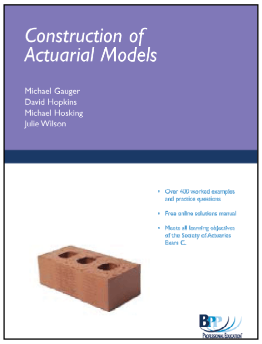

Construction of Actuarial Models

By J. Wilson, M. Hosking, D. Hopkins, and M. Gauger

Published by BPP Professional Education

Solutions to practice questions – Chapter 10

Solution 10.1

Let us use θ for the probability that a driver makes a claim in any year. Using the obvious notation, we have $\theta_B = 0.3$ and $\theta_G = 0.1$. The prior distribution for θ is:

$$\Pr(\theta = \theta_G) = \Pr(\theta = \theta_B) = ½$$

If we have no information, the probability that a driver makes no claims in the coming year can be obtained by conditioning on θ. If X is the number of claims:

$$\begin{aligned}\Pr(X=0) &= \Pr(X=0 \mid \theta=\theta_G)\Pr(\theta=\theta_G) + \Pr(X=0 \mid \theta=\theta_B)\Pr(\theta=\theta_B)\\ &= (0.9)(0.5)+(0.7)(0.5) = 0.8\end{aligned}$$

So the probability that this driver makes one claim in the coming year is 0.2.

Solution 10.2

We now have a policy with some prior information. We know that $X_1 = 0$, and we want to find the conditional probability $\Pr(X_2 = 0 \mid X_1 = 0)$.

Let's first find the posterior distribution for θ. We can use Bayes' Theorem to turn the probabilities around:

$$\begin{aligned}\Pr(\theta=\theta_G \mid X_1=0) &= \frac{\Pr(X_1=0 \mid \theta=\theta_G)\Pr(\theta=\theta_G)}{\Pr(X_1=0 \mid \theta=\theta_G)\Pr(\theta=\theta_G)+\Pr(X_1=0 \mid \theta=\theta_B)\Pr(\theta=\theta_B)}\\ &= \frac{(0.9)(0.5)}{(0.9)(0.5)+(0.7)(0.5)} = 0.5625\end{aligned}$$

So the posterior distribution for θ is that $\theta = \theta_G$ with probability 0.5625, and $\theta = \theta_B$ with probability 0.4375.

We can now find the predictive distribution for X_2 (the number of claims made in the next year):

$$\begin{aligned}\Pr(X_2=0 \mid X_1=0) &= \Pr(X_2=0 \mid \theta=\theta_G)\Pr(\theta=\theta_G \mid X_1=0)+\Pr(X_2=0 \mid \theta=\theta_B)\Pr(\theta=\theta_B \mid X_1=0)\\ &= (0.9)(0.5625)+(0.7)(0.4375) = 0.8125\end{aligned}$$

So that, given the fact that this driver has made no claims in the last year, the probability that he makes no claims in the coming year has risen slightly, from 0.8 to 0.8125.

Solution 10.3

The Bayesian premium is the expected value of this predictive distribution:

$$E\left[X_2 \mid X_1 = 0\right] = (0)(0.8125) + (1)(0.1875) = 0.1875$$

So the expected number of claims for the coming year for this driver is 0.1875.

Solution 10.4

Again we start by finding the posterior distribution for θ. We have:

$$\Pr(\theta = \theta_G \mid X_1 = X_2 = 0) = \frac{\Pr(X_1 = X_2 = 0 \mid \theta = \theta_G)\Pr(\theta = \theta_G)}{\Pr(X_1 = X_2 = 0 \mid \theta = \theta_G)\Pr(\theta = \theta_G) + \Pr(X_1 = X_2 = 0 \mid \theta = \theta_B)\Pr(\theta = \theta_B)}$$

$$= \frac{(0.9^2)(0.5)}{(0.9^2)(0.5) + (0.7^2)(0.5)} = 0.6231$$

Again, the posterior probability that $\theta = \theta_B$ is $1 - 0.6231 = 0.3769$.

