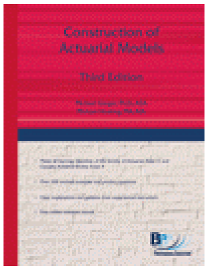

Construction of Actuarial Models

Third Edition

by Mike Gauger and Michael Hosking

Published by BPP Professional Education

Solutions to practice questions – Chapter 13

Solution 13.1

The interval $[0,0.24)$ of random numbers is transformed to $x=1$.

Since $0.24+0.49=0.73$, the interval $[0.24,0.73)$ is transformed to $x=2.3$.

The remaining random numbers in the interval $[0.73,1)$ are transformed to $x=5$.

Solution 13.2

With $\alpha=2.3$ and $\theta=198$, the CDF is:

$$F(x)=1-\left(\frac{\theta}{\theta+x}\right)^{\alpha}=1-\left(\frac{198}{198+x}\right)^{2.3}$$

This function is 1-1 on the support which is the interval $(0,\infty)$. So we need to solve the equation $F(x)=u$ for x in terms of u:

$$u=1-\left(\frac{198}{198+x}\right)^{2.3} \Leftrightarrow \frac{198}{198+x}=(1-u)^{1/2.3} \Leftrightarrow x=198\left(\left((1-u)^{-1/2.3}-1\right)\right)$$

Solution 13.3

We are looking for the 76th percentile of the future lifetime of a life currently age 90. The 76th death among the 100 lives occurs during the year of age between 92 and 93. The formula for the CDF on the range $2\le t\le 3$ is $F(t)=-0.44+0.48t$:

$$0.76=F(t)=-0.44+0.48t \Rightarrow t=\frac{1.20}{0.48}=2.5$$

Solution 13.4

We begin by finding the 35th percentile of future lifetime. The 35th death among the 100 lives occurs in the second year:

$$0.35 = F(t) = -0.18 + 0.35t \;\Rightarrow\; t = \frac{0.53}{0.35} = 1.514$$

$$L = \frac{1{,}000}{1.05^{1.514}} - 420\,\ddot{a}_{\overline{2}|} = 108.79$$

Solution 13.5

The inverse transformation for an exponential distribution with mean 30 is:

$$u = F(t) = 1 - e^{-t/\theta} \;\Rightarrow e^{-t/\theta} = 1-u \;\Rightarrow\; t = -\theta\ln(1-u) = -30\ln(1-u)$$

So with $u_1 = 0.56$, $u_2 = 0.21$ we have $t_1 = 24.63$ and $t_2 = 7.07$. So the simulated time of the first death is 7.07, and the simulated time of the second death is 24.63.

Solution 13.6

The 43th departure out of 100 lives occurs in the third year. So $u_1 = 0.43$ is transformed to $k = 2$. For $k = 2$ we have 4 of the 19 total departures are due to mode 1. So if $u_2 < 4/19$ we simulate the mode by 1. Otherwise the mode is 2. So with $u_2 = 0.18$ and $4/19 = 0.21$ we simulate the mode by 1.

Solution 13.7

The bootstrap approximation to the mean squared error is:

$$MSE_e\left(S_X^2\right) = E\left[\left(S_Y^2 - \text{var}(Y)\right)^2\right]$$

where Y is equally likely to take the value 1 or 5. The mean of the empirical distribution is obviously 3 and the variance is:

$$\text{var}(Y) = \left(1^2 \times 0.5 + 5^2 \times 0.5\right) - 3^2 = 4$$

So:

$$MSE_e\left(S_X^2\right) = E\left[\left(S_Y^2 - 4\right)^2\right]$$

We now have to compute S_Y^2 for each possible sample of size 2. We know that Y is equally likely to take the value 1 or 5. So the bootstrap samples are:

1,1 1,5 5,1 5,5

If the outcome is 1,1, then $\bar{Y} = 1$ and $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 = 0$.

If the outcome is 1,5, then $\bar{Y}=3$ and $S_Y^2=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=\sum_{i=1}^{2}\left(Y_i-3\right)^2=8$.

If the outcome is 5,1, then $\bar{Y}=3$ and $S_Y^2=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=\sum_{i=1}^{2}\left(Y_i-3\right)^2=8$.

If the outcome is 5,5, then $\bar{Y}=5$ and $S_Y^2=0$.

Since each of the samples has a probability of 0.25, we have:

$$MSE_e\left(S_X^2\right)=E\left[\left(S_Y^2-4\right)^2\right]$$
$$=0.25\left(0-4\right)^2+0.25\left(8-4\right)^2+0.25\left(8-4\right)^2+0.25\left(0-4\right)^2=16$$

Solution 13.8

We have to compute $S_Y'^2$ for each of the samples:

If the sample is 1, 1, then $\bar{Y}=1$ and $S_Y'^2=\dfrac{1}{n}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=0$.

