B.1 Revision, lifetime distributions, Lexis diagrams and the census approximation

1. (a) The log likelihood function is given by

$$\ell(\lambda) = \log\left(\prod_{k=1}^{n}\left(\lambda e^{-\lambda L_k}\right)\right) = n\log\lambda - \lambda\sum_{k=1}^{n} L_k$$

We differentiate w.r.t. λ to get

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{k=1}^{n} L_k, \qquad \ell''(\lambda) = -\frac{n}{\lambda^2} < 0.$$

ℓ' has its unique zero for

$$\lambda = \hat{\lambda} = \frac{n}{L_1 + \ldots + L_n}$$

and this is a maximum of ℓ, since $\ell'' < 0$. Therefore $\hat{\lambda}$ maximizes ℓ.

(b) i. Just apply (a) to the data and get

$$\hat{\lambda} = \frac{20}{l_1 + \ldots + l_{20}} = 0.002917.$$

ii. The Fisher information is given by

$$I_n(\lambda) = -\mathbb{E}(\ell''(\lambda)) = \frac{n}{\lambda^2} \tag{B.1}$$

so that, approximately $\hat{\lambda} \sim \mathcal{N}(\lambda, \lambda^2/n) \approx \mathcal{N}(\lambda, \hat{\lambda}^2/n)$, therefore

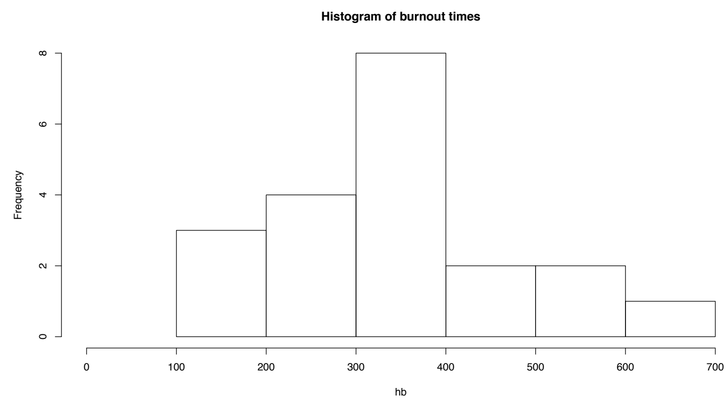
$$\begin{aligned} .95 = \mathbb{P}(|Z| < 1.96) &\approx \mathbb{P}\left(\left|\frac{\hat{\lambda}-\lambda}{\hat{\lambda}/\sqrt{n}}\right| < 1.96\right) \\ &= \mathbb{P}(\hat{\lambda} - 1.96\hat{\lambda}/\sqrt{n} < \lambda < \hat{\lambda} + 1.96\hat{\lambda}/\sqrt{n}) \\ &= \mathbb{P}\left(\frac{1}{\hat{\lambda} + 1.96\hat{\lambda}/\sqrt{n}} < \frac{1}{\lambda} < \frac{1}{\hat{\lambda} - 1.96\hat{\lambda}/\sqrt{n}}\right), \end{aligned}$$

so that $(\hat{\lambda} - 1.96\hat{\lambda}/\sqrt{n}, \hat{\lambda} + 1.96\hat{\lambda}/\sqrt{n}) = (0.001638, 0.004195)$ is an approximate 95% confidence interval for λ, and $(238.4, 610.3)$ an approximate 95% confidence interval for $1/\lambda$.

iii. Since $L_1 + \ldots + L_n \sim \Gamma(n, \lambda)$, we have $2\lambda(L_1 + \ldots + L_n) \sim \Gamma(n, \frac{1}{2}) = \chi^2_{2n}$ and so for $2n = 40$, and the lower and upper 2.5% quantiles

$$\begin{aligned} 0.95 = \mathbb{P}(24.43 < |X^2| < 59.34) &= \mathbb{P}(24.43 < 2\lambda n/\hat{\lambda} < 59.34) \\ &= \mathbb{P}\left(\frac{2n}{59.34\hat{\lambda}} < \frac{1}{\lambda} < \frac{2n}{24.43\hat{\lambda}}\right), \end{aligned}$$

so the exact 95% confidence interval for $\frac{1}{\lambda}$ is $(231.1, 561.3)$.



iv.

v. Expected numbers under Exp($\hat{\lambda}$) are $(e^{-100k} - e^{-100(k+1)})n$:

$$(5.1, 3.8, 2.8, 2.1, 1.6, 1.2, 0.9) \quad \text{and 2.6 for} > 700.$$

For the χ^2 test we require expected numbers above 5, so we keep the first bin, merge the next three to get 8.7 and the remainder to get 6.2 (alternatively merge next two and remainder). The data then is

Bin	0-100	100-400	400+	total
observed	0	15	5	20
expected	5.1	8.7	6.2	20

and we calculate the $\chi^2_{3-2} = \chi^2_1$ test statistic

$$\sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} = 9.84 \quad \Rightarrow \quad |Z| = \sqrt{\sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i}} = 3.14 \gg 1.96$$

so there is strong evidence against exponentiality.

