt}s(x+t)}{s(x+t)} = -{}_tp_x\,\mu(x+t)$$

Therefore, applying integration by parts

$$\mathring{e}_x = \int_0^\infty t\,{}_tp_x\mu(x+t)\,dt = -\int_0^\infty t\,d({}_tp_x) = -t\,{}_tp_x\Big|_0^\infty + \int_0^\infty {}_tp_x\,dt \tag{2}$$

Any realistic survival function will satisfy $\lim_{t\to\infty} ts(x+t)=0$, or, equivalently, $\lim_{t\to\infty} t\,{}_tp_x = 0$. (For human populations $s(x)$ should vanish at all for sufficiently large ages, $s(x)=0$ for all $x>150$, say). Therefore $t\,{}_tp_x\Big|_0^\infty = 0$ and we obtain Eq. (1) from Eq. (2).

Example. Exponentially distributed lifetime, $X \sim \text{Exp}(\lambda)$

$s(x) = e^{-\lambda x}$, $\mu(x) = -\frac{d}{dx}\ln s(x) = \lambda$, and ${}_tp_x = \frac{s(x+t)}{s(x)} = e^{-\lambda t}$.

$$\mathring{e}_x = \int_0^\infty {}_tp_x\,dt = \int_0^\infty e^{-\lambda t}\,dt = \frac{1}{\lambda}$$

is independent of age x (and coincides with $E(X)$). Hence using an exponential distribution over the entire lifespan is doubtful when a human population is concerned (however, exponential distributions can be applied over small age intervals).

The Curtate Expectation of Life

$T(x)$ is a continuous-type random variable. It expresses the *exact* time-until-death for (x). One can also associate with the future lifetime of (x) a discrete-type random variable. This discrete-type random variable is the number of years *completed* by (x) prior to death and is denoted by $K(x)$. Its probability mass function can be easily computed:

$$\begin{aligned} P(\,K(x)=k\,) &= P(k \le T(x) < k+1) \\ &= P(k < T(x) \le k+1) \qquad \text{[by (j) from Lecture 7]} \\ &= P(T(x)>k) - P(T(x)>k+1) \\ &= {}_kp_x - {}_{k+1}p_x. \end{aligned}$$

The expected value of $K(x)$ is called the *curtate-future-lifetime* of (x) and is denoted by e_x,

$$\begin{aligned} e_x = E(\,K(x)\,) &= \sum_{k=0}^\infty k\,P(K(x)=k) \\ &= \sum_{k=0}^\infty k\,{}_kp_x - \sum_{k=0}^\infty k\,{}_{k+1}p_x \\ &= \sum_{k=1}^\infty k\,{}_kp_x - \sum_{k=1}^\infty (k-1)\,{}_kp_x \end{aligned}$$

Therefore (cf. with Eq. (1))

$$e_x = \sum_{k=1}^\infty {}_kp_x. \tag{3}$$

By its definition, $K(x)$ satisfies the inequalities $K(x) \le T(x) \le K(x)+1$, hence taking expectations $E(K(x)) \le E(T(x)) \le E(K(x))+1$ and therefore for all ages x

$$e_x \le \mathring{e}_x \le e_x + 1.$$

Though $\mathring{e}_x$ and e_x are related, there is no explicit relationship expressing one through the other.

However,

$$\mathring{e}_x \simeq e_x + 1/2. \tag{4}$$

This approximate relation can be obtained by using linear interpolation. Indeed,

$$\begin{aligned}\mathring{e}_x = \int_0^\infty {}_tp_x\,dt &= \sum_{k=0}^\infty \int_k^{k+1} {}_tp_x\,dt \\ &= \sum_{k=0}^\infty \int_0^1 {}_{k+u}p_x\,du \\ &\simeq \sum_{k=0}^\infty \left[{}_kp_x \int_0^1 u\,du + {}_{k+1}p_x \int_0^1 (1-u)\,du \right]\end{aligned}$$

assuming linear interpolation

$${}_{k+u}p_x \simeq {}_kp_x u + {}_{k+1}p_x(1-u) \qquad \text{for all } 0 \le u \le 1 \text{ and integer } k.$$

Therefore

$$\mathring{e}_x \simeq \frac{1}{2}\sum_{k=0}^\infty {}_kp_x + \frac{1}{2}\sum_{k=0}^\infty {}_{k+1}p_x = \frac{1}{2}\left(1+\sum_{k=1}^\infty {}_kp_x\right) + \frac{1}{2}\sum_{k=1}^\infty {}_kp_x = \frac{1}{2} + \sum_{k=1}^\infty {}_kp_x = e_x + \frac{1}{2}.$$

SUMMARY.

