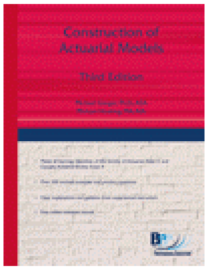

Construction of Actuarial Models

Third Edition

by Mike Gauger and Michael Hosking

Published by BPP Professional Education

Solutions to practice questions – Chapter 8

Solution 8.1

From the given baseline hazard function, we have:

$$H_0(x) = \int_0^x \frac{2}{10+s}\,ds = 2\ln(10+s)\Big|_0^x = 2\ln\left(\frac{10+x}{10}\right) = \ln\left(\left(\frac{10+x}{10}\right)^2\right)$$

$$\Rightarrow \quad S_0(x) = e^{-H_0(x)} = \left(\frac{10}{10+x}\right)^2$$

$$\Rightarrow \quad S_1(x) = \frac{S_0(2+x)}{S_0(2)} = \left(\frac{10}{10+(x+2)}\right)^2 \Big/ \left(\frac{10}{10+2}\right)^2 = \left(\frac{12}{12+x}\right)^2$$

Solution 8.2

$$h_0(x) = 1.2/(100-x) \;\Rightarrow\; H_0(x) = 1.2\ln\left(\frac{100}{100-x}\right)$$

$$\Rightarrow S_0(x) = e^{-H_0(x)} = \left(\frac{100}{100-x}\right)^{-1.2} = \left(\frac{100-x}{100}\right)^{1.2} \quad \text{for } 0 \le x < 100$$

$$\Rightarrow S_1(x) = S(1.05x) = \left(\frac{100-1.05x}{100}\right)^{1.2} = \left(\frac{95.23810-x}{95.23810}\right)^{1.2} \quad \text{for } 0 \le x < 95.23810$$

Solution 8.3

In general, if hazard functions are related by $h_1(x) = k\,h_0(x)$, then survival functions are related by $S_1(x) = \left(S_0(x)\right)^k$. So here we have:

$$S_1(x) = \sqrt{\frac{10}{10+x}} \quad \text{for } x > 0$$

Solution 8.4

The hazard function here is the one corresponding to the baseline survival function given in Question 8.3:

$$h_1(x) = 0.5h_0(x) = \frac{0.5}{10+x} \Rightarrow S_1(x) = \sqrt{\frac{10}{10+x}} \Rightarrow f_1(x) = h_1(x)S_1(x) = \frac{0.5\sqrt{10}}{(10+x)^{1.5}}$$

Solution 8.5

$$H_0(x) = \int_0^x 0.002t\,dt = 0.001x^2 \text{ for } x > 0 \Rightarrow S_0(x) = e^{-0.001x^2} \text{ for } x > 0$$

$$\Rightarrow {}_{10}p_{20} = \frac{S_0(30)}{S_0(20)} = \frac{e^{-0.001\times 30^2}}{e^{-0.001\times 20^2}} = e^{-0.5} = 0.60653$$

$$h_1(x) = 0.95h_0(x) \Rightarrow {}_{10}p_{20} = 0.60653^{0.95} = 0.62189$$

Solution 8.6

$$S_i(x) = S_j(x)^{K_{ij}} \Rightarrow 0.82 = S_i(60) = S_j(60)^{K_{ij}} = 0.80^{K_{ij}} \Rightarrow K_{ij} = \frac{\ln(0.82)}{\ln(0.80)} = 0.889$$

Solution 8.7

$$H_0(x) = \int_0^x h_0(t)\,dt = \int_0^x 0.015 + 0.01\times 1.01^t\,dt = 0.015x + \frac{0.01\left(1.01^x - 1\right)}{\ln(1.01)} \Rightarrow$$

$${}_{10}p_{60}^{\text{fn}} = \frac{S_0(70)}{S_0(60)} = \frac{\exp(-H_0(70))}{\exp(-H_0(60))} = \exp\left(-\left(H_0(70) - H_0(60)\right)\right)$$

$$= \exp\left(-\left(0.015(70-60) + \frac{0.01\left(1.01^{70} - 1.01^{60}\right)}{\ln(1.01)}\right)\right) = 0.71105$$

