Lecture 6

Comparing life tables

6.1 The Poisson model

Under the assumption of a constant hazard rate (force of mortality) $\mu_{x+\frac{1}{2}}$ over the year $(x, x+1]$, we may view the estimation problem as a chain of separate hazard rate estimation problems, one for each year of life. Each individual lives some portion of a year in the age interval $(x, x+1]$, the portion being 0 (if he dies before birthday x), 1 (if he dies after birthday $x+1$), or between 0 and 1 if he dies between the two birthdays. Suppose now we lay these intervals end to end, with a mark at the end of an interval where an individual died. It is not hard to see that what results is a Poisson process on the interval $[0, E_x^c]$, where E_x^c is the total observed years at risk.

Suppose we treat E_x^c as though it were a constant. Then if D_x represents the numbers dying in the year the model uses

$$\mathrm{P}\big\{D_x = k\big\} = \frac{\Big(\mu_{x+\frac{1}{2}} E_x^c\Big)^k \mathrm{e}^{-\mu_{x+\frac{1}{2}} E_x^c}}{k!}, \quad k = 0, 1, 2, \cdots$$

which is an approximation to the 2-state model, and which in fact yields the same likelihood.

The estimator for the constant force of mortality over the year is

$$\widetilde{\mu}_{x+\frac{1}{2}} = \frac{D_x}{E_x^c}, \quad \text{with estimate } \frac{d_x}{E_x^c}\ .$$

Under the Poisson model we therefore have that

$$\mathrm{var}\widetilde{\mu}_{x+\frac{1}{2}} = \frac{\mu_{x+\frac{1}{2}} E_x^c}{(E_x^c)^2} = \frac{\mu_{x+\frac{1}{2}}}{E_x^c}\ .$$

So the estimate will be

$$\mathrm{var}\widetilde{\mu}_{x+\frac{1}{2}} \approx \frac{d_x}{(E_x^c)^2}\ .$$

If we compare with the **2-state stochastic model** over year $(x, x+1)$, assuming constant $\mu = \mu_{x+\frac{1}{2}}$, then the likelihood is

$$L = \prod_1^n \mu^{\delta_i} \mathrm{e}^{-\mu t_i}\ ,$$

where $\delta_i = 1$ if life i dies and $t_i = b_i - a_i$ in previous terminology (see the binomial model). Hence

$$L = \mu^{d_x} \mathrm{e}^{-\mu E_x^c}$$

and so

$$\widehat{\mu} = \frac{D_x}{E_x^c} \ .$$

The estimator is exactly the same as for the Poisson model except that both D_x and E_x^c are random variables. Using asymptotic likelihood theory we see that the estimate for the variance is

$$\text{var}\widehat{\mu} \approx \frac{\mu^2}{d_x} \approx \frac{d_x}{(E_x^c)^2} \ .$$

6.2 Testing hypotheses for q_x and $\mu_{x+\frac{1}{2}}$

We note the following normal approximations:

(i) Binomial model:

$$D_x \sim B(E_x, q_x) \implies D_x \sim N(E_x q_x, E_x q_x\,(1 - q_x)$$

and

$$\widehat{q}_x = \frac{D_x}{E_x} \sim N\left(q_x, \frac{q_x\,(1 - q_x)}{E_x}\right) \ .$$

(ii) Poisson model or 2-state model:

$$D_x \sim N(E_x^c \mu_{x+\frac{1}{2}}, E_x^c \mu_{x+\frac{1}{2}})$$

and

$$\widehat{\mu}_{x+\frac{1}{2}} \sim N\left(\mu_{x+\frac{1}{2}}, \frac{\mu_{x+\frac{1}{2}}}{E_x^c}\right) \ .$$

Tests are often done using comparisons with a published **standard life table.** These can be from national tables for England and Wales published every 10 years, or insurance company data collected by the Continuous Mortality Investigation Bureau, or from other sources. (It needs to be a source

A superscript "s" denotes "from a standard table", such as q_x^s and $\mu^s_{x+\frac{1}{2}}$.