We now use the posterior distribution for θ to find the predictive distribution for X.

$$\begin{aligned}\Pr(X_3 = 0 \mid X_1 = X_2 = 0) &= \Pr(X_3 = 0 \mid \theta = \theta_G)\Pr(\theta = \theta_G \mid X_1 = X_2 = 0) \\ &\quad + \Pr(X_3 = 0 \mid \theta = \theta_B)\Pr(\theta = \theta_B \mid X_1 = X_2 = 0) \\ &= (0.9)(0.6231) + (0.7)(0.3769) = 0.8246\end{aligned}$$

Again, the fact that the policyholder now has made no claims for two years affords us a stronger assumption that $\theta = \theta_G$. The likelihood of making no claims next year has increased further, and the probability of making a claim is now 0.1754.

Solution 10.5

(a) The posterior distribution for θ in each case will assign increasingly large values for $\Pr(\theta = \theta_G)$, and decreasingly small values for $\Pr(\theta = \theta_B)$. As n tends to infinity, the probability that $\theta = \theta_G$ will tend to one, and the probability that $\theta = \theta_B$ will tend to zero.

(b) The expected value of the predictive distribution is just $0 \times \theta_G + 1 \times \theta_B = \theta_B$. So if θ_B tends to zero, so will the Bayesian premium.

If a policyholder makes a claim every year, the reverse situation applies. As we obtain more and more information, it becomes more and more likely that this policyholder is in fact a bad driver. So the probability that $\theta = \theta_B$ will tend to one, as will the Bayesian premium.

Solution 10.6

The estimate for the Poisson-gamma model is $\dfrac{(\alpha + \sum x_i)\beta}{1 + n\beta}$.

We have $\alpha = 19$, $\beta = 8$, $\sum x_i = 952$ and $n = 7$. Substituting these values in we get:

$$\frac{8(19 + 952)}{1 + 7 \times 8} = 136.28$$

Solution 10.7

Here we have $Z = \dfrac{\beta n}{1+\beta n} = \dfrac{8\times 7}{1+8\times 7} = \dfrac{56}{57} = 0.9825$.

We are placing lots of emphasis on the data and very little emphasis on the prior. So the answer to Question 10.6 comes out to be very close to the sample mean $\left(\dfrac{952}{7} = 136\right)$ and not very close to the prior mean $(19\times 8 = 152)$.

Solution 10.8

We could choose a gamma distribution with the same mean but a much smaller variance for example $Gamma(1900, 0.08)$ where the mean is the same $(1,900\times 0.08 = 152)$ but the variance is lower $(1,900\times 0.08^2 = 12.16)$. (Z in this case is only 0.36.)

We might do this if we are already pretty certain (before we collected the data for the last seven years) what the value of λ was likely to be. For example, we may have based our prior distribution on data collected over a very long previous period of time, in which case we may be already pretty certain what the value of λ is. In this case we might choose a prior distribution with a much smaller variance.

Solution 10.9

(Warning: the question uses μ for the unknown parameter as opposed to θ, which is used in the text. There is no standard notation, so you need to be careful when reading questions.)

Using the notation given in the text:

$$\sigma_1^2 = 20$$

$$\sigma_2^2 = 40$$

$$n = 50$$

So the credibility factor is:

$$\frac{n}{n+\sigma_1^2/\sigma_2^2} = \frac{50}{50+20/40} = 0.9901$$

Solution 10.10

In the solution of Example 10.1, we found that the predictive distribution for $X_2 \mid X_1 = 15$ was:

$$\Pr(X_2 = 12 \mid X_1 = 15) = \frac{109}{304}$$
$$\Pr(X_2 = 15 \mid X_1 = 15) = \frac{73}{304}$$
$$\Pr(X_2 = 20 \mid X_1 = 15) = \frac{61}{152}$$

$E\left[X_2 \mid X_1 = 15\right]$ is just the mean of this distribution:

$$E\left[X_2 \mid X_1 = 15\right] = 12\left(\frac{109}{304}\right) + 15\left(\frac{73}{304}\right) + 20\left(\frac{61}{152}\right) = 15.93$$