Solution 8.8

A female non-smoker ($z_1 = z_2 = 0$) corresponds to the baseline group. For a male 2-pack per day smoker ($z_1 = 1, z_2 = 2$) we have:

$$g(1,2) = e^{0.010\times 1 + 0.015\times 2} = e^{0.04} = 1.04081$$

So the relative risk is:

$$K = \frac{g(1,2)}{g(0,0)} = g(1,2) = 1.04081$$

Solution 8.9

Using the results of Solutions 8.8 and 8.9, we have:

$$_{10}p_{60}^{\text{ms}} = \left(_{10}p_{60}^{\text{fn}} \right)^K = 0.71105^{1.04081} = 0.70122$$

Solution 8.10

From Section 8.4, we have:

$$\frac{\partial \ln\left(L(K,\theta)\right)}{\partial K} = \frac{t}{K} - \frac{\sum x_{m,j}}{\theta}$$
$$\frac{\partial \ln\left(L(K,\theta)\right)}{\partial \theta} = \frac{\sum x_{f,i} + K\sum x_{m,j}}{\theta^2} - \frac{s+t}{\theta}$$

The second order partial derivatives are thus:

$$\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K^2} = -\frac{t}{K^2}$$
$$\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K \partial \theta} = \frac{\sum x_{m,j}}{\theta^2}$$
$$\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K^2} = -2 \times \frac{\sum x_{f,i} + K\sum x_{m,j}}{\theta^3} + \frac{s+t}{\theta^2}$$

Using the fact that $E\left[X_f\right] = \theta$ and $E\left[X_m\right] = \theta / K$ (see Section 8.4), we have:

$$-E\left[\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K^2}\right] = \frac{t}{K^2}$$
$$-E\left[\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K \partial \theta}\right] = -E\left[\frac{\sum X_{m,j}}{\theta^2}\right] = \frac{-t\theta/K}{\theta^2} = -\frac{t}{\theta K}$$
$$-E\left[\frac{\partial^2 \ln\left(L(K,\theta)\right)}{\partial K^2}\right] = 2 \times E\left[\frac{\sum X_{f,i} + K\sum X_{m,j}}{\theta^3}\right] - \frac{s+t}{\theta^2}$$
$$= 2 \times \frac{s\theta + tK(\theta/K)}{\theta^3} - \frac{s+t}{\theta^2} = \frac{s+t}{\theta^2}$$

So the information matrix is:

$$I(K,\theta) = \begin{pmatrix} \frac{t}{K^2} & -\frac{t}{\theta K} \\ -\frac{t}{\theta K} & \frac{s+t}{\theta^2} \end{pmatrix}$$

Solution 8.11

The asymptotic covariance matrix of the estimators is the inverse of the information matrix:

$$\begin{pmatrix} \text{var}(\hat{K}) & \text{cov}(\hat{K},\hat{\theta}) \\ \text{cov}(\hat{K},\hat{\theta}) & \text{var}(\hat{\theta}) \end{pmatrix} = I(K,\theta)^{-1} = \begin{pmatrix} \frac{t}{K^2} & -\frac{t}{\theta K} \\ -\frac{t}{\theta K} & \frac{s+t}{\theta^2} \end{pmatrix}^{-1}$$

$$= \frac{1}{\det(I(K,\theta))} \begin{pmatrix} \frac{s+t}{\theta^2} & \frac{t}{\theta K} \\ \frac{t}{\theta K} & \frac{t}{K^2} \end{pmatrix} = \frac{K^2\theta^2}{st} \begin{pmatrix} \frac{s+t}{\theta^2} & \frac{t}{\theta K} \\ \frac{t}{\theta K} & \frac{t}{K^2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{K^2(s+t)}{st} & \frac{K\theta}{s} \\ \frac{K\theta}{s} & \frac{\theta^2}{s} \end{pmatrix}$$

Solution 8.12

First determine the baseline (female) model:

$$h_0 = \frac{\alpha}{x} \quad \text{for } x \ge 1$$

$$H_0(x) = \int_1^x \frac{\alpha}{t} dt = \alpha \ln(x) \quad \text{for } x \ge 1$$

$$S_0(x) = e^{-H_0(x)} = e^{-\alpha \ln(x)} = \frac{1}{x^\alpha} \quad \text{for } x \ge 1$$

$$f_0(x) = h_0(x) S_0(x) = \frac{\alpha}{x^{\alpha+1}} \quad \text{for } x \ge 1$$