Complete Expectation of Life: $\mathring{e}_x = E(T(x))$,
where $T(x)$ is the time-until-death for a person age x:

$$\mathring{e}_x = \int_0^\infty {}_tp_x\,dt = \frac{1}{s(x)}\int_0^\infty s(x+t)\,dt.$$

Curtate Expectation of Life: $e_x = E(K(x))$, where $K(x)$ is the number of years completed by (x) prior to death: $e_x = \sum_{k=1}^\infty {}_kp_x$.

Approximate relationship: $\mathring{e}_x \simeq e_x + 1/2$

LECTURE 11

Life Tables

Random Survivorship Group

Consider a group of l_0 newborns whom we label by j, $j = 1,\dots,l_0$. Let X_j be the time-until-death for newborn j.

How many of the group will survive to age x? Denote this number by $\mathcal{N}(x)$:

$$\mathcal{N}(x) = \sum_{j=1}^{l_0} \mathbf{1}_j,$$

where $\mathbf{1}_j$ is the indicator for the survival of newborn j,

$$\mathbf{1}_j = \begin{cases} 1 & \text{if newborn } j \text{ survives to age } x \\ 0 & \text{otherwise} \end{cases}$$

Assume that X_j for all j has a common distribution specified by survival function $s(x)$. Then $\mathbf{1}_j \sim$ Bernoulli(p) where $p = P(X > x) = s(x)$ for all j and $E(\mathbf{1}_j) = s(x)$.

Therefore the expected number of survivors to age x from the group is

$$E(\mathcal{N}(x)) = E\left(\sum_{j=1}^{l_0} \mathbf{1}_j\right) = \sum_{j=1}^{l_0} E(\mathbf{1}_j) = l_0 s(x).$$

Under the assumption that $\mathbf{1}_j$, $j = 1,\dots,l_0$, are mutually independent, $\mathcal{L}(x) \sim \text{Bin}(l_0, s(x))$.

Notation:-

the expected number of survivors to age x in a group of l_0 newborns is denoted by the symbol l_x,

$$l_x = l_0 s(x); \qquad l_0 \text{ is called the } \textit{radix};$$

the expected number of deaths between ages x and $x+t$ is denoted by ${}_td_x$,

$${}_td_x = l_x - l_{x+t};$$

when $t = 1$ the prefix t is omitted, $d_x \equiv {}_1d_x$.

Life Tables

Tables containing estimates (or "observed" values) of the values of life-table functions for *integer* ages are called the life tables.

The first life table was published in 1693 by Edmund Halley, who based his table on the register of births and deaths of the city of Breslau (now Wrocław).

Edmund Halley, (b. 1656 – d. 1742), English astronomer and mathematician, was the first person to calculate the orbit of a comet which known now as Halley's Comet. (Also published the 1st meteorological chart (contained the distribution of prevailing winds over oceans) and the 1st magnetic charts of the Atlantic and Pacific areas.)

l_x, d_x, p_x, q_x, and $\mathring{e}_x$ are examples of the life-table functions. These functions are usually included in any life table.

Note the relationships (x is an integer):

$$\begin{aligned}
p_x &= \frac{l_{x+1}}{l_x} && \text{the probability that } (x) \text{ survives to age } x+1\\
q_x &= \frac{l_x - l_{x+1}}{l_x} && \text{the probability that } (x) \text{ dies before attaining age } x+1\\
d_x &= l_x - l_{x+1} && \text{the expected number dying aged } x \text{ last birthday from } l_0 \text{ newborns}\\
e_x &= \frac{1}{l_x}\sum_{k=1}^{\infty} l_{x+k}; && \text{the expected number of complete years lived after the } x\text{-th birthday}
\end{aligned} \tag{5}$$

