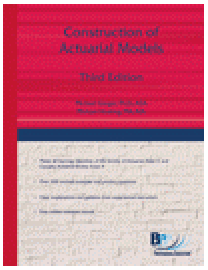

Construction of Actuarial Models

Third Edition

by Mike Gauger and Michael Hosking

Published by BPP Professional Education

Solutions to practice questions – Chapter 5

Solution 5.1

An estimate is a number, which is calculated using some sample data.

An estimator is a random variable. So its value depends on the outcome of some experiment and it has a statistical distribution.

Solution 5.2

The likelihood function is:

$$L = \prod_{i=1}^{n} f(x_i)$$
$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$
$$= C\sigma^{-n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

where C is a constant. Because there are now 2 unknown parameters, we have to differentiate with respect to each parameter and solve 2 simultaneous equations.
Taking logs, we obtain:

$$\ln L = \ln C - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

If we make the substitution $\nu = \sigma^2$, then this becomes:

$$\ln L = \ln C - \frac{n}{2}\ln\nu - \frac{1}{2\nu}\sum_{i=1}^{n}(x_i-\mu)^2$$

Differentiating with respect to μ and ν:

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\nu}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\nu}\left(\sum_{i=1}^{n}x_i - n\mu\right)$$

$$\frac{\partial \ln L}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{2\nu}\left[-n + \frac{1}{\nu}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

Setting these equal to 0 gives:

$$\sum_{i=1}^{n} x_i - n\hat{\mu} = 0 \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

and:

$$-n + \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

So what we have shown here is that the MLEs for μ and σ^2 in the normal distribution are the sample mean and the sample variance (calculated using a denominator of n). These results seem intuitively reasonable.

Solution 5.3

The estimator is:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\bar{X} + \bar{X}^2\right)$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - 2\bar{X}\sum_{i=1}^{n} X_i + n\bar{X}^2\right] = \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

and its expected value is:

$$E\left(\hat{\sigma}^2\right) = E\left[\frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)\right] = \frac{1}{n}\left[\sum_{i=1}^{n} E\left(X_i^2\right) - nE\left(\bar{X}^2\right)\right]$$

Since $X_i \sim N\left(\mu, \sigma^2\right)$ for $i = 1, 2, \ldots, n$, it follows that $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and:

$$E\left(X_i^2\right) = \text{var}\left(X_i\right) + \left[E\left(X_i\right)\right]^2 = \sigma^2 + \mu^2 \quad \text{for } i = 1, 2, \ldots, n$$

$$E\left(\bar{X}^2\right) = \text{var}\left(\bar{X}\right) + \left[E\left(\bar{X}\right)\right]^2 = \frac{\sigma^2}{n} + \mu^2$$

Substituting these into the expression for $E\left(\hat{\sigma}^2\right)$, we obtain:

$$E\left(\hat{\sigma}^2\right) = \frac{1}{n}\left[\sum_{i=1}^{n}\left(\sigma^2 + \mu^2\right) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right]$$

$$= \frac{1}{n}\left[n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right]$$

$$= \frac{n-1}{n}\sigma^2$$

Since $E\left(\hat{\sigma}^2\right) \neq \sigma^2$, $\hat{\sigma}^2$ is a biased estimator of σ^2. (But as theoretically guaranteed, it is asymptotically unbiased since the bias goes to zero as n increases.)

Solution 5.4

We need the mean and variance of $\bar{X}$. Using the standard results, we have:

$$E(\bar{X}) = \mu \qquad \text{and:} \qquad \text{var}(\bar{X}) = \sigma^2 / n$$

We can see from these that both the conditions for consistency (Section 5.4) are satisfied, and so $\bar{X}$ is a consistent estimator for μ.

In fact it is also true that $S^2 = \frac{1}{n-1}\sum (X_i - \bar{X})^2$ *is a consistent estimator for* σ^2. *You might like to check to see whether you can prove that this is also true.*

Solution 5.5

The MSE measures the amount of "squared deviation" of the estimator from the parameter. If this squared deviation is small, then the estimator is fairly close to the true value of the parameter. The smaller the MSE, the closer the estimator is on average, whatever the true value of the parameter.