Test statistics are generally obtained from the following:

Binomial:

$$z_x = \frac{d_x - E_x q_x^s}{\sqrt{E_x q_x^s\,(1 - q_x^s)}} \qquad \left(\approx \frac{O - E}{\sqrt{V}}\right)$$

Poisson/2-state:

$$z_x = \frac{d_x - E_x^c \mu^s_{x+\frac{1}{2}}}{\sqrt{E_x^c \mu^s_{x+\frac{1}{2}}}} \quad \left(\approx \frac{O - E}{\sqrt{V}}\right) \ .$$

Both of these are denoted as z_x since under a null hypothesis that the standard table is adequate $Z_x \sim N(0, 1)$ approximately.

6.2.1 The tests

χ^2 test

We take

$$X = \sum_{\text{all ages } x} z_x^2$$

This gives the sum of squares of standard normal random variables under the null hypothesis and so is a sum of $\chi^2(1)$. Therefore

$$X \sim \chi^2(m)\ ,\ \text{if } m = \#\ \text{years of study.}$$

H_0 : there is no difference between the standard table and the data,
H_A : they are not the same.
It is normal to use 5% significance and so the test fails if $X > \chi^2(m)_{0.95}$.
It tests large deviations from the standard table.
Disadvantages:
1. There may be a few large deviations offset by substantial agreement over part of the table. The test will not pick this up.
2. There might be bias, that is, although not necessarily large, all the deviations may be of the same sign.
3. There could be significant groups of consecutive deviations of the same sign, even if not overall.

Standardised deviations test

This tries to address point 1 above. Noting that each z_x is an observation from a standard normal distribution under H_0, the real line is divided into intervals, say 6 with dividing points at $-2, -1, 0, 1, 2$. The number of z_x in each interval is counted and compared with the expected number from a standard normal distribution. The test statistic is then

$$X = \sum_{\text{intervals}} \frac{(O-E)^2}{E} \sim \chi^2(5)$$

under the null hypothesis since this is Pearson's statistic. The problem here is that m is unlikely to be large enough to give approximate validity to the chi-square distribution. So this test is rarely appropriate.

Signs test

Test statistic X is given by

$$X = \#\{z_x > 0\}$$

Under the null hypothesis $X \sim B(m, \frac{1}{2})$, since the probability of a positive sign should be 1/2. This should be administered as a two-tailed test. It is under-powered since it ignores the size of the deviations but it will pick up small deviations of consistent sign, positive or negative, and so it addresses point 2 above.

Cumulative deviations test

This again addresses point 2 and essentially looks very similar to the logrank test between two survival curves. If instead of squaring $d_x - E_x q_x^s$ or $d_x - E_x^c \mu_{x+\frac{1}{2}}^s$, we simply sum then

$$\frac{\sum (d_x - E_x q_x^s)}{\sqrt{\sum E_x q_x^s (1 - q_x^s)}} \sim N(0,1), \quad \text{approximately}$$

and

$$\frac{\sum \left(d_x - E_x^c \mu_{x+\frac{1}{2}}^s\right)}{\sqrt{\sum E_x^c \mu_{x+\frac{1}{2}}^s}} \sim N(0,1) \quad \text{approximately.}$$

H_0 : there is no bias
H_A : there is a bias.

This test addresses point 2 again, which is that the chi-square test does not test for consistent bias.

Other tests

There are tests to deal with consecutive bias/runs of same sign. These are called the groups of signs test and the serial correlations test. Again a very large number of years, m, are required to render these tests useful.

6.2.2 An example

Table 6.2.2 presents imaginary data for men aged 90 to 95. The column ℓ_x lists the initial at risk, the number of men in the population on the census date, and d_x is the number of deaths from this initial population over the course of the year. E_x^c is the central at risk, estimated as $\ell_x - d_x/2$. Standard male British mortality for these ages is listed in column μ_x^s. (The columen $\mathring{\mu}_x$ is a graduated estimate, which will be discussed in section 6.3.

age	ℓ_x	d_x	E_x^c	$\hat{\mu}_{x+\frac{1}{2}}$	μ_x^s	z_x	$\mathring{\mu}_x$
90	40	10	35	0.29	0.202	1.1	0.25
91	35	8	31	0.258	0.215	0.52	0.28
92	22	4	18	0.20	0.236	−0.33	0.335
93	14	6	11	0.545	0.261	1.85	0.40
94	11	4	9	0.444	0.279	0.94	0.45
95	7	3	5.5	0.545	0.291	1.11	0.48

Table 6.1: Table of mortality rates for an imaginary old-people's home, with standard British male mortality given as μ_x^s, and graduated estimate $\mathring{\mu}_x$.