2. (a) i. Identify the survival function of T as

$$\begin{aligned}\mathbb{P}(T > t) &= \mathbb{P}(T_1 > t, \ldots, T_m > t) = \mathbb{P}(T_1 > t) \ldots \mathbb{P}(T_m > t) \\ &= \mathrm{e}^{-H_1(t)} \cdot \cdots \cdot \mathrm{e}^{-H_m(t)} \\ &= \exp\left\{ - \int_0^t (h_1(s) + \cdots + h_m(s)) \, \mathrm{d}s \right\},\end{aligned}$$

where $H_k(s) = \int_0^s h_k(u)du$ is the cumulative hazard

ii. Define $\hat{T}_k = \min\{T_i : i \neq k\}$, and $f_k(s) = h_k(s)\mathrm{e}^{-H_k(s)}$ is the density of T_k. We have

$$\begin{aligned}\mathbb{P}(T > t \cap K = k) &= \int_t^\infty \mathbb{P}(\hat{T}_k > s) f_k(s) \, \mathrm{d}s \\ &= \int_t^\infty \mathrm{e}^{-\sum_{i=1}^m H_i(s)} h_k(s) \, \mathrm{d}s.\end{aligned}$$

The density of T on $\{K = k\}$ is then

$$-\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{P}(T > t \cap K = k) =$$

The conditional probability of $\{K = k\}$ is then

$$\mathbb{P}(K = k \,|T = t) = \frac{-\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T > t \cap K = k)}{\sum_{i=1}^m -\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T > t \cap K = i)} = \frac{h_k(t)}{\sum_{i=1}^m h_i(t)}.$$

iii. If the T_k are exponential then the hazard functions are constant, so $\mathbb{P}(K = k|T = t)$ does not vary with t.

(b) By (a), the hazard function of T now is $k_1 t^n + \cdots + k_m t^n = (k_1 + \cdots + k_m)t^n$, so T has a Weibull distribution with rate parameter $k = k_1 + \cdots + k_m$ and exponent n.

(c) We first calculate the survival function and let $\lambda \to 0$ to get

$$\bar{F}(t) = \mathbb{P}(T > t|T \leq \omega) = \frac{e^{-\lambda t} - e^{-\lambda\omega}}{1 - e^{-\lambda\omega}} \to \frac{\omega - t}{\omega},$$

which is the survival function of the uniform distribution on $[0, \omega]$. This is not surprising since the exponential density for small λ is very flat initially, also after truncation and renormalisation.

We calculate the hazard function of the truncated exponential distribution via the density

$$f(t) = \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda\omega}} \qquad \Rightarrow \qquad h(t) = \frac{\lambda}{1 - e^{-\lambda(\omega - t)}}. \tag{B.2}$$

3. We have $X \sim \exp(1)$. We want the distribution of $Y = \Lambda^{-1}(X)$.
If we let F_X and F_Y be the corresponding cdfs, we have $F_X(x) = \mathbb{P}\{X \leq x\} = 1 - e^{-x}$, so

$$\begin{aligned} F_Y(y) &= \mathbb{P}\{Y \leq y\} \\ &= \mathbb{P}\{\Lambda^{-1}(X) \leq y\} \\ &= \mathbb{P}\{X \leq \Lambda(y)\} \text{ because } \Lambda \text{ is strictly increasing} \\ &= 1 - e^{-\Lambda(y)}, \end{aligned}$$

which is the cdf of a random variable with hazard rate λ.