You might have thought that there would have been other measures of this closeness that might be superior, for example $E(\theta - \hat{\theta})$ or $E(|\theta - \hat{\theta}|)$. However, there are problems with both of these apparently simpler possibilities. The first can be minimized by any unbiased estimator, and so is not sufficiently distinctive. The second has the problems of calculation that are associated with modulus functions. So the definition of mean squared error given here is preferable as a measure of "goodness-of-fit" to the true parameter value.

Solution 5.6

Let T_j denote the j th lifetime random variable. The PDF of T_j is:

$$f_{T_j}(t_j) = \frac{t_j^{39} e^{-t_j/\theta}}{\theta^{40}\,\Gamma(40)} \quad \text{for } t_j > 0$$

Assuming that the observed lifetimes are independent, the likelihood function is the product of the individual PDFs, *ie*:

$$L = C\prod_{j=1}^{10} \theta^{-40}\, e^{-t_j/\theta}$$

where C is a constant, *ie* it doesn't depend on θ.

Since we want to determine the value of θ that yields the maximum value of L, we have to differentiate L with respect to θ. However, as L involves a product of terms, our differentiation will be made easier if we first take (natural) logs. This turns the product into a sum. Also, since the log function is monotonically increasing, maximizing $\ln L$ is equivalent to maximizing L. We may use l to represent $\log L$.

Taking logs gives:

$$l = \ln L = \ln C - 400\ln\theta - \frac{1}{\theta}\sum_{j=1}^{10} T_j$$

Now, differentiating $\ln L$ with respect to θ, we have:

$$\frac{d\ln L}{d\theta} = -\frac{400}{\theta} + \frac{1}{\theta^2}\sum_{j=1}^{10} T_j = \frac{1}{\theta}\left(-\,400 + \frac{1}{\theta}\sum_{j=1}^{10} T_j\right)$$

Setting this equal to 0, we obtain the maximum likelihood estimator:

$$\hat{\theta} = \frac{1}{400}\sum_{j=1}^{10} T_j$$

And inserting the observed values of the T_j gives the maximum likelihood estimate:

$$\hat{\theta} = \frac{753}{400} = 1.8825$$

We should actually check that we have a maximum by differentiating the log-likelihood again. We obtain:

$$\frac{d^2 \ln L}{d\theta^2} = \frac{400}{\theta^2} - \frac{2}{\theta^3}\sum_{j=1}^{10} T_j$$

Evaluating this using the observed data and the estimate $\hat{\theta} = 1.8825$, gives:

$$\frac{d^2 \ln L}{d\theta^2} = \frac{400}{1.8825^2} - \frac{1,506}{1.8825^3} = -112.87$$

Since $\frac{d^2 \ln L}{d\theta^2} < 0$ when $\theta = 1.8825$, the critical point is a maximum, and the maximum likelihood estimate of θ is $\hat{\theta} = 1.8825$.

Now we turn to the asymptotic variance. From the above we have:

$$\frac{d^2 \ln L}{d\theta^2} = \frac{400}{\theta^2} - \frac{2}{\theta^3}\sum_{j=1}^{10} T_j$$

This has expected value:

$$\begin{aligned} E\left(\frac{d^2 \ln L}{d\theta^2}\right) &= \frac{400}{\theta^2} - \frac{2}{\theta^3}\sum_{j=1}^{10} E\left(T_j\right) \\ &= \frac{400}{\theta^2} - \frac{2}{\theta^3}\sum_{j=1}^{10} 40\theta \qquad \text{since } T_j \sim Gamma(40,\theta) \\ &= \frac{400}{\theta^2} - \frac{800}{\theta^2} \\ &= -\frac{400}{\theta^2} = -I(\theta) \end{aligned}$$

So the asymptotic variance of $\hat{\theta}$ is $\frac{1}{I(\theta)} = \frac{\theta^2}{400}$. As we have estimated the value of θ to be 1.8825, we estimate the asymptotic variance to be $\frac{1.8825^2}{400} = 0.00886$.

