MATH3085/6143 Survival Models – Worksheet 5 Solutions

1. i) ${}_{10}q_0 = 1 - {}_{10}p_0 = 1 - \frac{l_{10}}{l_0} = 1 - \frac{99172}{100000} = 0.00828.$

   ii) The number of woman years lived at ages above x years (OR total expected future complete lifetime of women aged x) is T_x, which can be computed as $\mathring{e}_x \times l_x$ when the complete life expectancy at age x is available. Therefore, the number of woman years lived at ages above 30 years is

$$T_{30} = \mathring{e}_{30} \times l_{30} = 49.94 \times 98617 = 4924933,$$

   and the number of woman years lived at ages above 40 years is

$$T_{40} = \mathring{e}_{40} \times l_{40} = 40.24 \times 97952 = 3941588.$$

   iii) a) The number of woman years lived between exact ages 30 and 40 years is just the difference

$$T_{30} - T_{40} = 4924933 - 3941588 = 983345.$$

   Of these, those who survived to exact ages 40 years each lived 10 years between exact ages 30 and 40 years, contributing a total of

$$l_{40} \times 10 = 97952 \times 10 = 979520$$

   woman years. Hence, the number of woman years lived between exact ages 30 and 40 years by those who died between that interval is given by the difference between the two quantities above,

$$983345 - 979520 = 3825.$$

   b) Number of women who die between exact ages 30 and 40 years is

$$l_{30} - l_{40} = 98617 - 97952 = 665.$$

   These women contributed a total of 3825 woman years within this age range (from (iii)(a)). Thus on average, each of them contributed

$$\frac{3825}{665} = 5.75$$

   woman years (notice how close this is to that obtained from assumption of UDD, which is 10/2=5). The average age at death of those who died between 30 and 40 years is thus

$$30 + 5.75 = 35.75.$$

2. i) ${}_5d_{30}$ is the number of deaths between exact ages 30 and 35 years:

$${}_5d_{30} = l_{30} - l_{35} = 98617 - 98359 = 258.$$

   ${}_4q_1$ is the probability that a female alive on her first birthday will die before attaining her fifth birthday:

$${}_4q_1 = 1 - {}_4p_1 = 1 - \frac{l_5}{l_1} = 1 - \frac{99243}{99368} = 0.00126.$$

   ${}_5L_{30}$ is the number of person-years lived between exact ages 30 and 35 years OR the total expected life within the age interval $[30, 35)$ of the l_{30} females. Computation of this quantity requires assumption of uniform distribution of deaths (UDD) between exact ages 30 and 35 years:

$${}_5L_{30} = \frac{5}{2}(l_{30} + l_{35}) = \frac{5}{2}(98617 + 98359) = 492440.$$

ii) The probability that a woman aged 20 years will die before she attains the age of 40 years is just

$$ {}_{20}q_{20} = 1 - {}_{20}p_{20} = 1 - \frac{l_{40}}{l_{20}} = 1 - \frac{97952}{98957} = 0.01016. $$

iii) Since this is an irregular life table with uneven spacing of the age intervals, it easier to work with ${}_{x_i - x_{i-1}}L_{x_{i-1}}$ to find $\mathring{e}_0$ first, and then use the approximation that

$$ e_0 \approx \mathring{e}_0 - \frac{1}{2} $$

to find the curtate life expectancy at birth (note the approximation is exact is UDD is assumed). Hence, the following table is constructed.

i	x_{i-1}	$l_{x_{i-1}}$	${}_{x_i - x_{i-1}}L_{x_{i-1}} = \frac{x_i - x_{i-1}}{2}(l_{x_{i-1}} + l_{x_i})$
1	0	100000	$\frac{1}{2}(100000 + 99368) = 99684$
2	1	99368	$\frac{4}{2}(99368 + 99243) = 397222$
3	5	99243	$\frac{5}{2}(99243 + 99172) = 496037.5$
4	10	99172	$\frac{5}{2}(99172 + 99098) = 495675$
5	15	99098	$\frac{5}{2}(99098 + 98957) = 495137.5$
6	20	98957	$\frac{5}{2}(98957 + 98797) = 494385$
7	25	98797	$\frac{5}{2}(98797 + 98617) = 493535$
8	30	98617	$\frac{5}{2}(98617 + 98359) = 492440$
9	35	98359	$\frac{5}{2}(98359 + 97952) = 490777.5$
10	40	97952	

