Actuarial Models: solutions example sheet 9

Answer to 1
Using formula (9.1) in the lecture notes, we have
\[
\ell(\lambda, \alpha) = C + \sum_{i=1}^{n} \left\{ \delta_i \log \mu(\tilde{t}_i) - A(\tilde{t}_i) \right\},
\]
with \( A(t) = \int_{0}^{t} \mu(s)ds = (\lambda t)^{\alpha} \) the cumulative hazard function and \( C \in \mathbb{R} \) an unimportant constant. Hence
\[
\ell(\lambda, \alpha) = C + \sum_{i=1}^{n} \left\{ \delta_i \left[ \log \alpha + \alpha \log \lambda + (\alpha - 1) \log \tilde{t}_i \right] - \lambda^{\alpha} \tilde{t}_i^{\alpha} \right\}.
\]

Answer to 2
Assume that the failure time corresponding to the treatment group is exponentially distributed with parameter \( \mu_1 \) and the failure time corresponding to the control group is exponentially distributed with parameter \( \mu_2 \). Denote by \( \delta^{(1)} \) and \( \delta^{(2)} \) the total number of deaths in the treatment group respectively control group and by \( \tilde{t}^{(1)} \) and \( \tilde{t}^{(2)} \) the total exposure time of the treatment group respectively the control group. The likelihood of \( \mu_1 \) given the observations of the treatment group is given by
\[
L^{(1)}(\mu_1) = \mu_1^{\delta^{(1)}} e^{-\mu_1 \tilde{t}^{(1)}} = \mu_1^{9} e^{-359\mu_1}
\]
and thus the log-likelihood is equal to
\[
\ell^{(1)}(\mu_1) = \log L^{(1)}(\mu_1) = 9 \log \mu_1 - 359\mu_1.
\]
It is easy to find the unique maximiser of \( \ell^{(1)}(\mu) \) and thus the maximum likelihood estimator of \( \mu_1 \), and it is given by \( \hat{\mu}_1 = 9/359 = 0.0251 \). Similarly, for the control group the log-likelihood of \( \mu_2 \) is given by
\[
\ell^{(2)}(\mu_2) = 21 \log \mu_2 - 182\mu_2
\]
and consequently the mle of \( \mu_2 \) is given by \( \hat{\mu}_2 = 21/182 = 3/26 = 0.115 \).

We now want to test the hypothesis
\[
H_0 : \mu_1 = \mu_2 \quad \text{versus} \quad H_A : \mu_1 \neq \mu_2.
\]
We do this by using the generalised likelihood ratio test. Let \( \ell(\mu_1, \mu_2) \) be the log-likelihood of the parameters \( \mu_1 \) and \( \mu_2 \) given the observations of both groups. By independence of the individuals, we have
\[
\ell(\mu_1, \mu_2) = \ell^{(1)}(\mu_1) + \ell^{(2)}(\mu_2).
\]
The test statistic for the generalised likelihood ratio test is given by
\[
2(\max_{\mu_1, \mu_2} \ell(\mu_1, \mu_2) - \max_{\mu} \ell(\mu, \mu)).
\]

We easily see that
\[
\max_{\mu_1, \mu_2} \ell(\mu_1, \mu_2) = \ell_1(\hat{\mu}_1) + \ell_2(\hat{\mu}_2) = -108.52.
\]

Further
\[
\ell(\mu, \mu) = 9 \log \mu - 359\mu + 21 \log \mu - 182\mu = 30 \log \mu - 541\mu.
\]

and thus
\[
\max_{\mu} \ell(\mu, \mu) = \ell(30/541, 30/541) = -116.77.
\]

Hence the test statistic is equal to
\[
2(\max_{\mu_1, \mu_2} \ell(\mu_1, \mu_2) - \max_{\mu} \ell(\mu, \mu)) = 16.49.
\]

Under \(H_0\) this value is a sample from a chi-squared distribution with 1 degree of freedom. Therefore we reject \(H_0\) if this sample is unlikely to come from a chi-squared distribution with 1 degree of freedom. To be more precise, we reject \(H_0\) at significance level 0.05 if 16.49 > \(\lambda\) where \(\lambda > 0\) is defined such that \(\Pr(\chi^2_1 \geq \lambda) = 0.05\). From statistical tables we find that \(\lambda = 3.841\) and so we reject \(H_0\), which means that there is evidence (at the 5% level of significance) that there is a difference between the parameter values of the two exponential distributions.