4. (a) The full table of deaths d_x, lives at risk ℓ_x and total time at risk $\tilde{\ell}_x$ aged x is

x	0	1	2	3	4
d_x	45	9	8	4	3
ℓ_x	69	24	15	7	3
$\tilde{\ell}_x$	35.63	19.15	10.20	4.94	1.32

(b) The discrete method based on curtate lifetimes $K^{(1)}, \ldots, K^{(n)}$, $n = 69$, factorises the likelihood

$$\prod_{j=1}^{69} p_K(K^{(j)}) = \prod_{x=0}^{\infty} (1 - q_x)^{\ell_x - d_x} q_x^{dx} \tag{B.3}$$

and differentiation of each factor leads to maximum likelihood estimators $\hat{q}_x^{(0)} = d_x/\ell_x$.
The continuous method based on $T^{(1)}, \ldots, T^{(n)}$, $n = 69$, and the assumption of constant forces of mortality between integer ages, factorises the likelihood

$$\prod_{j=1}^{69} f_T(T^{(j)}) = \prod_{x=0}^{\infty} \mu_{x+\frac{1}{2}}^{d_x} \exp\left\{-\tilde{\ell}_x \mu_x\right\} \tag{B.4}$$

and differentiation of each factor leads to maximum likelihood estimators $\hat{q}_x = 1 - \exp\{-d_x/\tilde{\ell}_x\}$.

(c) From the formulas obtained in (b) we calculate

x	0	1	2	3	4
$\hat{q}_x^{(0)}$	0.65	0.38	0.53	0.57	1
$\hat{q}_x$	0.717	0.375	0.543	0.555	0.898

$\hat{q}_0^{(0)} < \hat{q}_0$ since the total time $\tilde{\ell}_0$ spent at risk is very short. We can see this directly from the data. Most subject dying in the first year die very early (e.g. three subjects die the day after their transplant). This actually suggests that the force of mortality is not constant over the first year, but much higher initially.

$\hat{q}_4 < \hat{q}_4^{(0)} = 1$ allows survival beyond the maximal observed age under the continuous method. The specification of the distribution estimate is not complete, but with no data we get no estimate. Some methods of graduation will allow to extrapolate beyond maximal age.

By both methods, the one-year survival probabilities indicate a bathtub behaviour, decreasing initially and then increasing.

(d) i. Under the estimates from curtate lifetimes and the assumption of independent uniform fractional part,

$$\mathbb{P}(T > 0.25) = \mathbb{P}(T > 1) + \mathbb{P}(K = 0, S > 0.25)$$
$$\text{is estimated by} \qquad (1 - \hat{q}_0^{(0)}) + \hat{q}_0^{(0)}\frac{3}{4} = 0.837.$$

ii. Under the estimates from continuous lifetimes and the assumption of constant force of mortality between integer ages

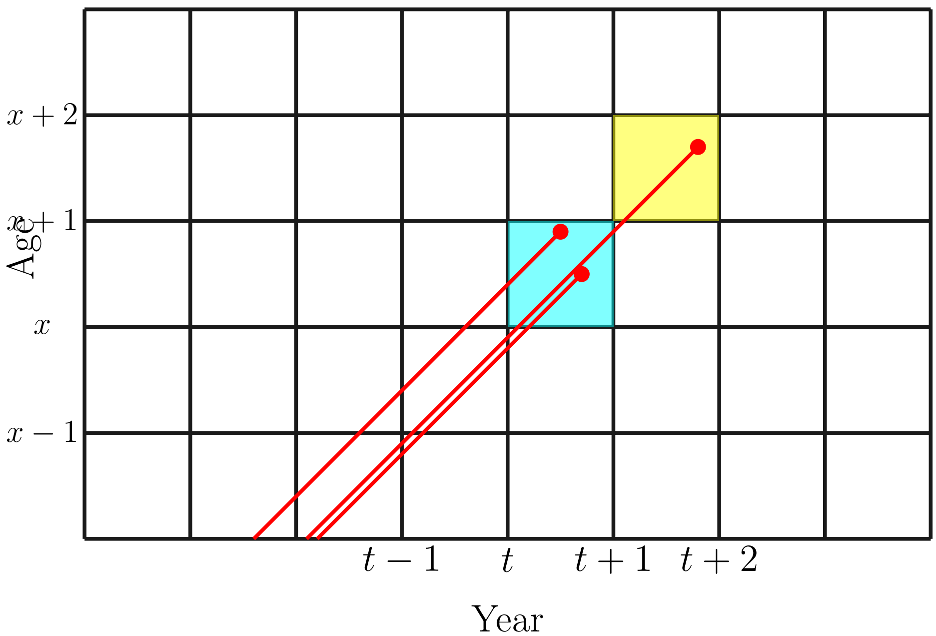
$$\mathbb{P}(T > 0.25) = \exp\left\{-\int_0^{0.25} \mu_t dt\right\}$$
$$\text{is estimated by} \qquad \exp\{-0.25\hat{\mu}_{0+\frac{1}{2}}\} = (1 - \hat{q}_0)^{0.25} = 0.729.$$

iii. Again we can apply the discrete or continuous method (formally for units of three months). The continuous method assumes constancy of forces of mortality over each three-month period and gives an estimate

$$\exp\left\{-\frac{d}{4\tilde{\ell}}\right\} = \exp\left\{-\frac{31}{4 \times 12.58}\right\} = 0.540.$$