We note substantial differences between the estimates $\hat{\mu}_x$ and the standard mortality μ_x^s, but none of them is extremely large relative to the standard error: The largest z_x is 1.85. We test the two-sided alternative hypothesis, that the mortality rates in the old-people's home are different from the standard mortality rates, with a χ^2 test, adding up the z_x^2. The observed X^2 is 7.1, corresponding to an observed significance level $p = 0.31$. (Remember that we have 6 degrees of freedom, not 5, because these z_x are independent. This is not an incidence table.)

6.3 Graduation

Graduation is what statisticians would call "smoothing". Suppose that a company has collected its own data, producing estimates for either q_x or $\mu_{x+\frac{1}{2}}$. The estimates may be rather irregular from year to year and this could be an artefact of the population the company happens to have in a particular scheme. The underlying model should probably (but not necessarily) be smoother than the raw estimates. If it is to be considered for future predictions, then smoothing should be considered. This is called graduation.

There is always a tradeoff in smoothing procedures. Without smoothing, real patterns get lost in the random noise. Too much smoothing, though, can swamp the data in the model, so that the final estimate reflects more our choice of model than any truth gleaned from the data.

6.3.1 Parametric models

We may fit a formula to the data. Possible examples are

$$\mu_x = \mu \qquad \text{(Exponential);}$$
$$\mu_x = Be^{\theta x} \qquad \text{(Gompertz);}$$
$$\mu_x = A + Be^{\theta x} \qquad \text{(Makeham)}$$

The Gompertz can be a good model for a population of middle older age groups. The Makeham model has an extra additive constant which is sometimes used to model "intrinsic mortality", which is supposed to be independent of age. We could use more complicated formulae putting in polynomials in x.

6.3.2 Reference to a standard table

Here q_x^0, μ_x^0 represent the graduated estimates. We could have a linear dependence

$$q_x^0 = a + bq_x^s, \quad \mu_x^0 = a + b\mu_x^s$$

or possibly a translation of years

$$q_x^0 = q_{x+k}^s, \quad \mu_x^0 = \mu_{x+k}^s$$

In general there will be some assigned functional dependence of the graduated estimate on the standard table value. These are connected with the notions of accelerated lifetimes and proportional hazards, which will be central topics in the second part of the course.

6.3.3 Nonparametric smoothing

We effectively smooth our data when we impose the assumption that mortality rates are constant over a year. We may tune the strength of smoothing by requiring rates to be constant over longer intervals. This is a form of local averaging, and there are more and less sophisticated versions of this. In `Matlab` or `R` the methods available include kernel smoothing, orthogonal polynomials, cubic splines, and LOESS. These are beyond the scope of this course.

In Figure 6.1 we show a very simple example. The mortality rates are estimated by individual years or by lumping the data in five year intervals. The green line shows a moving average of the one-year estimates, in a window of width five years.