**Answer to 3**

(a) Similarly, as for the previous exercise, the likelihood of \(\mu\) given the data is equal to
\[
L(\mu) = \mu^4 e^{-36\mu}
\]

and the mle of \(\mu\) is given by \(\hat{\mu} = 4/36 = 1/9 = 0.11\). Further, the observed Fisher information is given by
\[
\hat{I} = -\frac{d^2}{d\mu^2} \log L(\mu)|_{\mu=\hat{\mu}} = \frac{4}{\hat{\mu}^2} = 324.
\]

From maximum likelihood theory we know that approximately \(\hat{\mu} \sim \mathcal{N}(\mu, 1/324)\). Hence an approximate 95% confidence interval for \(\mu\) using the mle \(\hat{\mu}\) is given by
\[
\hat{\mu} \pm 1.96 \sqrt{\frac{1}{324}} = \frac{1}{9} \pm 1.96 \frac{1}{18} = [0.0022, 0.22].
\]

(b) From the notes we have that the score function/statistic
\[
U(\mu) = \frac{d}{d\mu} \log L(\mu) = \frac{4}{\mu} - 36.
\]
is approximately $\mathcal{N}(0, \tilde{I}(\mu))$ distributed with
\[
\tilde{I}(\mu) := -\frac{d^2}{d\mu^2} \log L(\mu) = \frac{4}{\mu^2},
\]
i.e. $U(\mu) \simeq \mathcal{N}(0, 4/\mu^2)$ or $\frac{1}{2}\mu U(\mu) \simeq \mathcal{N}(0, 1)$. Since
\[
\frac{1}{2}\mu U(\mu) = 2 - 18\mu.
\]
This means that approximately 95% of the times
\[-1.96 \leq 2 - 18\mu \leq 1.96
\]
or
\[0.0022 \leq \mu \leq 0.22.
\]

**Answer to 4**

(a) (i) We first determine from the data the values for $t_i$ (the $i$th ordered observed genuine failure time), $d_i$ (the observed number of failures/divorces at $t_i$) and $r_i$ (the observed number of marriages at risk of divorce just before $t_i$). These values are given in Table 1. Then we compute the Kaplan-Meier estimator $\hat{S}(t)$ at the failure times $t_i$ using the recursion,
\[
\hat{S}(t_i) = (1 - \frac{d_i}{r_i}) \hat{S}(t_{i-1}), \quad i = 2, 3, \ldots.
\]
The results are given in the last column of Table 1. Since $\hat{S}(t)$ is right-continuous and piecewise constant in between the genuine failure times $t_i$, we have that the value of $\hat{S}(t)$ for $t \in [t_i, t_{i+1})$ is given by $\hat{S}(t) = \hat{S}(t_i)$. Further, $\hat{S}(t) = 0$ for $t \in [0, t_1) = [0, 1.5)$.

<table>
<thead>
<tr>
<th>$t_i$</th>
<th>$d_i$</th>
<th>$r_i$</th>
<th>$\hat{S}(t_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>1</td>
<td>20</td>
<td>0.950</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>19</td>
<td>0.800</td>
</tr>
<tr>
<td>3.5</td>
<td>2</td>
<td>15</td>
<td>0.693</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>13</td>
<td>0.587</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>10</td>
<td>0.528</td>
</tr>
<tr>
<td>6.5</td>
<td>1</td>
<td>8</td>
<td>0.462</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>6</td>
<td>0.385</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>6</td>
<td>0.289</td>
</tr>
</tbody>
</table>

Table 1: Table corresponding to Exercise 4(a)(i).