Here, $4\tilde{\ell}$ is the total number of time units (as calculated from years $\tilde{\ell}$) at risk during first three-month unit.
The discrete method is based on one-unit death probabilities and gives $1 - d/\ell = 1 - 31/69 = 0.551$ as an estimate for the first-unit survival probability.
These estimates are much smaller reflecting a higher risk to die initially. In fact, this suggests that neither assumption i. nor ii. is optimal. An initially decreasing force of mortality would be better.

5. (a) i. The turquoise region corresponds to age x in year t. The same individuals are age $x + 1$ in year $t + 1$, and this portion of their lifelines falls in the yellow region.



ii. Just write the quantities as integrals and sums and interchange the order of integration and summation:

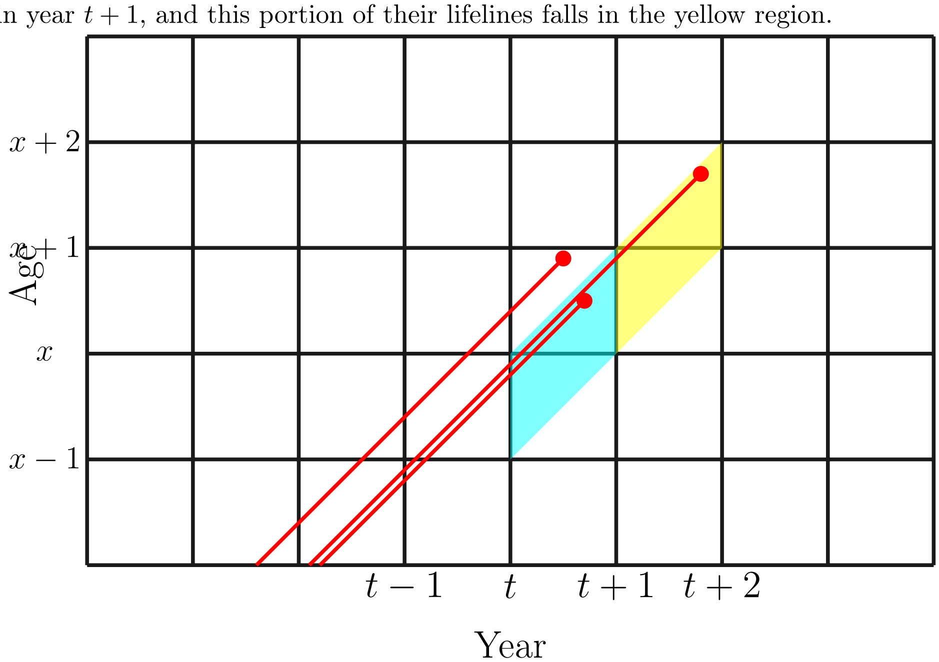
$$E_x^c = \sum_{i=1}^{n} \int_{a_i}^{b_i} dt = \int_K^{K+N} \sum_{i=1}^{n} 1_{\{a_i \le t \le b_i\}} dt = \int_K^{K+N} P_{x,t} dt. \tag{B.5}$$

iii. Under the assumption of piecewise linearity we calculate

$$E_x^c = \sum_{k=K}^{K+N-1} \int_0^1 (rP_{x,k} + (1-r)P_{x,k+1}) dr = \sum_{k=K}^{K+N-1} \frac{P_{x,k} + P_{x,k+1}}{2}. \tag{B.6}$$

The assumption of piecewise linear $P_{x,t}$ cannot hold exactly since $P_{x,t} \in \mathbb{N}$, but for large n this is negligible.

(b) i. The turquoise region corresponds to age x in year t. The same individuals are age $x+1$ in year $t+1$, and this portion of their lifelines falls in the yellow region.



ii. We can define

$$P^{(2)}_{x,t} = \#\text{ lives at risk at time } t \text{ with } x\text{th birthday in calendar year } \lfloor t \rfloor, \tag{B.7}$$

but also ought to adjust

$$E^{c,2}_x = \int_K^{K+N} P^{(2)}_{x,t}\,dt \approx \sum_{k=K}^{K+N-1} \frac{P_{x,k} + P_{x+\mathbf{1},k+1}}{2} \tag{B.8}$$

since it is more natural to assume that the cohort of lives $P_{x,k}$ with x-th birthday in calendar year k changes linearly to $P_{x+1,k+1}$ since this will count the same people (being age k at the start of year x and age $k+1$ at the start of year $x+1$).