Similarly, we found that the predictive distribution for $X_2 \mid X_1 = 20$ was:

$$\Pr(X_2 = 12 \mid X_1 = 20) = \frac{79}{256}$$
$$\Pr(X_2 = 15 \mid X_1 = 20) = \frac{61}{256}$$
$$\Pr(X_2 = 20 \mid X_1 = 20) = \frac{29}{64}$$

$E\left[X_2 \mid X_1 = 20\right]$ is therefore:

$$E\left[X_2 \mid X_1 = 20\right] = 12\left(\frac{79}{256}\right) + 15\left(\frac{61}{256}\right) + 20\left(\frac{29}{64}\right) = 16.34$$

Solution 10.11

Using the notation in the chapter, we have:

$$\sigma_1^2 = 400 \qquad \sigma_2^2 = 1200 \qquad n = 10$$

So the credibility factor is:

$$Z = \frac{n}{n + \sigma_1^2 / \sigma_2^2} = \frac{10}{10 + 400/1200} = 0.96774$$

So the posterior mean is:

$$0.96774 \times 426 + 0.03226 \times 600 = 431.613$$

The posterior variance is:

$$\frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + n\sigma_2^2} = \frac{400 \times 1200}{400 + 10 \times 1200} = 38.7097$$

So the posterior probability that μ is greater than 450 is:

$$1 - \Phi\left(\frac{450 - 431.613}{\sqrt{38.7097}}\right) = 1 - \Phi(2.9553) = 0.0016$$

or about 0.16%.

The predictive distribution is $N(431.613, 38.7097 + 400) = N(431.613, 438.7097)$. So we now have a probability of:

$$1 - \Phi\left(\frac{450 - 431.613}{\sqrt{438.7097}}\right) = 1 - \Phi(0.87785) = 0.1900$$

So the probability is now about 19%.

Solution 10.12

Since this is a random sample and we are interested in a proportion, we use the beta/binomial model. First we need the beta distribution parameters. Using the formulas for the mean and variance of the beta distribution, we have:

$$\frac{a}{a+b} = 0.6 \qquad \text{and:} \qquad \frac{ab}{(a+b)^2(a+b+1)} = 0.1$$

Solving these simultaneous equations, we find that $a = 0.84$ and $b = 0.56$.

Using the formulas for the parameters of the posterior distribution in the beta/binomial model, we have:

$$a^* = a + x = 187.84$$

$$b^* = b + m - x = 63.56$$

where here $a = 0.84$, $b = 0.56$, $m = 250$, $n = 1$ and $x = 187$.

So the Bayesian estimate of the true underlying proportion in favour is:

$$\frac{a^*}{a^* + b^*} = \frac{187.84}{187.84 + 63.56} = 0.74718$$

and the credibility factor is:

$$Z = \frac{mn}{a + b + mn} = \frac{250}{251.4} = 0.99443$$

Solution 10.13

Using the new value of the proportion in favour of 0.74718, the number in favour in a sample of 250 should have a $\mathit{Binomial}(250, 0.74718)$ distribution. This has mean equal to 186.795, and variance equal to 47.226. So, using a normal approximation to this distribution with a continuity correction, we have:

$$\Pr(X > 200) = 1 - \Phi\left(\frac{200.5 - 186.795}{\sqrt{47.226}}\right) = 1 - \Phi(1.9943) = 0.0231$$

So the probability is about 2.3%.

Solution 10.14

The likelihood function, based on this random sample, will have the form:

$$L(q) = pq^{x_1 - 1} \times pq^{x_2 - 1} \times \cdots \times pq^{x_n - 1} = p^n q^{\sum x_i - n}$$

The beta distribution has the form:

$$f_{prior}(q) \propto q^{a-1}(1-q)^{b-1}$$

The posterior distribution is proportional to the product of these two expressions:

$$f_{post}(q) \propto q^{a-1+\sum x_i - n}(1-q)^{b-1+n}$$

We can see that this has the form of another beta distribution, but with different parameters. So the beta distribution is a conjugate prior in this case.