(ii) We need to estimate $1 - S(5)$. We have $\hat{S}(5) = \hat{S}(4) = \frac{44}{75} = 0.587$. Hence the estimate of the probability that a couple divorces before they reach their 5-year anniversary and based on the Kaplan-Meier estimator, is given by $\frac{31}{75} = 0.413$. 
(b) (i) The hazard function $\mu(t)$ and cumulative hazard function $A(t) = \int_0^t \mu(s)ds$ are given by

$$
\mu(t) = \begin{cases} 
\theta_1 & \text{if } 0 < t \leq 5, \\
\theta_2 & \text{if } 5 < t \leq 10, \\
\theta_3 & \text{if } t > 10.
\end{cases}
$$

$$
A(t) = \begin{cases} 
\theta_1 t & \text{if } 0 < t \leq 5, \\
5\theta_1 + (t - 5)\theta_2 & \text{if } 5 < t \leq 10, \\
5\theta_1 + 5\theta_2 + (t - 10)\theta_3 & \text{if } t > 10.
\end{cases}
$$

Using formula (9.1) in the lecture notes, we have

$$
\ell(\theta_1, \theta_2, \theta_3) = C + \sum_{i=1}^{20} \{ \delta_i \log \mu(\tilde{t}_i) - A(\tilde{t}_i) \}
$$

where $\tilde{t}_1, \ldots, \tilde{t}_{20}$ are the 20 numbers, from left to right, given in Table 2 on the Exercise sheet and $\delta_i = 0$ if $\tilde{t}_i$ is a censored value (recall that the censored values in Table 2 on the Exercise sheet are indicated by a *) and $\delta_i = 1$ otherwise. Further $C$ is an unimportant constant. Inserting the data and the parametric forms of $\mu(t)$ and $A(t)$, we get

$$
\ell(\theta_1, \theta_2, \theta_3) = C + \sum_{i=1}^{10} \{ \delta_i \log \mu(\tilde{t}_i) - A(\tilde{t}_i) \} + \sum_{i=11}^{15} \{ \delta_i \log \mu(\tilde{t}_i) - A(\tilde{t}_i) \}
$$

$$
\quad \quad \quad + \sum_{i=15}^{20} \{ \delta_i \log \mu(\tilde{t}_i) - A(\tilde{t}_i) \}
$$

$$
= C + 8 \log \theta_1 - 29.5\theta_1 + 3 \log \theta_2 - 25\theta_2 - 9.5\theta_2 + \log \theta_3 - 25\theta_1 - 25\theta_2 - 22\theta_3
$$

$$
= C + 8 \log \theta_1 - 79.5\theta_1 + 3 \log \theta_2 - 34.5\theta_2 + \log \theta_3 - 22\theta_3.
$$

It is easy to see that in order to maximise $\ell$ with respect to $\theta_1$, we need to maximise $g(\theta_1) = 8 \log \theta_1 - 79.5\theta_1$. Since $g''(\theta_1) = -\frac{8}{\theta_1^2} < 0$ and $g''(8/79.5) = 0$, it follows that the mle of $\theta_1$ is given by $\hat{\theta}_1 = \frac{8}{79.5} = 0.101$. Similarly, the mles of $\theta_1$ and $\theta_2$ are given by respectively $\hat{\theta}_2 = \frac{3}{34.5} = 0.087$ and $\hat{\theta}_3 = \frac{1}{22} = 0.045$.

(ii) With $T$ the survival time representing time from marriage until divorce, we need to estimate

$$
\Pr(T > 9|T > 4) = \frac{S(9)}{S(4)} = e^{A(9) - A(4)} = e^{-\theta_1 - 4\theta_2},
$$

cf. (1.3) and (1.4) in the notes. Hence we can estimate this probability by

$$
e^{-\hat{\theta}_1 - 4\hat{\theta}_2} = 0.639.
$$

(iii) An exponential distribution would correspond to $\theta_1 = \theta_2 = \theta_3$. Hence we want to test

$$
H_0 : \theta_1 = \theta_2 = \theta_3 \quad \text{versus} \quad H_A : \theta_1 \neq \theta_2 \text{ or } \theta_2 \neq \theta_3.
$$
We use the likelihood ratio test for this with the test statistic
\[ \Lambda = 2 \left( \max_{\theta_1, \theta_2, \theta_3} \ell(\theta_1, \theta_2, \theta_3) - \max_{\hat{\theta}} \ell(\hat{\theta}, \hat{\theta}, \hat{\theta}) \right). \]