Solution 5.7

From the *Tables*, we have:

$$E(X)=e^{\mu+\frac{1}{2}\sigma^2}$$

and:

$$E\left(X^2\right)=e^{2\mu+2\sigma^2}$$

So:

$$\begin{aligned}\text{var}(X)&=E\left(X^2\right)-\left[E\left(X\right)\right]^2\\&=e^{2\mu+2\sigma^2}-e^{2\mu+\sigma^2}\\&=e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)\end{aligned}$$

Equating the theoretical mean to the sample mean and the theoretical variance to the sample variance, we obtain the equations:

$$e^{\mu+\frac{1}{2}\sigma^2}=600$$

$$e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)=1,600$$

Squaring the first of these and substituting into the second gives:

$$600^2\left(e^{\sigma^2}-1\right)=1,600$$

So:

$$\sigma^2=\ln\left(\frac{1,600}{600^2}+1\right)=0.00443$$

and:

$$\mu=\ln 600-\frac{1}{2}\sigma^2=6.39471$$

Note. The *Tables* refer to μ and σ as the parameters. Here we have called μ and σ^2 the parameters. No difference is intended.

Solution 5.8

From the Tables:

$$E(X^{-1})=\theta^{-1}\Gamma(2)=\frac{1}{\theta}$$

(Note that for an integer n, the gamma function is $\Gamma(n)=(n-1)!$)

(Note also that the formula is valid for $k<1$ for this distribution.)

Taking the sample values from the Example 5.11, we have:

$$\sum x_i^{-1}=0.266917$$

Equating these, we find that $\hat{\theta}=56.1972$.

Solution 5.9

The median lies between the 4th and 5th of these values.
The 4th value, or $400/9 = 44.44$ th percentile is 547.

The 5th value, or $500/9 = 55.56$ th percentile is 609.

Interpolating gives:

$$x_{0.5} \approx 547 + \frac{(50 - 400/9)}{(500/9 - 400/9)} \times (609 - 547) = 578$$

Solution 5.10

To find the percentile matching estimate of θ, we equate the theoretical median to the sample median:

$$F_X(1,045) = 0.5 \Rightarrow 1 - e^{-1,045/\theta} = 0.5$$
$$\Rightarrow e^{-1,045/\theta} = 0.5$$
$$\Rightarrow -\frac{1,045}{\theta} = \ln 0.5$$
$$\Rightarrow \theta = -\frac{1,045}{\ln 0.5} = 1,507.616$$

Solution 5.11

From the *Tables*, the distribution function of the inverse Weibull distribution is:

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

Equating the theoretical median to the sample median, and the theoretical 90th percentile to the 90th percentile of the sample:

$$e^{-(\theta/1,000)^\tau} = 0.5$$
$$e^{-(\theta/2,500)^\tau} = 0.9$$

Taking logs:

$$\left(\frac{\theta}{1,000}\right)^\tau = -\ln 0.5$$
$$\left(\frac{\theta}{2,500}\right)^\tau = -\ln 0.9$$

So:

$$\left(\frac{\theta}{1,000}\right)^\tau \Big/ \left(\frac{\theta}{2,500}\right)^\tau = 2.5^\tau = \frac{\ln 0.5}{\ln 0.9}$$

and taking logs again:

$$\tau \ln 2.5 = \ln\left(\frac{\ln 0.5}{\ln 0.9}\right) \Rightarrow \tau = 2.05596$$

We can also find the estimate for θ by substituting back if necessary.

Solution 5.12

Let X_i denote the amount of the ith claim. The posterior distribution of θ is given by the conditional probabilities $\Pr\left(\theta = 1,000 \mid \sum_{i=1}^{5} X_i = 6,258\right)$, $\Pr\left(\theta = 1,200 \mid \sum_{i=1}^{5} X_i = 6,258\right)$ and $\Pr\left(\theta = 1,500 \mid \sum_{i=1}^{5} X_i = 6,258\right)$.