For the male survival model we will simplify the notation. Instead of the standard notation $h(x \mid z=1) = g(1)h_0(x)$ we will simply write $h_1(x)$. Furthermore, it will simplify the likelihood function if we write $K = g(1) = e^\beta$:

$$h_1(x) = \frac{\alpha K}{x} \quad \text{for } x \ge 1$$

$$H_1(x) = \int_0^x h_1(t) dt = \alpha K \ln(x) \quad \text{for } x \ge 1$$

$$S_1(x) = e^{-H_1(x)} = \frac{1}{x^{\alpha K}} \quad \text{for } x \ge 1$$

$$f_1(x) = h_1(x) S_1(x) = \frac{\alpha K}{x^{\alpha K+1}} \quad \text{for } x \ge 1$$

Now we can write down the likelihood function:

$$\begin{aligned} L(K,\alpha) &= \prod_{i=1}^{s} f_0\left(x_{f,i}\right) \times \prod_{j=1}^{t} f_1\left(x_{m,j}\right) \\ &= \prod_{i=1}^{s} \frac{\alpha}{\left(x_{f,i}\right)^{\alpha+1}} \times \prod_{j=1}^{t} \frac{\alpha K}{\left(x_{m,j}\right)^{\alpha K+1}} \\ &= \alpha^{s+t} K^t \times \frac{1}{\prod_{i=1}^{s} \left(x_{f,i}\right)^{\alpha+1}} \times \frac{1}{\prod_{j=1}^{t} \left(x_{m,j}\right)^{\alpha K+1}} \end{aligned}$$

Solution 8.13

The log-likelihood function is:

$$\ln\left(L(K,\alpha)\right) = (s+t)\ln(\alpha) + t\ln(K) - (\alpha+1)\sum \ln\left(x_{f,i}\right) - (\alpha K+1)\sum \ln\left(x_{m,j}\right)$$

The partial derivatives are:

$$\begin{aligned} \frac{\partial \ln\left(L(K,\alpha)\right)}{\partial K} &= \frac{t}{K} - \alpha \sum \ln\left(x_{m,j}\right) \\ \frac{\partial \ln\left(L(K,\alpha)\right)}{\partial \alpha} &= \frac{s+t}{\alpha} - \sum \ln\left(x_{f,i}\right) - K\sum \ln\left(x_{m,j}\right) \end{aligned}$$

The simultaneous solutions are:

$$\hat{\alpha} = \frac{s}{\sum \ln\left(x_{f,i}\right)}$$ (inverse of the sample mean of the log of female lifetimes)

$$\hat{K} = \frac{\sum \ln\left(x_{f,i}\right)/s}{\sum \ln\left(x_{m,j}\right)/t}$$ (ratio of sample means of log-lifetimes)

Solution 8.14

To simplify matters we will use the following notation:

female: $g(0) = e^{\beta \times 0} = 1$

male: $g(1) = e^{\beta \times 1} = e^{\beta} = K$

The modified risk set data is:

i	1	2	3	4
y_i	2	5	8	9
$s_i = 1$	m	f	f	m
R_i	$2m, 2f$	$1m, 2f$	$1m, 1f$	$1m$

Using the formula for the partial likelihood function and the notation introduced here, the partial likelihood function is:

$$L(K) = \prod_{i=1}^{4} \frac{K_i}{\sum_{j \text{ in } R_i} K_j}$$

$$= \underbrace{\frac{K}{2+2K}}_{\substack{R_1:\,2f,2m \\ \text{male death}}} \times \underbrace{\frac{1}{2+K}}_{\substack{R_2:\,2f,m \\ \text{female death}}} \times \underbrace{\frac{1}{1+K}}_{\substack{R_3:\,1f,1m \\ \text{female death}}} \times \underbrace{\frac{K}{K}}_{\substack{R_4:\,1m \\ \text{male death}}}$$