iii. The first estimate is approximating $\int_x^{x+1} \mu_s\,\mathrm{d}s$; that is, the average of μ_s over the age interval from x to $x+1$, assuming age-specific mortality doesn't change with time. (This is just μ_x if μ_s is assumed constant on $[x, x+1)$.) The second estimate clearly includes the experience of individuals at ages between $x-1$ and $x+1$, but weighted toward the middle (in proportion to the width of the parallelogram in the above figure.) In fact, the estimate $\tilde{\mu}_x = d^{(2)}_x / E^{c,2}_x$ will be approximating

$$\int_{x-1}^{x+1} \Big(1 - \big|s - x\big|\Big)\,\mu_s\,\mathrm{d}s.$$

6. (a) Let us write e_x and p_x for the figures in the table, and $\tilde{e}_x$ and $\tilde{p}_x$ for the figures after we change the rates in the first two years.

We have

$$e_0 = p_0\left(1 + p_1\left(1 + e_2\right)\right),$$
$$\tilde{e}_0 = \tilde{p}_0\left(1 + \tilde{p}_1\left(1 + \tilde{e}_2\right)\right)$$

Since we change only the rates before year 2, we have $e_2 = \tilde{e}_2$. Then solving for $\tilde{e}_0$, we have

$$\tilde{e}_0 = \tilde{p}_0\left(1 + \tilde{p}_1\left(\frac{e_0/p_0 - 1}{p_1}\right)\right).$$

With $e_0 = 44.83$, $p_0 = 0.839$, $p_1 = 0.946$, $\tilde{p}_0 = 0.995$, $\tilde{p}_1 = 0.9996$ we obtain $\tilde{e}_0 = 56.12$, an increase of 11.29 years.

(b) It is clear from the table that the rates q_x $(x = 19, \ldots, 25)$ are much larger than we would normally expect. Comparing them with the rates just before and after, it looks like a plausible first approximation would be to replace all these rates by a rate around 0.0035. So we could work with a model where the mortality is constant with $q = 0.0035 = 1 - 0.9965$ for those 7 years. (Of course this is very rough, and there are all sorts of things we ignore including the various effects of the war on mortality, even after it had finished).

We can represent e_0 as a sum $A + B + C$ of three terms:

$$A = \text{expected number of whole years lived up to age 19},$$
$$B = \text{expected number of whole years lived up between 19 and 26},$$
$$C = \text{expected number of whole years lived after age 26}.$$

Let us write $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{C}$ for the new values once we change the rates.

The change to the rates between ages 19 and 26 makes no difference to the first term, so $\widetilde{A} = A$.

We have

$$\begin{aligned}
C &= {}_{26}p_0 e_{26} \\
&= \frac{\ell_{26}}{\ell_0} e_{26} \\
&= 0.6014 \times 42.75 \\
&= 25.71.
\end{aligned}$$

The change to the rates makes no difference to e_{26}, but we have a new value for the probability of surviving to age 26, giving

$$\begin{aligned}
\widetilde{C} &= (0.9965)^7 \frac{\ell_{19}}{\ell_0} e_{26} \\
&= (0.9965)^7 \times 0.7309 \times 42.75 \\
&= 30.49.
\end{aligned}$$

Finally, we can find B using $B + C = e_{19}\ell_{19}/\ell_0$ as in the previous calculation, so

$$\begin{aligned}
B &= \frac{\ell_{19}}{\ell_0} e_{19} - C \\
&= 0.7309 \times 41.57 - 25.71 \\
&= 4.67.
\end{aligned}$$

In the new model with constant rate between age 19 and age 26, we can use

$$\begin{aligned}
\widetilde{B} &= {}_{19}\tilde{p}_0 \left({}_1\tilde{p}_{19} + {}_2\tilde{p}_{19} + \cdots + {}_7\tilde{p}_{19} \right) \\
& \frac{\ell_{19}}{\ell_0} \sum_{k=1}^{7} 0.9965^k \\
&= \frac{\ell_{19}}{\ell_0} \frac{0.9965 - 0.9965^8}{0.0035} \\
&= 0.7381 \times 6.903 \\
&= 5.10.
\end{aligned}$$

The total change in life expectancy at birth is $\widetilde{B} + \widetilde{C} - B - C$ which comes to $30.49 + 5.10 - 25.71 - 4.67 = 5.21$, giving a new life expectancy of around 50.