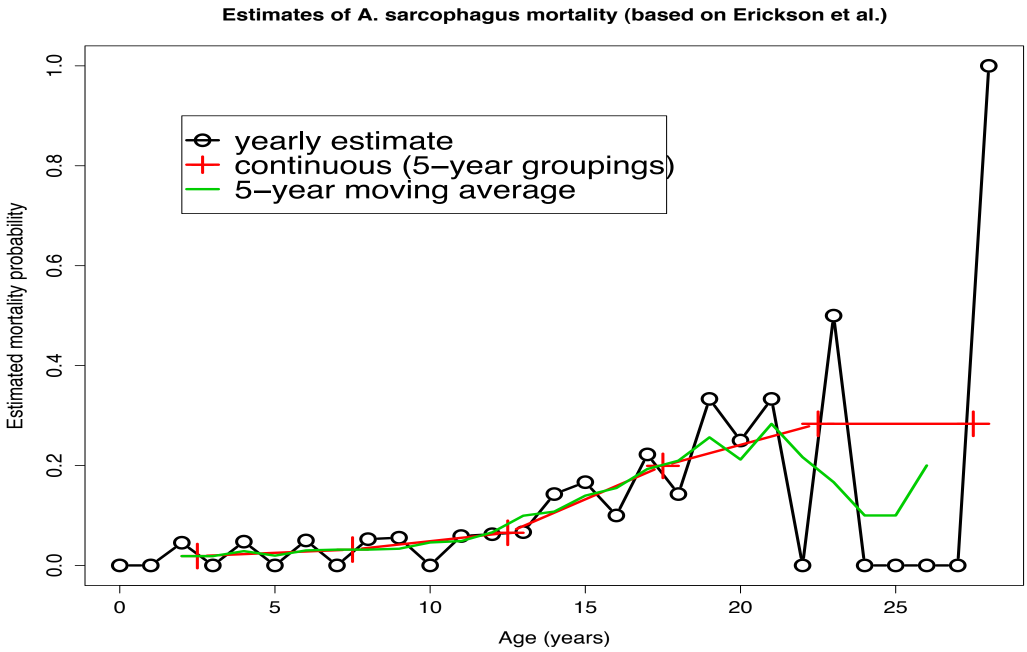

Figure 6.1: Different smooothings for *A. sarcophagus* mortality from Table 1.1.

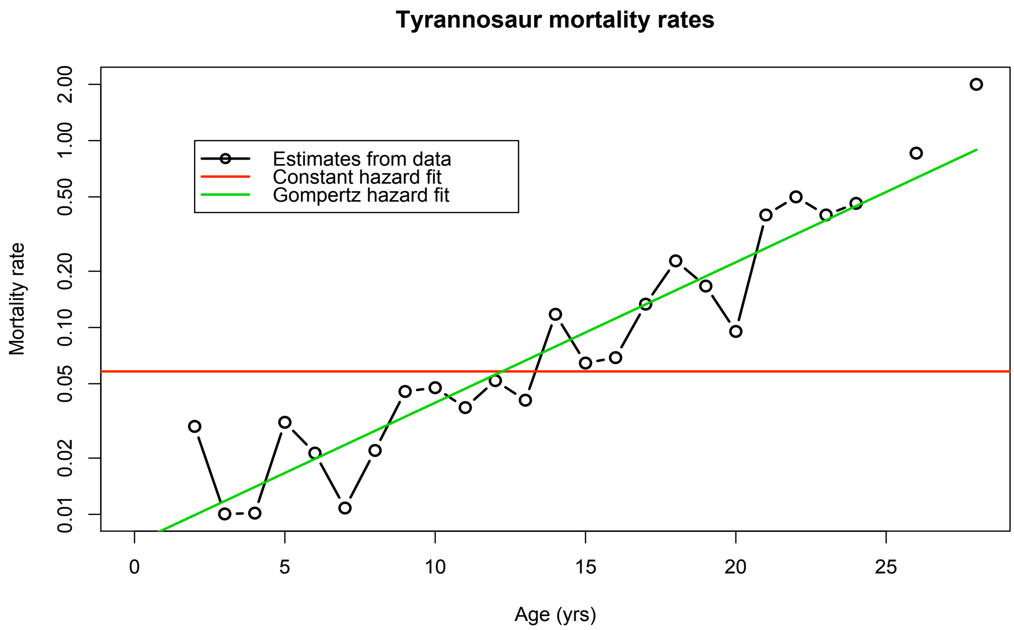

Figure 6.2: Estimated tyrannosaurus mortality rates from Table 4.3, together with exponential and Gompertz fits.

6.3.4 Methods of fitting

1. In any of the models (binomial, Poisson, 2-state) set (say) $q_x = a + bq_x^s$ in the likelihood and use maximum likelihood estimators for the unknown parameters a, b and similarly for μ_x and other functional relationships with the standard values.

2. Use weighted least squares and minimise

$$\sum_{\text{all ages } x} w_x \left(\widehat{q}_x - q_x^0\right)^2 \qquad \text{or}$$
$$\sum_{\text{all ages } x} w_x \left(\widehat{\mu}_{x+\frac{1}{2}} - \mu^0_{x+\frac{1}{2}}\right)^2$$

as appropriate. For the weights suitable choices are either E_x or E_x^c respectively. Alternatively we can use $1/\text{var}$, where the variance is estimated for $\widehat{q}_x$ or $\widehat{\mu}_{x+\frac{1}{2}}$, respectively.