Also recall that we are given $\mathring{e}_{40} = 40.237$ years, which allows us to obtain the remaining person years to be lived by ages above 40 years:

$$ T_{40} = \mathring{e}_{40} l_{40} = 40.237 \times 97952 = 3941294.6. $$

Now

$$ \begin{aligned} T_0 &= \sum_{i=1}^{\infty} {}_{x_i - x_{i-1}}L_{x_{i-1}} \\ &= \sum_{i=1}^{9} {}_{x_i - x_{i-1}}L_{x_{i-1}} + \sum_{i=10}^{\infty} {}_{x_i - x_{i-1}}L_{x_{i-1}} \\ &= \sum_{i=1}^{9} {}_{x_i - x_{i-1}}L_{x_{i-1}} + T_{40} \\ &= (99684 + 397222 + \cdots + 490777.5) + 3941294.6 \\ &= 7896188.1. \end{aligned} $$

The complete life expectancy at birth is then

$$ \mathring{e}_0 = \frac{T_0}{l_0} = \frac{7896188.1}{100000} = 78.96 $$

years. And curtate life expectancy at birth is approximated as

$$ e_0 \approx \mathring{e}_0 - \frac{1}{2} = 78.46. $$

3. i) UDD implies a linear interpolation of the l_x function within the age range 40-45, hence

$$l_{41} = \frac{4}{5}l_{40} + \frac{1}{5}l_{45} = \frac{4}{5} \times 10000 + \frac{1}{5} \times 9883 = 9976.6.$$

ii) The event that a life aged exactly 40 years will die between ages 44 and 46 years requires that life to survive from age 40 to 44 years, then dying within 2 years (between ages 44 and 46 years) thereafter. Mathematically, the required probability is

$${}_4p_{40} \times {}_2q_{44} = \frac{l_{44}}{l_{40}}\left(1 - \frac{l_{46}}{l_{44}}\right),$$

where

$$\begin{aligned} l_{44} &= \frac{1}{5}l_{40} + \frac{4}{5}l_{45} = 9906.4 \\ l_{46} &= \frac{4}{5}l_{45} + \frac{1}{5}l_{50} = 9844.8, \end{aligned}$$

obtained using linear interpolations through the assumptions of UDD within age intervals 40-45 and 45-50 respectively. Hence,

$${}_4p_{40} \times {}_2q_{44} = \frac{9906.4}{10000}\left(1 - \frac{9844.8}{9906.4}\right) = 0.00616.$$

4. i) Suppose a_i and b_i are defined such that $x+a_i$ is the age at which individual i enters the observation, while $x+b_i$ is the age at which individual i exits the observation if death had not happened. Denote also y_i as the death indicator so that $y_i = 1$ if individual dies during the study and 0 otherwise. The following table can then be constructed.

Life i	y_i	a_i	b_i	$b_i - a_i$
1	0	0	$\frac{6}{12}$	$\frac{6}{12}$
2	0	$\frac{1}{12}$	1	$\frac{11}{12}$
3	1	$\frac{1}{12}$	1	$\frac{11}{12}$
4	0	$\frac{2}{12}$	1	$\frac{10}{12}$
5	1	$\frac{3}{12}$	1	$\frac{9}{12}$
6	0	$\frac{4}{12}$	1	$\frac{8}{12}$
7	1	$\frac{5}{12}$	1	$\frac{7}{12}$
8	0	$\frac{7}{12}$	1	$\frac{5}{12}$
9	1	$\frac{8}{12}$	1	$\frac{4}{12}$
10	0	$\frac{9}{12}$	1	$\frac{3}{12}$
	$d_{60} = 4$			$\sum_{i=1}^{10}(b_i - a_i) = \frac{74}{12}$