This expression can be simplified to the following:

$$L(K) = \frac{K}{2(1+K)^2(2+K)}$$

The log-partial likelihood function is:

$$\ln\big(L(K)\big) = \ln(K) - 2\ln(1+K) - \ln(2+K) - \ln(2)$$

It is maximized at the unique critical point:

$$0 = \frac{d\ln\big(L(K)\big)}{dK} = \frac{1}{K} - \frac{2}{1+K} - \frac{1}{2+K}$$

$$\Rightarrow\ 0 = (1+K)(2+K) - 2K(2+K) - K(1+K) \quad \text{(numerator)}$$

$$\Rightarrow\ 0 = 1 - K - K^2$$

$$\Rightarrow\ \hat{K} = \frac{1 \pm \sqrt{5}}{-2} = 0.618 \text{ or } -1.618 \text{ (impossible since } K > 0\text{)}$$

$$\Rightarrow \hat{\beta} = \ln\big(\hat{K}\big) = -0.481$$

Solution 8.15

We obtained the partial likelihood estimate $\hat{K} = 0.618$ in Solution 8.14. We will need the following:

i	1	2	3	4
y_i	2	5	8	9
$\dfrac{s_i}{\sum_{j \text{ in } R_i} K_j}$	$\dfrac{1}{2+2K}$	$\dfrac{1}{2+K}$	$\dfrac{1}{1+K}$	$\dfrac{1}{K}$
	= 0.309	0.382	0.618	1.618

Using the (unrounded) values in the last line in the display above leads to:

$$\hat{H}_0(x) = \sum_{y_i \le x} \frac{s_i}{\sum_{j \text{ in } R_i} K_j} = \begin{cases} 0 & 0 \le x < 2 \\ 0.309 & 2 \le x < 5 \\ 0.691 & 5 \le x < 8 \\ 1.309 & 8 \le x < 9 \\ 2.927 & 9 \le x \end{cases}$$

Solution 8.16

$$\hat{S}_0(7) = \exp\left(-\hat{H}_0(7)\right) = 0.501 \quad \text{(Solution 8.15)}$$

$$\hat{K} = 0.618 \text{ (Solution 8.14)}$$

$$\Rightarrow \quad \hat{S}_0(7)^{\hat{K}} = 0.652$$

Solution 8.17

$$\begin{aligned} \text{var}\left(\hat{\beta}_1 z_1 + \hat{\beta}_2 z_2\right) &= (z_1)^2 \text{var}\left(\hat{\beta}_1\right) + (z_2)^2 \text{var}\left(\hat{\beta}_2\right) + 2 z_1 z_2 \text{cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) \\ &= 0.00005(z_1)^2 + 0.00002(z_2)^2 + 0.00002 z_1 z_2 \end{aligned}$$

Solution 8.18

The distribution of $\hat{\beta}_1 z_1 + \hat{\beta}_2 z_2$ is approximately normal with mean $\beta_1 z_1 + \beta_2 z_2$ and variance equal to the result in Solution 8.17.

Solution 8.19

From Solution 8.18, it follows that the distribution of $-\hat{\beta}_1 + 6\hat{\beta}_2$ is approximately normal with mean $-\beta_1 + 6\beta_2$ and variance:

$$E\left[-\hat{\beta}_1 + 6\hat{\beta}_2\right] = -\beta_1 + 6\beta_2 = 0.06200$$

$$\text{var}\left(-\hat{\beta}_1 + 6\hat{\beta}_2\right) = 0.00005 \times (-1)^2 + 0.00002 \times (6)^2 + 0.00002 \times (-6) = 0.00065$$

The 90% confidence interval for $-\beta_1 + 6\beta_2$ is thus:

$$0.06200 \pm 1.645 \times \sqrt{0.00065}$$

This interval is: $(0.020, 0.104)$.

Solution 8.20

The 95% confidence interval for $-\beta_1 + 6\beta_2$ (working exactly as in the previous question) is:

$$0.06200 \pm 1.96\sqrt{0.00065} = (0.01203,\ 0.11197)$$

The relative risk of a female who consumes 8 fluid ounces of alcohol per day (Life 1) to a male who consumes 2 fluid ounces per day (Life 2) is:

$$K_{12} = \frac{g(0,8)}{g(1,2)} = \frac{e^{8\beta_2}}{e^{\beta_1 + 2\beta_2}} = e^{-\beta_1 + 6\beta_2}$$

So we can obtain the 95% confidence interval for the relative risk by applying the exponential function to the endpoints of the interval just obtained:

$$\left(e^{0.0120}\ ,\ e^{0.1120}\right) = (1.012\ ,\ 1.118)$$