The first of these is given by:

$$\Pr\left(\theta = 1,000 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{f\left(\sum_{i=1}^{5} X_i = 6,258 \,\middle|\, \theta = 1,000\right)\Pr(\theta = 1,000)}{f\left(\sum_{i=1}^{5} X_i = 6,258\right)}$$

However, since the X_i's are independent $Exp(\theta)$ random variables, $\sum_{i=1}^{5} X_i \sim Gamma(5,\theta)$. (This is a well-known result and is easily proved using moment generating functions.)

So:

$$f\left(\sum_{i=1}^{5} X_i = 6,258 \,\middle|\, \theta\right) = \frac{(6,258/\theta)^5 e^{-6,258/\theta}}{6,258\,\Gamma(5)} = K\theta^{-5} e^{-6,258/\theta}$$

where $K = \dfrac{6,258^4}{\Gamma(5)}$ (a constant), and:

$$\Pr\left(\theta = 1,000 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{K(1,000)^{-5} e^{-6,258/1,000} \Pr(\theta = 1,000)}{f\left(\sum_{i=1}^{5} X_i = 6,258\right)}$$

$$= \frac{K(1,000)^{-5} e^{-6,258/1,000} \times \frac{1}{3}}{f\left(\sum_{i=1}^{5} X_i = 6,258\right)}$$

Similarly:

$$\Pr\left(\theta = 1,200 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{K(1,200)^{-5} e^{-6,258/1,200} \times \frac{1}{3}}{f\left(\sum_{i=1}^{5} X_i = 6,258\right)}$$

and:

$$\Pr\left(\theta = 1,500 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{K(1,500)^{-5} e^{-6,258/1,500} \times \frac{1}{3}}{f\left(\sum_{i=1}^{5} X_i = 6,258\right)}$$

The probability in the denominator is:

$$f\left(\sum_{i=1}^{5} X_i = 6,258\right) = K\Big[(1,000)^{-5}\, e^{-6,258/1,000} \times \tfrac{1}{3} + (1,200)^{-5}\, e^{-6,258/1,200} \times \tfrac{1}{3}$$

$$+(1,500)^{5}\, e^{-6,258/1,500} \times \tfrac{1}{3}\Big]$$

$$= K\Big[6.385574\times 10^{-19} + 7.2799231\times 10^{-19}$$

$$+ 6.7693329\times 10^{-19}\Big]$$

$$= 2.0432830\times 10^{-18} \times K$$

So:

$$\Pr\left(\theta = 1,000 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{6.383574\times 10^{-19} \times K}{2.0432830\times 10^{-18} \times K} = 0.31$$

$$\Pr\left(\theta = 1,200 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{7.2799231\times 10^{-19} \times K}{2.0432830\times 10^{-18} \times K} = 0.36$$

$$\Pr\left(\theta = 1,500 \,\middle|\, \sum_{i=1}^{5} X_i = 6,258\right) = \frac{6.7693329\times 10^{-19} \times K}{2.0432830\times 10^{-18} \times K} = 0.33$$

Solution 5.13

The PDF of the prior distribution of λ is:

$$f_{prior}(\lambda) = \frac{(2,800/\lambda)^2\, e^{-2,800/\lambda}}{\lambda\,\Gamma(2)} \qquad \lambda > 0$$

The likelihood function is given by:

$$L(\lambda) = \prod_{i=1}^{5} f_X(x_i) = \prod_{i=1}^{5} \frac{1}{\lambda} e^{-x_i/\lambda} = \frac{1}{\lambda^5} e^{-\sum x_i/\lambda} = \frac{1}{\lambda^5} e^{-7,504/\lambda}$$

So the PDF of the posterior distribution of λ is:

$$f_{post}(\lambda) = C\,\frac{1}{\lambda^8} e^{-10,304/\lambda} \qquad \lambda > 0$$

where C is a constant. Hence the posterior distribution of λ is inverse gamma with parameters $\alpha = 7$ and $\theta = 10,304$.