If the sample is either 1, 5 or 5, 1, then $\bar{Y}=3$ and $S_Y'^2=\dfrac{1}{n}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=\dfrac{4+4}{2}=4$

If the sample is 5, 5, then $\bar{Y}=5$ and $S_Y'^2=\dfrac{1}{n}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=0$

So the bootstrap approximation to the MSE is:

$$MSE_e\left(S_X'^2\right)=E\left[\left(S_Y'^2-4\right)^2\right]$$
$$=0.25\left(0-4\right)^2+0.25\left(4-4\right)^2+0.25\left(4-4\right)^2+0.25\left(0-4\right)^2=8$$

Solution 13.9

The distribution function for the Pareto distribution is:

$$F(x) = 1 - \left(\frac{\theta}{\theta + x}\right)^{\alpha}$$

Setting this equal to u and rearranging, we get:

$$x = \theta(1-u)^{-1/\alpha} - \theta$$

So using u_1, u_2 and u_3 from the table above we get:

$$x_1 = \theta(1-0.49)^{-1/\alpha} - \theta = 118.76$$

$$x_2 = \theta(1-0.36)^{-1/\alpha} - \theta = 77.22$$

$$x_3 = \theta(1-0.27)^{-1/\alpha} - \theta = 53.85$$

These are our simulated values from the Pareto distribution.

Solution 13.10

For the $Poisson(4)$ distribution we have:

$\Pr(0) = 0.0183$ $\Pr(1) = 0.0732$ $\Pr(2) = 0.1465$

and so on. So we can assign random values from this distribution as follows:

Value from $U[0,1)$	Assigned Poisson value
$(0, 0.0183)$	0
$(0.0183, 0.0915)$	1
$(0.0915, 0.2380)$	2

and so on.

So applying the value of 0.11 from the tables, we have a simulated Poisson value of 2.

So we now have 2 claims in our compound Poisson distribution. So the simulated value of S is (using the first two x values from the previous solution):

$$s = 118.76 + 77.22 = 195.98$$

This is a simulated value from the compound Poisson distribution.

Solution 13.11

We could start by finding the CDF of the mixture distribution, and then using the inversion method to generate values from the appropriate distribution. However, it will be easier to proceed as follows:

(1) Use the first u-value to determine which of the two exponential distributions we are sampling from.

(2) Having decided this, generate a single value from the appropriate exponential distribution.

Using this method here, we start by saying that if our first value is in the range $[0,0.6)$ we will use the exponential distribution with parameter 0.02, and if it is in the range $(0.6,1)$ we will use the exponential distribution with parameter 0.01.

Here we have $u_5 = 0.52$, so we will use the first exponential distribution.

So simulating from an exponential distribution with parameter 0.02 (*ie* mean 50):

$$X = -\theta \ln(1-U) = -50\ln(1-U)$$
$$u_6 = 0.92 \Rightarrow x = 126.29$$

Solution 13.12

We first find the distribution of the time at which the benefit will be paid. This will be $t=1$ if the policyholder dies in the first year, $t=2$ if he dies in the second year, or $t=3$ otherwise.

The probability of death in the first year is:

$$q_x = 0.002$$

For death to occur in the second year, the policyholder has to survive the first year:

$$p_x q_{x+1} = {}_2q_x - {}_1q_x = 0.004 - 0.002 = 0.002$$

So the probability of death in the second year is also 0.002.

So the probability that the benefit is paid at the end of the third year is .996.

The value of the loss function will be (dependent on when the benefit is paid):

Death in the first year: $L = 1000v - 300 = 643.40$

Death in the second year: $L = 1000v^2 - 300(1+v) = 306.98$

Death in the third year/survival $L = 1000v^3 - 300(1+v+v^2) = -10.40$

So we want to simulate a random variable that takes these values with the relevant probabilities.

The question tells us to assume that small random numbers correspond to small losses, *ie* to large survival ages. So we will assign the random values:

$(0,0.996)$ to a loss of -10.40 (profit of 10.40)

$(0.996,0.998)$ to a loss of 306.98

$(0.998,1)$ to a loss of 643.40

Our random value here is $u_7 = 0.24$, so we get a loss of -10.40, *ie* the payment is made at the end of Year 3.

Solution 13.13

The distribution function for the Weibull distribution is (from the Tables):

$$F(x) = 1 - e^{-(x/\theta)^\tau}$$

Setting this equal to u and rearranging, we get:

$$x = \theta\left[-\log(1-u)\right]^{1/\tau}$$

So the values of our simulated distribution are:

$$x_1 = 5{,}000\left(-\log(1-.59)\right)^{½} = 4{,}721.22$$

$$x_2 = 5{,}000\left(-\log(1-.58)\right)^{½} = 4{,}656.99$$

$$x_3 = 5{,}000\left(-\log(1-.29)\right)^{½} = 2{,}926.13$$

We now need the sample lower quartile. For a sample of size 3, we take the smallest value as the sample lower quartile (the middle value is the median and the largest the upper quartile). So our simulated lower quartile value is 2,926.