Using the approximation

$${}_{b-a}q_{x+a} \approx (b-a)q_x,$$

a simple estimator of q_x (using method of moments) is given by

$$\tilde{q}_x = \frac{d_x}{\sum_{i=1}^{10}(b_i - a_i)} = \frac{d_x}{E_x^0}.$$

Hence, we obtain

$$\tilde{q}_{60} = \frac{d_{60}}{\sum_{i=1}^{10}(b_i - a_i)} = \frac{4}{\frac{74}{12}} = 0.6486.$$

ii) Either, assuming a uniform distribution of deaths, ${}_tp_{60} = 1 - {}_tpq_{60} = 1 - tq_{60}$ and

$$\mu_{60+t} = -\frac{\mathrm{d}}{\mathrm{d}t}\log {}_tp_{60} = \frac{q_{60}}{1 - tq_{60}}$$

and therefore

$$\tilde{\mu}_{60} = \tilde{q}_{60} = 0.6486.$$

Or, assuming that the force of mortality is constant ($= \mu_{60}$) between exact ages 60 and 61 years, we have $p_{60} = \exp(-\mu_{60})$ and hence $\mu_{60} = -\log(1 - q_{60})$ and therefore

$$\tilde{\mu}_{60} = -\log(1 - 0.6486) = 1.045.$$

5. The relevant data are:

Life i	y_i	a_i	b'_i	$b'_i - a_i$
1	0	$\frac{9}{12}$	1	$\frac{3}{12}$
2	0	$\frac{3}{12}$	1	$\frac{9}{12}$
3	1	$\frac{4}{12}$	$\frac{10}{12}$	$\frac{6}{12}$
4	1	$\frac{2}{12}$	$\frac{5}{12}$	$\frac{3}{12}$
5	0	$\frac{5}{12}$	$\frac{8}{12}$	$\frac{3}{12}$
6	0	$\frac{6}{12}$	$\frac{10}{12}$	$\frac{4}{12}$
7	0	0	$\frac{7}{12}$	$\frac{7}{12}$
8	0	0	$\frac{3}{12}$	$\frac{3}{12}$
	$d_{70} = 2$			$\sum_{i=1}^{8}(b'_i - a_i) = \frac{38}{12}$

The estimator under the Poisson model is the same as the two-state model, and assumes constant force of mortality μ_x for $x \in [70, 71)$; hence

$$\hat{\mu}_{70} = \hat{\mu}_{70.5} = \frac{d_{70}}{E^C_{70}} = \frac{d_{70}}{\sum_{i=1}^{8}(b'_i - a_i)} = \frac{2}{\frac{38}{12}} \approx 0.6316.$$

6. i) We first obtain y_i, a_i, b_i and b'_i.

Life i	y_i	a_i	b_i	b'_i	$b_i - a_i$	$b'_i - a_i$
1	0	$\frac{11}{12}$	1	1	$\frac{1}{12}$	$\frac{1}{12}$
2	0	$\frac{10}{12}$	1	1	$\frac{2}{12}$	$\frac{2}{12}$
3	0	$\frac{5}{12}$	1	1	$\frac{7}{12}$	$\frac{7}{12}$
4	0	$\frac{4}{12}$	1	1	$\frac{8}{12}$	$\frac{8}{12}$
5	1	0	$\frac{11}{12}$	$\frac{2}{12}$	$\frac{11}{12}$	$\frac{2}{12}$
6	0	0	$\frac{9}{12}$	$\frac{9}{12}$	$\frac{9}{12}$	$\frac{9}{12}$
7	0	0	$\frac{7}{12}$	$\frac{7}{12}$	$\frac{7}{12}$	$\frac{7}{12}$
8	1	0	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{4}{12}$	$\frac{2}{12}$
	$d_{60} = 2$				$\sum(b_i - a_i) = \frac{49}{12}$	$\sum(b'_i - a_i) = \frac{38}{12}$

