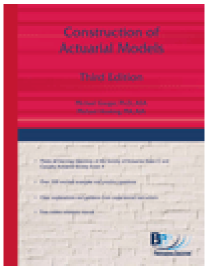


Construction of Actuarial Models

Third Edition

by Mike Gauger and Michael Hosking

Published by BPP Professional Education

Solutions to practice questions – Chapter 7

Solution 7.1

We have:

$$S(x)=\left(\frac{100-x}{100}\right)^{1.02}$$

The PDF is obtained as the negative derivative of the survival function:

$$f(x)=-S'(x)=-1.02\left(\frac{100-x}{100}\right)^{0.02}\left(-\frac{1}{100}\right)=\frac{1.02(100-x)^{0.02}}{100^{1.02}}$$

Solution 7.2

$$h(x)=2/(75+x)\ ,\ x\geq 0\quad\Rightarrow\quad H(x)=\int_0^x\frac{2}{75+y}dy=2\ln\left(\frac{75+x}{75}\right)\quad\Rightarrow$$

$$S(x)=\exp(-H(x))=\exp\left(-2\ln\left(\frac{75+x}{75}\right)\right)=\left(\frac{75}{75+x}\right)^2\quad\text{for }\ x>0$$

Solution 7.3

$$f(x)=-S'(x)=-\left(\frac{75^2}{(75+x)^2}\right)'=\frac{2\times 75^2}{(75+x)^3}\quad\text{for }\ x>0$$

Solution 7.4

$$h(1)=\frac{\Pr(X=1)}{\Pr(X\geq 1)}=\frac{0.4}{1.0}\ ,\quad h(2)=\frac{\Pr(X=2)}{\Pr(X\geq 2)}=\frac{0.6}{0.6}\quad,\quad\text{zero otherwise}$$

Solution 7.5

In general we have $H(x) = \sum_{x_i \le x} h(x)$. So we have:

$$H(1.2) = h(1) = 0.4 \ , \ H(2) = h(1) + h(2) = 1.4 \ , \ H(2.5) = h(1) + h(2) = 1.4$$

Solution 7.6

There are 6 of the 8 lives surviving at time 1.2. So we have $\hat{S}(1.2) = 6/8 = 0.75$.

The exact and approximate variances of this estimator are:

$$\text{var}\left(\hat{S}(1.2)\right) = \frac{S(1.2)\left(1-S(1.2)\right)}{8} \approx \frac{\hat{S}(1.2)\left(1-\hat{S}(1.2)\right)}{8} = \frac{0.75 \times 0.25}{8} = 0.02344$$

Solution 7.7

The empirical estimate of the cumulative hazard function is:

$$\hat{H}(1.2) = -\ln \underbrace{\left(\hat{S}(1.2)\right)}_{\text{Solution 7.6}} = -\ln(0.75) = 0.28768$$

The derivative of the transformation $y = -\ln(x)$ is $-1/x$. So according to the delta method, the approximate variance of the estimator of the empirical hazard function at age 1.2 is:

$$\text{var}\left(\hat{H}(1.2)\right) \approx \left(-\frac{1}{\hat{S}(1.2)}\right)^2 \text{var}\left(\hat{S}(1.2)\right) = \underbrace{\frac{1}{0.75^2} \times 0.02344}_{\text{Solution 7.6}} = 0.04167$$

Solution 7.8

Using the formula $f_n(x) = \dfrac{n_i / n}{x_i - x_{i-1}}$ for $x_{i-1} < x \le x_i$ and for $i = 1, 2, ..., k$, we obtain the following:

x **(hours after 9am)**	$f_{5000}(x)$
$0 \le x < 1$	0.044
$1 \le x < 2$	0.092
$2 \le x < 3$	0.136
$3 \le x < 4$	0.14
$4 \le x < 5$	0.13
$5 \le x < 7$	0.11
$7 \le x < 10$	0.0793

Solution 7.9

Using the results of Solution 7.8, we have for $1 \le x \le 2$:

$$F_{5,000}(x) = \int_0^x f_{5,000}(t)\,dt = \int_0^1 0.044\,dt + \int_1^x 0.092\,dt = 0.044 + 0.092(x-1)$$
$$= 0.092x - 0.048$$