Clearly, by the definition of an mle,
\[ \max_{\theta_1, \theta_2, \theta_3} \ell(\theta_1, \theta_2, \theta_3) = \ell(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) = C - 40.789. \]

Further,
\[ \max_{\theta} \ell(\theta, \theta, \theta) = C + \max_{\theta} (12 \log \theta - 136) = C + \frac{12}{136} - \frac{12 \times 12}{136} = C - 41.133. \]

Hence \( \Lambda = 2 \times (41.133 - 40.789) = 0.689. \) Under \( H_0, \) \( \Lambda \) should be a sample from a \( \chi^2 \)-distribution with \( 3 - 1 = 2 \) degrees of freedom. Since \( \Pr(\chi^2 > 5.991) = 0.05 \) and \( \Lambda < 5.991, \) we do not reject \( H_0 \) at the 5% significance level. This means that at the 5% significance level, the piecewise constant hazard function that we considered, does not perform better than the constant (over the whole positive half line) hazard function (which recall corresponds to the exponential distribution).

(c) The expected duration of a marriage is given by \( \mathbb{E}[T] = \int_0^\infty S(t)dt \) and can thus be estimated by \( \int_0^\infty \hat{S}(t)dt \), where \( \hat{S}(t) \) is an estimator for \( S(t) \). Now the non-parametric estimator, i.e. the Kaplan-Meier estimator, does not give a reliable estimate of the survival function for values \( t > 14 \), leading to an unreliable estimate for \( \mathbb{E}[T] \). This is because there are no data points in the region \( t > 14 \) and since with non-parametric estimation no specific form of the survival function is assumed, one cannot really extrapolate the estimator to the region \( t > 14 \). Note that strictly speaking, the Kaplan-Meier estimator equals 0.321 for all \( t \geq 14 \). This means that the Kaplan-Meier estimator says that a divorce cannot take place more than 14 years after marriage, which is a rather strong conclusion to make and in particular leads to an estimated expected duration of \( +\infty \). On the other hand the parametric estimation procedure leads to a more natural/smooth progress of the survival function for large \( t \) and so can provide a more sensible estimate of the expected duration. In conclusion: the above parametric estimator is better than the non-parametric one (i.e. the Kaplan-Meier estimator), because it allows more naturally for extrapolation of the survival function beyond the largest observed value and this is necessary to estimate \( \mathbb{E}[T] = \int_0^\infty S(t)dt \).

Note that the estimate of \( \mathbb{E}[T] \) based on the above parametric estimation procedure is equal to
\[
\int_0^\infty \hat{S}(t)dt = \int_0^5 e^{-\hat{\theta}_1 t}dt + \int_5^{10} e^{-5\hat{\theta}_1 - \hat{\theta}_2(t-5)}dt + \int_{10}^\infty e^{-5(\hat{\theta}_1 + \hat{\theta}_2) - \hat{\theta}_3(t-10)}dt \\
= \frac{1}{\hat{\theta}_1} \left( 1 - e^{-5\hat{\theta}_1} \right) + \frac{1}{\hat{\theta}_2} \left( e^{-5\hat{\theta}_1} - e^{-5(\hat{\theta}_1 + \hat{\theta}_2)} \right) + \frac{1}{\hat{\theta}_3} e^{-5(\hat{\theta}_1 + \hat{\theta}_2)} \\
= 3.929 + 2.452 + 8.612 \\
= 14.99.
\]
Answer to 5

(a) (i) Obvious, since in this case \( S(t) = e^{-\lambda t} \).