The hypothesis tests we have already covered above can be used to test the graduation fit to the data, replacing $q_x^s, \mu^s_{x+\frac{1}{2}}$ by the graduated estimates. **Note that in the χ^2 test we must reduce the degrees of freedom of the χ^2 distribution by the number of parameters estimated in the model for the graduation.** For example if $q_x^0 = a + bq_x^s$, then we reduce the degrees of freedom by 2 as the parameters a, b are estimated.

6.3.5 Examples

Standard life table

We graduate the estimates in Table 6.2.2, based on the standard mortality rates listed in the column μ_x^s, using the parametric model $\mathring{\mu}_x = a + b\mu_x^s$. The log likelihood is

$$\ell = \sum d_x \log \mathring{\mu}_{x+\frac{1}{2}} - \mathring{\mu}_{x+\frac{1}{2}} E_x^c.$$

We maximise by solving the equations

$$0 = \frac{\partial \ell}{\partial a} = \sum \left(\frac{d_x}{\hat{a} + \hat{b}\mu^s_{x+\frac{1}{2}}} - E_x^c \right)$$
$$0 = \frac{\partial \ell}{\partial b} = \sum \left(\frac{d_x \mu^s_{x+\frac{1}{2}}}{\hat{a} + \hat{b}\mu_x^s} - \mu^s_{x+\frac{1}{2}} E_x^c \right).$$

We can solve these equations numerically, to obtain $\hat{a} = -0.279$ and $\hat{b} = 2.6$. This yields the graduated estimates $\mathring{\mu}$ tabulated in the final column of Table 6.2.2. Note that these estimates have the virtue of being, on the one hand, closer to the observed data than the standard mortality rates; on the other hand smoothly and monotonically increasing.

If we had used ordinary least squares to fit the mortality rates, we would have obtained very different estimates: $\tilde{a} = -0.472$ and $\tilde{b} = 3.44$, because we would be counting the Weighted least squares, with weights proportional to E_x^c (inverse variance) solves this problem, more or less, and gives us estimates $\hat{a}^* = -0.313$ and $\hat{b}^* = 2.75$ very close to the MLE.

In Figure 6.2 we plot the mortality rate estimates for the complete population of tyrannosaurs described in Table 1.1, on a logarithmic scale, together with two parametric model fits: the exponential model, with one parameter μ estimated by

$$\hat{\mu} = \frac{1}{\bar{t}} = \frac{n}{t_1 + \cdots + t_n} \approx \frac{n}{k_1 + \cdots + k_n + n/2} = 0.058,$$

where $t_1, \dots, t_n$ are the n lifetimes observed, and $k_i = \lfloor t_i \rfloor$ the curtate lifetimes; and the Gompertz model $\mu_s = Be^{\theta s}$, estimated by

$$\begin{aligned} \hat{\theta} \text{ solves } \frac{Q'(\hat{\theta})}{Q(\hat{\theta}) - 1} - \frac{1}{\hat{\theta}} &= \bar{t}, \\ \hat{B} &:= \frac{\hat{\theta}}{Q(\hat{\theta}) - 1}, \\ \text{where } Q(\theta) &:= \frac{1}{n} \sum e^{\theta t_i}. \end{aligned}$$

This yields $\hat{\theta} = 0.17$ and $\hat{B} = 0.0070$. It seems apparent to the eye that the exponential fit is quite poor, while the Gompertz fit might be pretty good. It is hard to judge the fit by eye, though, since the quality of the fit depends in part on the number of individuals at risk that go into the individual mortality-rate estimates, something which does not appear in the plot.

To test the hypothesis, we compute the predicted number of deaths in each age class $d_x^{(\mathrm{exp})} = l_x \cdot q_x^{(\mathrm{exp})}$ if there is a constant $\mu_x = \hat{\mu} = 0.058$, meaning that $q_x^{(\mathrm{exp})} = 0.057$, and $d_x^{(\mathrm{Gom})} = l_x \cdot q_x^{(\mathrm{Gom})}$ if

$$q_x = q_x^{(\mathrm{Gom})} := 1 - \exp\left\{ -\frac{\hat{B}}{\hat{\theta}} e^{\hat{\theta} x} \left(e^{\hat{\theta}} - 1 \right) \right\},$$

which is obtained by integrating the Gompertz hazard.