Solution 5.14

The PDF of the prior distribution of λ is:

$$f_{prior}(\lambda) = \frac{(\lambda/2)^8 \, e^{-\lambda/2}}{\lambda\,\Gamma(8)} \qquad \lambda > 0$$

Now let N_j denote the number of claims in year j. Since $N_j \sim Poi(\lambda)$, the likelihood function is given by:

$$L(\lambda) = \Pr(N_1 = 12)\Pr(N_2 = 12) = \frac{e^{-\lambda}\lambda^{12}}{12!}\frac{e^{-\lambda}\lambda^{12}}{12!} = \frac{e^{-2\lambda}\lambda^{24}}{(12!)^2}$$

The PDF of the posterior distribution of λ is therefore:

$$f_{post}(\lambda) = C\,\lambda^{31}\,e^{-5\lambda/2} \quad \lambda > 0$$

where C is a constant. We recognize this as the PDF of a $Gamma(32, 0.4)$ distribution. The mean of this distribution is $32 \times 0.4 = 12.8$, so the required Bayesian estimate of λ is 12.8.

Solution 5.15

(i) Let T_i denote the lifetime random variable for the i th life to die and T_j denote the lifetime random variable for the j th life to be censored. The likelihood function is:

$$\begin{aligned}
L &= \prod_{i\in D} f_{T_i}(t_i) \prod_{j\in C} \Pr(T_j > 10) \\
&= \prod_{i\in D} \frac{1}{\theta} e^{-t_i/\theta} \prod_{j\in C} e^{-10/\theta} \\
&= \frac{1}{\theta^4} \exp\left(-\sum_{i\in D} t_i/\theta\right)(e^{-10/\theta})^3 \\
&= \frac{1}{\theta^4} e^{-20.16/\theta} e^{-30/\theta} \\
&= \frac{1}{\theta^4} e^{-50.16/\theta}
\end{aligned}$$

(ii) The log-likelihood is:

$$\log L = -4\log\theta - \frac{50.16}{\theta}$$

Differentiating with respect to θ gives:

$$\frac{d\log L}{d\theta} = -\frac{4}{\theta} + \frac{50.16}{\theta^2}$$

Setting the derivative equal to 0 and rearranging, we then obtain:

$$\theta = \frac{50.16}{4} = 12.54$$

Finally, since:

$$\frac{d^2 \log L}{d\theta^2} = \frac{4}{\theta^2} - \frac{100.32}{\theta^3} = -0.0254 \text{ when } \theta = 12.54$$

$\log L$ has a maximum turning point at $\theta = 12.54$. Thus 12.54 is the maximum likelihood estimate of θ.

A quicker way to establish this result is to note that the log-likelihood is the PDF of an inverse gamma random variable with parameters $\alpha = 3$ and $\theta = 50.16$. This is maximized at the mode.

The formula for the mode of the inverse gamma distribution is given in the Tables as:

$$\frac{\theta}{\alpha+1}=\frac{50.16}{4}=12.54$$

Solution 5.16

Let X_i denote the i th loss random variable. The observed values of the last six loss amounts are:

1250, 1535, 1490, 1604, 2205, 2090

Since only losses in excess of \$1000 are observed, the likelihood function is given by:

$$L=\prod_{i=1}^{6}\frac{f_{X_i}(x_i)}{\Pr(X_i>1000)}$$

Now, since X_i has a Pareto distribution with parameters α and θ, we have:

$$f_{X_i}(x_i)=\frac{\alpha\theta^{\alpha}}{(x_i+\theta)^{\alpha+1}}$$

and:

$$\Pr(X_i>1000)=\left(\frac{\theta}{1000+\theta}\right)^{\alpha}$$

So:

$$L(\theta)=\prod_{i=1}^{6}\left(\frac{\alpha\theta^{\alpha}}{(x_i+\theta)^{\alpha+1}}\right)\left(\frac{\theta}{1000+\theta}\right)^{-\alpha}$$

$$=\prod_{i=1}^{6}\frac{\alpha(1000+\theta)^{\alpha}}{(x_i+\theta)^{\alpha+1}}$$

$$=\frac{\alpha^{6}(1000+\theta)^{6\alpha}}{\left[(1250+\theta)(1535+\theta)(1490+\theta)(1604+\theta)(2205+\theta)(2090+\theta)\right]^{\alpha+1}}$$