(ii) We have
\[ S(t) = e^{-\int_0^t \mu(s)\,ds} = e^{-\int_0^t \alpha \lambda^n s^{\alpha-1} \,ds} = e^{-(\lambda t)^\alpha} \]
and thus
\[ \log(- \log S(t)) = \log((\lambda t)^\alpha) = \alpha \log \lambda + \alpha \log t. \]

(iii) With \( S(t) = \frac{1}{1 + (\lambda t)^\alpha} \) it follows that \( 1 - S(t) = \frac{1}{1 + (\lambda t)^\alpha} \) and so
\[ -\log \frac{S(t)}{1 - S(t)} = -\log \left( \frac{1}{(\lambda t)^\alpha} \right) = \alpha \log \lambda + \alpha \log t. \]

(iv) We have
\[ S(t) = e^{-\int_0^t \mu(s)\,ds} = e^{-\int_0^t \beta e^{\gamma s} \,ds} = e^{-\frac{\beta}{\gamma} (e^{\gamma t} - 1)} \]
and thus
\[ \log \left( -\log \left( \frac{S(t+1)}{S(t)} \right) \right) = \log \{\log(S(t)) - \log(S(t+1))\} \]
\[ = \log \left\{ \frac{\beta}{\gamma} (e^{\gamma(t+1)} - 1) - \frac{\beta}{\gamma} (e^{\gamma t} - 1) \right\} \]
\[ = \log \left( \frac{\beta e^{\gamma t} (e^\gamma - 1)}{\gamma} \right) \]
\[ = \log \left( \frac{\beta (e^{\gamma t} - 1)}{\gamma} \right) + \gamma t. \]

(b) What is common about the four results in (i)-(iv) is that in each case some function \( g \) of \( S(t) \) is a linear function of either \( t \) or \( \log t \). We can exploit this in the following way. Suppose that we have available the failure times of \( n \) individuals, some of whom may be censored. Let \( t_1, t_2, \ldots, t_k \) be the ordered genuine failure times and let \( \hat{S}(t) \) be the Kaplan-Meier estimator based on these results.

(i) If indeed the failure times are exponentially distributed with parameter \( \lambda \), then \( \hat{S}(t) \) will be very close to \( S(t) = e^{-\lambda t} \) at all \( t \) and in particular at \( t = t_1, t_2, \ldots, t_k \).
In this case, if we were to plot \(-\log(\hat{S}(t_i))\) against \( t_i, i = 1, 2, \ldots, k \) we should obtain a scatter plot which is a near straight line passing through the origin. The converse of this argument indicates that if on plotting \(-\log(\hat{S}(t_i))\) versus \( t_i, i = 1, 2, \ldots, k \), we obtain a near straight line passing through the origin, then we should take that as graphical evidence that the failure time distribution is exponential. Furthermore, if we fit a straight line to this plot, then its slope forms a rough estimate for \( \lambda \).

(ii) Following similar arguments, we see that if on plotting \( \log(-\log(\hat{S}(t_i))) \) versus \( \log(t_i), i = 1, 2, \ldots, k \), we get a near straight line scatter plot, then we could take
that as graphical evidence that the failure time distribution is Weibull and on fitting this straight line, we have the rough estimates
\[ \hat{\alpha} = \text{slope}, \quad \hat{\lambda} = \exp\left(\frac{\text{intercept}}{\text{slope}}\right). \]

(iii) Again following similar arguments, we see that if on plotting
\[ -\log\left(\frac{\hat{S}(t_i)}{1 - \hat{S}(t_i)}\right) \quad \text{against} \quad \log t_i, \quad i = 1, 2, \ldots, k, \]
we get a near straight line scatter plot, then we could take that as graphical evidence that the lifetime distribution is log-logistic and on fitting this straight line, we have the rough estimates
\[ \hat{\alpha} = \text{slope}, \quad \hat{\lambda} = \exp\left(\frac{\text{intercept}}{\text{slope}}\right). \]

(iv) Finally, we see by similar arguments as before that if on plotting
\[ \log\left(-\log\left(\frac{\hat{S}(t + 1)}{\hat{S}(t)}\right)\right) \quad \text{against} \quad \log(t_i), \quad i = 1, 2, \ldots, k, \]
we get a near straight line scatter plot, then we could take this as graphical evidence that the failure distribution is the Gompertz distribution. Clearly, however, this is only going to work if the sample size is sufficiently large to have consecutive ordered failure times less than one unit of time apart. If we do get a near straight line plot, then on fitting the straight line, we can construct the following rough estimates of the Gompertz parameters:
\[ \hat{\gamma} = \text{slope}, \quad \hat{\beta} = \frac{\text{slope} \times e^{\text{intercept}}}{e^{\text{slope}} - 1}. \]