Solution 13.14

We cannot express the distribution function of the normal distribution as an algebraic expression. However, we can use the equation percentiles argument together with the normal tables, to generate values (initially) of an $N(0,1)$ distribution. We can then use the result that if Z is $N(0,1)$, $X = 100Z + 400$ is $N(400,10000)$.

Consider our first random number, 0.33. This is the 33rd percentile of the $U(0,1)$ distribution. The corresponding percentile of $N(0,1)$ is the value z for which $\Phi(z) =$ 0.33. From the Tables, this is $z = -0.44$. So the corresponding value from the $N(400,10000)$ distribution is $x = 100z + 400 = 356$.

Repeating the process for the other three random numbers, we obtain the sample values:

$$\Phi^{-1}(0.54) = 0.10 \;\Rightarrow\; x = 100(0.10) + 400 = 410$$

$$\Phi^{-1}(0.51) = 0.025 \;\Rightarrow\; x = 100(0.025) + 400 = 402.5$$

$$\Phi^{-1}(0.70) = 0.525 \;\Rightarrow\; x = 100(0.525) + 400 = 452.5$$

So our random sample is 356, 410, 402.5, 452.5.

Solution 13.15

We take our random numbers in pairs. Using 0.33 and 0.54 initially, we obtain using the formula given the corresponding values:

$$z_1 = \sqrt{-2\ln 0.33}\cos(2\pi(0.54)) = -1.442$$

and: $$z_2 = \sqrt{-2\ln 0.33}\sin(2\pi(.54)) = -0.370$$

We can now use the same process as before to generate values from $N(400,10000)$:

$$x_1 = 400 + 100z_1 = 256 \qquad \text{and:} \qquad x_2 = 400 + 100z_2 = 363$$

Using the other pair of random numbers in the same way gives us the values 369 and 290.

Solution 13.16

The distribution of the total score obtained when two dice are thrown is as follows:

$\Pr(X=2) = 1/36 \qquad \Pr(X=3) = 2/36 \qquad \Pr(X=4) = 3/36 \qquad \ldots$

$\Pr(X=7) = 6/36 \qquad \Pr(X=8) = 5/36 \qquad \Pr(X=9) = 4/36 \qquad \ldots$

$\Pr(X=12) = 1/36$

We can use this distribution directly with the inversion method. Alternatively (and probably more easily) we can simulate two individual dice scores and then add them together.

If we adopt the latter approach (using 0 to 1/6 to correspond to a score of 1, 1/6 to 2/6 to correspond to a score of 2, and so on, our values of 0.49 and0.36 give us two simulated dice scores of 3 and 3. Our simulated total is therefore 6. Repeating the process gives us scores of 2 and 1 (total 3) and 4 and 6 (total 10).

Solution 13.17

The bootstrap approximation is:

$$E[(\bar{Y} - \mu_Y)^2]$$

where Y has the empirical distribution, $\Pr(Y=1) = \Pr(Y=2) = ½$.

The distribution of $\bar{Y}$ is:

$\Pr(\bar{Y}=1) = 1/4 \qquad \Pr(\bar{Y}=1.5) = 1/2 \qquad \Pr(\bar{Y}=2) = 1/4$

So the mean of $\bar{Y}$ is 1.5, and:

$$E[(\bar{Y} - 1.5)^2] = (1-1.5)^2 \times 1/4 + (1.5-1.5)^2 \times 1/2 + (2-1.5)^2 \times 1/4 = 1/8$$

Solution 13.18

We first need the sample variance:

$$s^2 = \frac{1}{9}\left(\left(\sum x_i^2\right) - 10\bar{x}^2\right) = \frac{1}{9}\left(56 - 10\times 2.2^2\right) = 0.8444$$

So the bootstrap estimate is:

$$\frac{(n-1)s^2}{n^2} = \frac{9}{100}\times 0.8444 = 0.076$$

Solution 13.19

We now have, for the random sample in Solution 13.17:

1,1	$\bar{Y} = 1$	$s_Y^2 = 0$
1,2	$\bar{Y} = 1.5$	$s_Y^2 = 1/2$
2,1	$\bar{Y} = 1.5$	$s_Y^2 = 1/2$
2,2	$\bar{Y} = 2$	$s_Y^2 = 0$

The variance of Y is:

$$\text{var}(Y) = 1^2 \times 1/2 + 2^2 \times 1/2 - 1.5^2 = 0.25$$

So the mean squared error is:

$$1/4 \times 0.25^2 \times 4 = 0.0625$$

Solution 13.20

The chi square statistic for the sample data is:

$$\frac{(8-5)^2}{5} + \frac{(2-5)^2}{5} = 3.6$$

We need to calculate the corresponding sample statistics for the 10 sets of simulated data. In order, we obtain from these:

0.4	0.4
6.4	0
0.4	0
0	0
0.4	1.6

The estimated p-value is the proportion of these values that exceeds 3.6, which is 1 in 10 or 10%. So our estimated p-value is 10%.