Solution 5.17

The last two values in the list represent censored observations. So the likelihood function becomes:

$$L=\prod_{i=1}^{4}\frac{f_{X_i}(x_i)}{\Pr(X_i>1000)}\left[\frac{\Pr(X_i>2000)}{\Pr(X_i>1000)}\right]^{2}$$

Now:

$$\Pr(X_i>2000)=\left(\frac{\theta}{2000+\theta}\right)^{\alpha}$$

So the likelihood is:

$$L=\left(\frac{1000+\theta}{2000+\theta}\right)^{2\alpha}\frac{\alpha^{4}(1000+\theta)^{4\alpha}}{\left[(1250+\theta)(1535+\theta)(1490+\theta)(1604+\theta)\right]^{\alpha+1}}$$

Solution 5.18

The posterior PDF was found to be:

$$f_\Theta\left(\theta \mid X_1 = 500, X_2 = 275\right) = 38{,}133.01 \times \frac{e^{-775/\theta}}{\theta^2} \quad \text{for } 100 \le \theta \le 200$$

The posterior probability that $\Theta > 150$ is determined by integrating the posterior PDF over the region from 150 to 200:

$$\Pr\left(\Theta > 150 \mid X_1 = 500, X_2 = 275\right) = \int_{150}^{200} f_\Theta\left(\theta \mid X_1 = 500, X_2 = 275\right) d\theta$$

$$= \int_{150}^{200} 38{,}133.01 \times \frac{e^{-775/\theta}}{\theta^2} d\theta = 38{,}133.01\left(e^{-775/\theta}\Big|_{150}^{200}\right)$$

$$= 49.20389\left(e^{-3.875} - e^{-5.167}\right) = 0.741$$

Solution 5.19

The Bayesian estimate in this case is the mean of the posterior distribution. Using this result from Solution 5.18 above, we have:

$$\hat{\theta} = E\left[\Theta \mid X_1 = 500, X_2 = 275\right] = \int_{100}^{200} \theta \times f_\Theta\left(\theta \mid X_1 = 500, X_2 = 275\right) d\theta$$

$$= 38{,}133.01 \times \int_{100}^{200} \frac{e^{-775/\theta}}{\theta} d\theta$$

Note: There is no easy way to find an anti-derivative so a numerical approximation would be required to approximate the Bayesian estimate.

Solution 5.20

See Example 5.19. The Bayesian estimate is the posterior mean. The posterior distribution is gamma with:

$$\alpha^* = \alpha + \sum_{i=1}^{n} x_i = 2 + 0 + 1 = 3$$

$$\theta^* = \frac{\theta}{1 + n\theta} = \frac{0.2}{1 + 0.4} = \frac{1}{7}$$

So the Bayesian estimate is the gamma mean $\alpha^*\theta^* = 3/7$.

Solution 5.21

Since the annual number of claims is a Bernoulli variable we have:

$$f_X\left(1 \mid Q = q\right) = q$$

Multiplying this by the prior pdf, we have:

$$f_Q(q \mid X = 1) = c \times q^2 \times q = c\,q^3 \quad \text{for } 0 \le q \le 1$$

So the posterior distribution of q is beta with parameters $a = 4$ and $b = 1$.

Integrating the posterior PDF between 0 and h gives:

$$\int_0^h \frac{\Gamma(5)}{\Gamma(4)\Gamma(1)} q^3\, dq = \int_0^h 4q^3\, dq = h^4$$

Setting this expression equal to 0.025 gives $h = 0.3976$, and setting it equal to 0.975 gives $h = 0.9937$.
So $(0.3976, 0.9937)$ is a 95% Bayesian confidence interval for q.