Solution 7.10

We can linearly interpolate since the ogive is piecewise linear. The median time corresponds to the 2,500th ticket sold. You can see from the data that 2,060 tickets have been sold by $x = 4$ (1pm), and 2,710 tickets have been sold by $x = 5$ (2pm). Therefore the median time is approximated by:

$$0.50 = F_{5,000}(x_{0.5}) = \frac{2,060}{5,000} + \underbrace{\left(\frac{2,710-2,060}{5,000}\right)}_{\text{slope}}(x_{0.5} - 4)$$

$$\Rightarrow \quad x_{0.5} = 4.67692 \text{ hours}$$

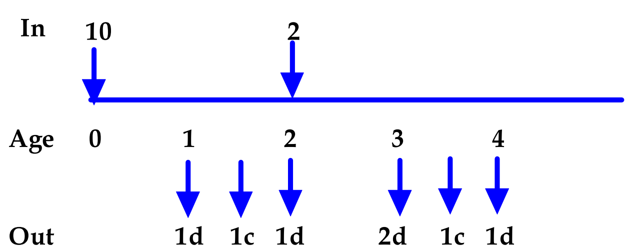
Solution 7.11

By 5pm one-third of the tickets sold between 4pm and 7pm will have been sold. So the probability that a ticket is sold by 5pm ($x = 8$):

$$F_{5,000}(8) = 1 - \underbrace{\frac{2}{3} \times \frac{1,190}{5,000}}_{\text{sold in } [8,10]} = 0.841$$

Solution 7.12

Here is the age diagram indicating all activity prior to time 4.1:



So the important data are:

i	1	2	3	4
y_i	1	2	3	4
r_i	10	8	9	6
s_i	1	1	2	1

Solution 7.13

$$\hat{S}(4.1) = \prod_{y_i \le 4.1} \frac{r_i - s_i}{r_i} = \prod_{i=1}^{4} \frac{r_i - s_i}{r_i} = \frac{9}{10} \times \frac{7}{8} \times \frac{7}{9} \times \frac{5}{6} = \frac{49}{96} = 0.51042$$

Solution 7.14

Using results from Solution 7.13:

$$\sigma_s^2(4.1) = \sum_{y_i \le 4.1} \frac{s_i}{r_i(r_i - s_i)} = \frac{1}{10 \times 9} + \frac{1}{8 \times 7} + \frac{2}{9 \times 7} + \frac{1}{6 \times 5} = 0.09405$$

$$\text{var}\left(\hat{S}(4.1)\right) \approx \left(\hat{S}(4.1)\right)^2 \sigma_s^2(4.1) = 0.02450$$

Solution 7.15

Using results from Solutions 7.13 and 7.14, we have:

linear $\hat{S}(4.1) \pm 1.645\sqrt{\text{var}\left(\hat{S}(4.1)\right)}$ is the interval $(0.253\ ,\ 0.728)$

transformed $U = \exp\left(\dfrac{z_{\alpha/2}\ \sigma_s(4.1)}{\ln\left(\hat{S}(4.1)\right)}\right) = \exp\left(\dfrac{1.645 \times 0.30667}{-0.67253}\right) = 0.47231$

$$\left(\hat{S}(4.1)^{1/U}\ ,\ \hat{S}(4.1)^{U}\right) = (0.241\ ,\ 0.729)$$

Solution 7.16

Using results from Solution 7.12:

$$\hat{H}(4.1) = \sum_{y_i \le 4.1} \frac{s_i}{r_i} = \sum_{i=1}^{4} \frac{s_i}{r_i} = \frac{1}{10} + \frac{1}{8} + \frac{2}{9} + \frac{1}{6} = 0.61389$$

$$\text{var}\left(\hat{H}(4.1)\right) \approx \sum_{y_i \le 4.1} \frac{s_i}{r_i^2} = \sum_{i=1}^{4} \frac{s_i}{r_i^2} = \frac{1}{10^2} + \frac{1}{8^2} + \frac{2}{9^2} + \frac{1}{6^2} = 0.07809$$

Solution 7.17

$$\theta = \exp\left(\frac{z_{\alpha/2}\sqrt{\text{var}\left(\hat{H}(4.1)\right)}}{\hat{H}(4.1)}\right) = \exp\left(\frac{1.645 \times 0.27945}{0.61389}\right) = 2.11453$$

$$\left(\frac{\hat{H}(4.1)}{\theta}\ ,\ \theta\hat{H}(4.1)\right) = (0.290\ ,\ 1.298)$$

Solution 7.18

Using the results of Solution 7.16, the Nelson-Aalen estimate is:

$$\hat{S}^{\text{NA}}(4.1) = \exp\left(-\hat{H}(4.1)\right) = \exp(-0.61389) = 0.54124$$

The derivative of the transformation $y = e^{-x}$ is $-e^{-x}$. So, by the delta method, the approximate variance of the underlying estimator is:

$$\text{var}\left(\hat{S}^{\text{NA}}(4.1)\right) \approx \left(-e^{-\hat{H}(4.1)}\right)^2 \text{var}\left(\hat{H}(4.1)\right) = \left(-e^{-0.61389}\right)^2 \times 0.07809 = 0.02288$$

Solution 7.19

The empirical distribution is: $f_4(1) = 0.25$, $f_4(1.8) = 0.5$, $f_4(4) = 0.25$. Using $b = 1$, we have:

$$k_1^t(x) = \frac{b - |1-x|}{b^2} = 1 - |1-x| \text{ for } 0 \le x \le 2 \Rightarrow k_1^t(1.2) = 0.8$$

$$k_{1.8}^t(x) = \frac{b - |1.8-x|}{b^2} = 1 - |1.8-x| \text{ for } 0.8 \le x \le 2.8 \Rightarrow k_{1.8}^t(1.2) = 0.4$$

