A.2 Estimation of lifetime distributions

1. List some of the reasons for changes in mortality rates over the last 150 years (in, for example, the UK population, but of course more widely if you like).

2. The survival times (in days after transplant) for the original $n = 69$ members of the Stanford Heart Transplant Program were as follows:

Survival time after heart transplant (days)									
15	3	624	46	127	64	1350	280	23	10
1024	39	730	136	1775	1	836	60	1536	1549
54	47	51	1367	1264	44	994	51	1106	897
253	147	51	875	322	838	65	815	551	66
228	65	660	25	589	592	63	12	499	305
29	456	439	48	297	389	50	339	68	26
30	237	161	14	167	110	13	1	1	

This data is also at `http://www.stats.ox.ac.uk/~martin/BS3b/StanfordHeart`.

The aim of this exercise is to construct the associated lifetable.

(a) Complete the following table of counts d_x of associated curtate survival times (in years=365 days), counts ℓ_x of subjects alive exactly x years after their transplant, total time $\tilde{\ell}_x$ spent alive between x and $x+1$ years after their transplant, by all subjects:

x	0	1	2	3	4
d_x			8	4	3
ℓ_x					
$\tilde{\ell}_x$		19.148	10.203	4.937	1.315

(b) Use the discrete method to calculate maximum likelihood estimates of q_x, $x = 0, \ldots, 4$. Use the continuous method to calculate maximum likelihood estimates of $\mu_{x+1/2}$, $x = 0, \ldots, 4$, under the assumption of constant mortality throughout each year, and the induced estimates of q_x. Comment on the differences.

(c) Estimate the probability to survive for 3 months

i. assuming fractional and integer parts of lifetimes are independent, and the fractional part is uniform;

ii. assuming the force of mortality is constant over the first year;

iii. directly from the data using the discrete method.

3. In a certain population, the force of mortality of lifetimes T is believed to be constant over ages $x_{j-1} \leq x < x_j$, $j \geq 1$, where $x_0 = 0$. Denote these unknown constants by γ_j, $j \geq 1$. You observe n full lifetimes $T^{(1)}, \ldots, T^{(n)} \sim T$ sampled from this population.

(a) Determine the likelihood function of the sample, in terms of the parameters γ_j, $j \geq 1$.

(b) Let L_j be the total time spent alive between ages x_{j-1} and x_j. Express L_j explicitly in terms of $T^{(1)}, \ldots, T^{(n)}$.

(c) Show that a maximum likelihood estimator for $\gamma_j \in [0, \infty)$, $j \geq 1$, is given by

$$\hat{\gamma}_j = \frac{D_j}{L_j} \text{ if } L_j > 0.$$

where D_j is the number of deaths between ages x_{j-1} and x_j.

(d) Denote $h_j = \mathbb{P}(T \leq x_j | T > x_{j-1})$, $j \geq 1$. Express h_j, $j \geq 1$, in terms of γ_j, $j \geq 1$ and deduce maximum likelihood estimators for the new parameters.

(e) Discretise $K = \sup\{x_j : j \geq 0, x_j \leq T\}$ and express the probability mass function p_K of K in terms of h_j, $j \geq 1$.

(f) Derive maximum likelihood estimators for h_j based on an observation of the discrete data $K^{(1)}, K^{(2)}, \ldots, K^{(n)}$.

4. Erickson *et al.* analysed 22 skeletons of *A. sarcophagus.* The observed (curtate) ages at death in years were 2,4,6,8,9,11,12,13,14,14,15,15,16,17,17,18,19,19,20,21,23,28.

(a) Estimate directly the (curtate) life expectancy of this population.

(b) Construct an approximate 95% confidence interval for this life expectancy.

(c) We estimated one-year death probabilities by $\hat{q}_x^d = d_x/\ell_x$. Show that the life expectancy predicted from this estimated distribution must be the same as that computed directly from the observed lifetimes.

(d) Using the approximation that the fractional part of the lifetimes are independent of the integer part and uniformly distributed, adapt your answers to (a) and (b) for the full life expectancy.

(e) Estimate mortality rates for the population, under the assumption that these rates are constant over five-year intervals $[0, 5]$, $[5, 10]$, etc. (Consider two approaches: either use the "uniform fractional part" assumption to approximate the total exposed to risk $\tilde{\ell}_x$ given the curtate lifetimes, or consider directly the likelihood of the observed curtate data). Show how to estimate again the life expectancy for the population, based on the life table obtained.

5. Attached is an excerpt from a cohort life table for men in England and Wales born in 1894, including curtate life expectancies. (Data from the Human Mortality Database at `http://www.mortality.org`). Using the given data:

(a) Estimate the change to e_0, the curtate life expectancy at birth, if the mortality rate in the first two years of life were reduced to modern-day levels (say $q_0 = 0.005$, $q_1 = 0.0004$).

(b) Make a rough estimate of the change to e_0 if the increases in mortality due to the 1914-18 war and the 1918-19 influenza pandemic had not occurred.

AGE x	ℓ_x	q_x	e_x
0	100000	0.16134	44.83
1	83866	0.05398	52.39
	$\vdots$	$\vdots$	
14	74067	0.0022	45.99
15	73904	0.00237	45.09
16	73729	0.0026	44.2
17	73538	0.00301	43.31
18	73316	0.00313	42.44
19	73087	0.00787	41.57
20	72512	0.01836	40.9
21	71181	0.03218	40.65
22	68890	0.04424	40.98
23	65842	0.06194	41.86
24	61764	0.02088	43.59
25	60474	0.00551	43.51
26	60141	0.00385	42.75
27	59910	0.00384	41.91
28	59680	0.00391	41.07
29	59446	0.00377	40.23
30	59222	0.00386	39.38
31	58994	0.00367	38.53
32	58777	0.0038	37.67
33	58554	0.00399	36.81
34	58320	0.00445	35.96
35	58061	0.0046	35.11

Cohort life table for male births in 1894 in England and Wales