$$\begin{aligned}
l_x &= l_0 s(x)\\
&\equiv l_0 P(X > x) && \text{the expected number of survivors to age } x \text{ from } l_0 \text{ newborns}\\
{}_t p_x &= \frac{l_{x+t}}{l_x} && \text{the probability that } (x) \text{ survives to age } x+t\\
{}_t q_x &= \frac{l_x - l_{x+t}}{l_x} && \text{the probability that } (x) \text{ dies before attaining age } x+t\\
{}_{t|u} q_x &= \frac{l_{x+t} - l_{x+t+u}}{l_x} && \text{the probability that } (x) \text{ dies aged between } x+t \text{ and } x+t+u \text{ years}\\
{}_t d_x &= l_x - l_{x+t} && \text{the expected number of those who will die aged between } x \text{ and } x+t-1 \text{ last birthday from } l_0 \text{ newborns}
\end{aligned} \tag{6}$$

The expressions relating the functions p, q, or e to l_x can be derived by expressing these functions in terms of the survival function $s(x)$ and using that $l_x = l_0 s(x)$, e.g.

$${}_t q_x = \frac{s(x) - s(x+t)}{s(x)} = \frac{l_0\, s(x) - l_0\, s(x+t)}{l_0\, s(x)} = \frac{l_x - l_{x+t}}{l_x}$$

or

$$e_x = \sum_{k=1}^{\infty} {}_k p_x = \sum_{k=1}^{\infty} \frac{s_{x+k}}{s_x} = \sum_{k=1}^{\infty} \frac{l_0\, s_{x+k}}{l_0\, s_x} = \sum_{k=1}^{\infty} \frac{l_{x+k}}{l_x}.$$

In this course we will be using two life tables: *English Life Table No. 12 – Males* and *A1967-70*.

Life tables are not constructed by observing l_0 newborns until the last survivor dies (one would need to wait too long!). Instead life tables are based on estimates of probabilities of death, given survival to various ages derived from the experience of the entire population under investigation.

Thus the English Life Table No. 12 – Males was constructed on the basis of the mortality experienced by the entire population of England in 1960, 1961 and 1962.

The A1967-70 table is based on the experience, within these years, of lives assured by UK life assurance companies.

Example. Toy example of constructing a life table

Suppose you know q_x for all integer x (or, equivalently, $p_x = 1 - q_x$. Then you can construct a life table using the relations in (5) and $l_{x+1} = p_x l_x$.
For instance, consider an animal population with a lifespan of 5 years (i.e. the animals live at most 5 years) with

$$\begin{aligned}
p_0 &= 0.5 && = P(T(0) > 1) = P(X > 1)\\
p_1 &= 0.4 && = P(T(1) > 1) = P(X > 2 | X > 1)\\
p_2 &= 0.3 && = P(T(2) > 1) = P(X > 3 | X > 2)\\
p_3 &= 0.2 && = P(T(3) > 1) = P(X > 4 | X > 3)\\
p_4 &= 0.1 && = P(T(4) > 1) = P(X > 5 | X > 4)\\
p_5 &= 0 && = P(T(5) > 1) = P(X > 6 | X > 5)
\end{aligned}$$

Set the radix: $l_0 = 10000$, say. Then (as $l_{x+1} = l_x \times p_x$)

$l_1 = l_0 p_0 = 10000 \times 0.5 = 5000$, $l_2 = l_1 p_1 = 5000 \times 0.4 = 2000$, etc.;

$d_0 = l_0 - l_1 = 10000 - 5000 = 5000$; $d_1 = l_1 - l_2 = 5000 - 2000 = 3000$, etc.;

$e_0 = \frac{l_1+l_2+l_3+l_4+l_5}{l_0} = \frac{5000+2000+600+120+12}{10000} = 0.7732$, hence $\mathring{e}_0 \simeq e_0 + \frac{1}{2} = 1.2732$

$e_1 = \frac{l_2+l_3+l_4+l_5}{l_1} = \frac{2000+600+120+12}{5000} = 0.5464$, hence $\mathring{e}_1 \simeq e_1 + \frac{1}{2} = 1.0464$, etc.

x	l_x	d_x	p_x	q_x	$\mathring{e}_x$	x
0	10000	5000	0.5	0.5	1.2732	0
1	5000	3000	0.4	0.6	1.0464	1
2	2000	1400	0.3	0.7	0.8660	2
3	600	480	0.2	0.8	0.7200	3
4	120	108	0.1	0.9	0.6000	4
5	12	12	0	1	0.5000	5

SUMMARY:
The relations in (5) and (6).