$k_4^t(1.2) = 0$ (1.2 is more than a bandwidth below 4)

So the kernel smoothed approximation of $f(1.2)$ is:

$$\begin{aligned} f_4^{ks}(1.2) &= f_4(1)k_1^t(1.2) + f_4(1.8)k_{1.8}^t(1.2) + f_4(4)k_4^t(1.2) \\ &= 0.25 \times 0.8 + 0.50 \times 0.4 + 0.25 \times 0 = 0.40 \end{aligned}$$

Solution 7.20

For this question, we have:

$$F_4^{ks}(1.2) = f_4(1)K_1^t(1.2) + f_4(1.8)K_{1.8}^t(1.2) + f_4(4)K_4^t(1.2)$$

where:

$$K_1^t(1.2) = \underbrace{\int_0^1 k_1^t(x)\,dx}_{1/2 \text{ (symmetry)}} + \int_1^{1.2} k_1^t(x)\,dx = 0.5 + \int_1^{1.2} 1 - |1-x|\,dx$$

$$= 0.5 + \int_1^{1.2} 2 - x\,dx = 0.5 + 0.18 = 0.68$$

$$K_{1.8}^{t}(1.2) = \int_{0.8}^{1.2} k_{1.8}^{t}(x)\,dx = \int_{0.8}^{1.2} 1-|1.8-x|\,dx$$

$$= \int_{0.8}^{1.2} x-0.8\,dx = 0.4 - 0.32 = 0.08$$

$K_{4}^{t}(1.2) = 0$ since 1.2 is more than a bandwidth to the left of 4

Finally, we have:

$$F_4^{ks}(1.2) = f_4(1)K_1^t(1.2) + f_4(1.8)K_{1.8}^t(1.2) + f_4(4)K_4^t(1.2)$$
$$= 0.25 \times 0.68 + 0.50 \times 0.08 + 0.25 \times 0 = 0.21$$

Solution 7.21

In working this series of problems notice that all left truncation is at $x = 200$ which is the left endpoint of the second interval. All right censoring is at $x = 1000$ which is the right endpoint of the interval $(750, 1000]$. We have $\alpha = 1$ and $\beta = 0$ since new entrants arrive at the start of the interval and censors occur at the end of the interval. The starting value is $P_0 = 480$. The recursive rule is:

$P_{j+1} = P_j + d_j - u_j - x_j$
d_j = new entrants in $[c_j, c_{j+1})$
u_j = censors in $(c_j, c_{j+1}]$
x_j = number of losses in $(c_j, c_{j+1}]$

So:

i	Range	P_i	d_i	u_i	x_i
0	$(0, 200]$	480	0	0	30
1	$(200, 500]$	450	500	0	285
2	$(500, 750]$	665	0	0	220
3	$(750, 1000]$	445	0	40	230
4	$(1000, 1500]$	175	0	0	115
6	$(1500, 2000]$	60	0	0	60

Solution 7.22

Since $\alpha = 1$ and $\beta = 0$, the risk set r_j is determined from: $r_j = P_j + \alpha d_j - \beta u_j = P_j + d_j$. So using the table in Solution 7.21 together with this formula, we have:

$r_0 = 480$, $r_1 = 950$, $r_2 = 665$, $r_3 = 445$, $r_4 = 175$, $r_5 = 60$

Solution 7.23

We will first have to estimate the survival function at the borders of the intervals. Using the formulas

$$\hat{p}_i = \frac{r_i - x_i}{r_i} \qquad \hat{S}(c_i) = \hat{p}_0 \times \hat{p}_1 \times \cdots \times \hat{p}_{i-1}$$

we have:

i	0	1	2	3	4	5
x_i	30	285	220	230	115	60
r_i	480	950	665	445	175	60
$p_0 \times \cdots \times p_i$	0.9375	0.65625	0.43914	0.21217	0.07274	0

From this table we can see that:

$$\hat{S}(500) = 0.65625 \qquad \hat{S}(750) = 0.43914$$

So the median loss can be determined from linear interpolation:

$$0.50 = \hat{S}(x_{0.5}) = \hat{S}(500) + \frac{\hat{S}(750) - \hat{S}(500)}{750 - 500} \times (x_{0.5} - 500)$$

$$= 0.65625 + \frac{0.43914 - 0.65625}{250} \times (x_{0.5} - 500)$$

$$\Rightarrow \quad x_{0.5} = 679.92$$

Solution 7.24

We can again use linear interpolation with the results in Solution 7.23:

$$\hat{S}(600) = \hat{S}(500) + \frac{\hat{S}(750) - \hat{S}(500)}{750 - 500} \times (600 - 500)$$

$$= 0.65625 - 0.00087 \times 100 = 0.56941$$

$$\Rightarrow \hat{F}(600) = 1 - \hat{S}(600) = 0.43059$$