It matters little how we choose to interpret the deviations in the column $z_x^{(\mathrm{exp})}$ — with values going up as high as 6.65, it is clear that these could not have come from a normal distribution, and we must reject the null hypothesis that these lifetimes came from an exponential distribution.

As for the Gompertz model, the deviations are all quite moderate. We compute $\sum z_x^2 = 26.1$. There are 29 categories, but we have estimated 2 parameters, so this needs to be compared to the χ^2 distribution with 27 degrees of freedom. The cutoff for a test at the 0.05 level is 40.1, so we do not reject the null hypothesis.

As you have already learned, the χ^2 approximation doesn't work very well when the expected numbers in some categories are too low. This is certainly the case in this case, with d_x^{Gom} as low as 0.7. (That is, we are using a normal approximation with mean 0.7 for a quantity which takes on integer values. That obviously cannot be right.) The solution is to lump categories together. If we replace the first 10 years by a single category, it will have an expected number of deaths equal to $0.171 \cdot 103 = 17.7$, as compared with exactly 17 observed deaths, producing a z value of 0.18. Similarly, we cut off the final three years (since collectively they correspond to the certain event that all remaining individuals die), leaving us with $\sum z_x^2 = 11$ on 14 degrees of freedom. Again, this is a perfectly ordinary value of the χ^2 variable, and we do not reject the hypothesis of Gompertz mortality.

age	l_x	d_x	$q_x^{(\mathrm{exp})}$	$d_x^{(\mathrm{exp})}$	$z_x^{(\mathrm{exp})}$	$q_x^{(\mathrm{Gom})}$	$d_x^{(\mathrm{Gom})}$	$z_x^{(\mathrm{Gom})}$
0	103	0	0.057	5.8	-2.48	0.007	0.7	-0.97
1	103	0	0.057	5.8	-2.48	0.009	0.9	-0.97
2	103	3	0.057	5.8	-1.20	0.011	1.1	1.81
3	100	1	0.057	5.7	-2.02	0.013	1.3	-0.25
4	99	1	0.057	5.6	-2.00	0.015	1.5	-0.41
5	98	3	0.057	5.5	-1.11	0.018	1.8	0.94
6	95	2	0.057	5.4	-1.50	0.021	2.0	-0.02
7	93	1	0.057	5.3	-1.91	0.025	2.4	-0.90
8	92	2	0.057	5.2	-1.45	0.030	2.8	-0.47
9	90	4	0.057	5.1	-0.50	0.036	3.2	0.45
10	86	4	0.057	4.9	-0.40	0.042	3.6	0.19
11	82	3	0.057	4.6	-0.78	0.050	4.1	-0.56
12	79	4	0.057	4.5	-0.23	0.059	4.7	-0.32
13	75	3	0.057	4.2	-0.62	0.070	5.3	-1.02
14	72	8	0.057	4.1	2.00	0.083	6.0	0.87
15	64	4	0.057	3.6	0.21	0.098	6.2	-0.95
16	60	4	0.057	3.4	0.34	0.115	6.9	-1.17
17	56	7	0.057	3.2	2.22	0.135	7.6	-0.22
18	49	10	0.057	2.8	4.47	0.159	7.8	0.87
19	39	6	0.057	2.2	2.63	0.186	7.2	-0.51
20	33	3	0.057	1.9	0.85	0.216	7.1	-1.75
21	30	10	0.057	1.7	6.56	0.252	7.6	1.03
22	20	8	0.057	1.1	6.65	0.292	5.8	1.07
23	12	4	0.057	0.7	4.15	0.336	4.0	-0.02
24	8	3	0.057	0.5	3.90	0.386	3.1	-0.06
25	5	0	0.057	0.3	-0.55	0.440	2.2	-1.98
26	5	3	0.057	0.3	5.26	0.498	2.5	0.46
27	2	0	0.057	0.1	-0.35	0.559	1.1	-1.59
28	2	2	0.057	0.1	5.78	0.623	1.2	1.10

Table 6.2: Life table for tyrannosaurs, with fit to exponential and Gompertz models, and based on data from Table 1.1.