Notice that life 2 is not counted as dead for our study because death occurs after age 61. The likelihood for the observed data is given by

$$\begin{aligned} L &= \prod_{i=1}^{n} {}_{b_i - a_i}q_{x+a_i}^{y_i}(1 - {}_{b_i - a_i}q_{x+a_i})^{1-y_i} \\ &= (1 - {}_{\frac{1}{12}}q_{60\frac{11}{12}})(1 - {}_{\frac{2}{12}}q_{60\frac{10}{12}})(1 - {}_{\frac{7}{12}}q_{60\frac{5}{12}})(1 - {}_{\frac{8}{12}}q_{60\frac{4}{12}})\,{}_{\frac{11}{12}}q_{60}(1 - {}_{\frac{9}{12}}q_{60})(1 - {}_{\frac{7}{12}}q_{60})\,{}_{\frac{4}{12}}q_{60}. \end{aligned}$$

Under the assumption of UDD,

$${}_tq_x = tq_x \text{ for } t \in [0,1),$$

and

$${}_{b_i-a_i}q_{x+a_i} = \frac{(b_i-a_i)q_x}{1-a_iq_x} \Rightarrow 1 - {}_{b_i-a_i}q_{x+a_i} = \frac{1-b_iq_x}{1-a_iq_x}.$$

Therefore, the likelihood can be expressed in terms of q_{60}:

$$\begin{aligned} L(q_{60}) &= \frac{11}{12}q_{60} \times \frac{4}{12}q_{60} \times \frac{1-q_{60}}{1-\frac{11}{12}q_{60}} \times \frac{1-q_{60}}{1-\frac{10}{12}q_{60}} \times \frac{1-q_{60}}{1-\frac{5}{12}q_{60}} \times \frac{1-q_{60}}{1-\frac{4}{12}q_{60}} \\ &\quad \times (1-\frac{9}{12}q_{60}) \times (1-\frac{7}{12}q_{60}) \\ &\propto q_{60}^2(1-q_{60})^4(1-\frac{9}{12}q_{60})(1-\frac{7}{12}q_{60}) \times \frac{1}{1-\frac{11}{12}q_{60}} \times \frac{1}{1-\frac{10}{12}q_{60}} \times \frac{1}{1-\frac{5}{12}q_{60}} \times \frac{1}{1-\frac{4}{12}q_{60}} \\ &\propto \frac{q_{60}^2(1-q_{60})^4(12-9q_{60})(12-7q_{60})}{(12-11q_{60})(12-10q_{60})(12-5q_{60})(12-4q_{60})}. \end{aligned}$$

The log-likelihood is then

$$\begin{aligned} l(q_{60}) &= \log L(q_{60}) \\ &= C + 2\log q_{60} + 4\log(1-q_{60}) + \log(12-9q_{60}) + \log(12-7q_{60}) - \log(12-11q_{60}) \\ &\quad - \log(12-10q_{60}) - \log(12-5q_{60}) - \log(12-4q_{60}), \end{aligned}$$

(or anything equivalent) which can be maximized numerically w.r.t. q_{60} to obtain the m.l.e. $\hat{q}_{60}$.

ii) Using the approximation that

$${}_{b_i-a_i}q_{x+a_i} \approx (b_i-a_i)q_x$$

to yield a simple estimator

$$\tilde{q}_{60} = \frac{d_{60}}{\sum_{i=1}^{8}(b_i-a_i)} = \frac{2}{\frac{49}{12}} = 0.4898.$$

iii) Assuming UDD we have (see Question 4)

$$\tilde{\mu}_{60.5} = \frac{\tilde{q}_{60}}{1-\frac{1}{2}\tilde{q}_{60}} = 0.6487.$$

Or, assuming constant force of mortality, we have (see Question 4)

$$\tilde{\mu}_{60.5} = \tilde{\mu}_{60} = -\log(1-\tilde{q}_{60}) = -\log(1-0.4898) = 0.6730.$$

iv) Using the two-state model, the estimator is

$$\hat{\mu}_{60.5} = \frac{d_{60}}{E_{60}^C} = \frac{2}{\sum_{i=1}^{8}(b_i'-a_i)} = \frac{2}{\frac{38}{12}} = 0.6316.$$

Hence, the estimate obtained, 0.6316 is slightly different from either of the estimates in (iii). The chief difference is that both the estimates in (iii) rely on the approximation used in (ii) and do not account for actual times of death, whereas the estimate in (iii) is an exact m.l.e. and explicitly incorporates death dates.