Solution 5.22

The likelihood function is now:

$$L = f_X(1 | Q = q) f_X(0 | Q = q) = q(1-q)$$

Multiplying this by the prior PDF, we have:

$$f_Q(q | X_1 = 1, X_2 = 0) \ = c \times q^2 \times q(1-q) = c q^3 (1-q) \ \text{ for } 0 \le q \le 1 \qquad 0 < q < 1$$

So the posterior distribution of Q is beta with parameters $a = 4$ and $b = 2$.
Integrating the posterior PDF between 0 and h gives:

$$\int_0^h \frac{\Gamma(6)}{\Gamma(4)\Gamma(2)} q^3(1-q)\, dq = \int_0^h \left(20q^3 - 20q^4\right) dx = 5h^4 - 4h^5$$

When $h = 0.2836$, $5h^4 - 4h^5 = 0.0250$. Also, when $h = 0.9473$, $5h^4 - 4h^5 = 0.9750$. So $(0.2836, 0.9473)$ is a 95% Bayesian confidence interval for q.

Solution 5.23

The likelihood function is:

$$L = \prod_{i=1}^{100} f(x_i) = \prod_{i=1}^{100} \frac{1}{\theta} e^{-x_i/\theta} = \frac{1}{\theta^{100}} e^{-\sum_{i=1}^{100} x_i \big/ \theta} = \frac{1}{\theta^{100}} e^{-100\bar{x}/\theta}$$

Taking logs:

$$\ln L = -100 \ln \theta - \frac{100\bar{x}}{\theta}$$

Differentiating with respect to θ:

$$\frac{d \ln L}{d\theta} = -\frac{100}{\theta} + \frac{100\bar{x}}{\theta^2}$$

Setting this equal to 0:

$$\frac{100}{\theta} = \frac{100\bar{x}}{\theta^2} \Rightarrow \theta = \bar{x} = 800$$

The second derivative of the log-likelihood is:

$$\frac{d^2 \ln L}{d\theta^2} = \frac{100}{\theta^2} - \frac{200\bar{x}}{\theta^3} = -0.00015625 \ \text{ when } \theta = 800$$

So the maximum likelihood estimate of θ is $\hat{\theta} = 800$. The corresponding maximum likelihood estimator of θ is $\tilde{\theta} = \bar{X}$. This estimator is asymptotically normally distributed and its variance is estimated by:

$$\frac{1}{0.00015626} = 6400$$

So a 95% confidence interval for θ is:

$$800 \pm 1.960\sqrt{6400} = 800 \pm 156.8 = (643.2, 956.8)$$

Solution 5.24

The pdf is:

$$f(x)=\frac{\alpha(500)^{\alpha}}{x^{\alpha+1}}$$

So the likelihood function is:

$$L=\prod_{i=1}^{100}\frac{\alpha\left(500\right)^{\alpha}}{x_i^{\alpha+1}}=\frac{\alpha^{100}\,500^{100\alpha}}{\prod x_i^{\alpha+1}}$$

Taking logs:

$$\ln L=100\ln\alpha+100\alpha\ln 500-\left(\alpha+1\right)\sum_{i=1}^{100}\ln x_i$$

Differentiating with respect to α:

$$\frac{d\ln L}{d\alpha}=\frac{100}{\alpha}+100\ln 500-\sum_{i=1}^{100}\ln x_i$$

Setting this equal to 0:

$$\frac{100}{\alpha}+100\ln 500-\sum_{i=1}^{100}\ln x_i=0\Rightarrow\alpha=\frac{100}{\sum\ln x_i-100\ln 500}$$

The second derivative is:

$$\frac{d^2\ln L}{d\alpha^2}=-\frac{100}{\alpha^2}$$

So the maximum likelihood estimate of α is:

$$\hat{\alpha}=\frac{100}{658.617-100\ln 500}=2.691$$

and the estimated standard error is:

$$\sqrt{\frac{\hat{\alpha}^2}{100}}=0.269$$

So a 99% confidence interval for α is:

$$2.691\pm 2.576\times 0.269=\left(1.998,3.384\right)$$