The new beta parameters are:

$$a^* = a + \sum x_i - n \qquad \text{and:} \qquad b^* = b + n$$

The Bayesian estimate is the mean of this new posterior distribution:

$$\frac{a^*}{a^* + b^*} = \frac{a + \sum x_i - n}{a + b + \sum x_i}$$

We now want to show that this can be expressed as a weighted average of the prior mean (which is $\frac{a}{a+b}$), and the estimate of q from the sample (which can be calculated using maximum likelihood estimation to be $1 - \frac{n}{\sum x_i}$ – you should check that you can agree this expression). So:

$$\frac{a + \sum x_i - n}{a + b + \sum x_i} = \frac{a}{a + b + \sum x_i} + \frac{\sum x_i - n}{a + b + \sum x_i} = \frac{a}{a+b} \times \frac{a+b}{a + b + \sum x_i} + \left(1 - \frac{n}{\sum x_i}\right) \times \frac{\sum x_i}{a + b + \sum x_i}$$

We can now see that the posterior mean has the correct form, and that the credibility factor is the weighting attached to the MLE, which is:

$$Z = \frac{\sum x_i}{a + b + \sum x_i}$$

We can now find the posterior estimate for q:

$$\frac{a + \sum x_i - n}{a + b + \sum x_i} = \frac{4 + 257 - 100}{4 + 6 + 257} = 0.6030$$

Solution 10.15

We can see from the formula for Z that we found in the previous question that if the sum of the sample values increases, then this will increase Z. This is sensible, since this will usually correspond to the situation where we have more sample data. This means that our MLE for q will be based on more sample data, and hence we will expect it to be more reliable. So we want to place more emphasis on the MLE, and less on the sample mean.

If a and b increase, we see that the value of Z will decrease. This is again sensible, since we will choose large values of a and b if we want a prior distribution with a small variance, *ie* we are already pretty sure that we know what the correct value of q is. If we are already sure, then we want to place lots of emphasis on the prior estimate and less on the MLE, and this is exactly what a decrease in the value of Z will do.

Solution 10.16

We want the conditional probability $\Pr(TI \mid X)$, where X represents the event that we actually observed, *ie* one claim of amount 150. We can use Bayes' theorem to turn the probability around in the usual way:

$$\Pr(TI \mid X) = \frac{\Pr(X \mid TI)\Pr(TI)}{\Pr(X \mid TI)\Pr(TI) + \Pr(X \mid TII)\Pr(TII)}$$

We cannot calculate an exact expression for the probability of getting a claim amount of 150, since the claim distribution is continuous. However, we can use a proportioning argument since all we need is the ratio between the two probabilities. Since the portfolio contains equal numbers of the two types of risk, we can assume that the prior probabilities of each type are ½.

So:

$$\begin{aligned}\Pr(TI \mid X) &= \frac{\Pr(X \mid TI)\Pr(TI)}{\Pr(X \mid TI)\Pr(TI) + \Pr(X \mid TII)\Pr(TII)}\\ &= \frac{e^{-2}2^1/1! \times 0.01e^{-0.01\times 150} \times ½}{\left(e^{-2}2^1/1! \times 0.01e^{-0.01\times 150} \times ½\right) + \left(e^{-3}3^1/1! \times 0.005e^{-0.005\times 150} \times ½\right)}\\ &= 0.63127\end{aligned}$$

So the posterior probability that the risk is Type I is about 63%.

Solution 10.17

Since the claims are 6, 2, 2 and 4, we have :

$$f_{post}(\theta) = \frac{\theta e^{-\theta/6}}{6^2} \cdot \frac{\theta e^{-\theta/2}}{2^2} \cdot \frac{\theta e^{-\theta/2}}{2^2} \cdot \frac{\theta e^{-\theta/4}}{4^2} \cdot \frac{e^{-\theta/5}}{5} \propto \theta^4 \cdot e^{-\theta(6^{-1}+2^{-1}+2^{-1}+4^{-1}+5^{-1})}$$

We recognize this as having the form of a gamma distribution with parameters:

$$\frac{1}{\theta^*} = \frac{1}{5} + \frac{1}{x_1} + \cdots + \frac{1}{x_4} = 1.61667$$

and: $\alpha^* = 4$

So our Bayesian estimate is:

$$\alpha^* \theta^* = 2.4742$$

Solution 10.18

Type I risks can produce an aggregate claim amount of 2 if there is one claim for 2 units. The probability of this happening is $0.2 \times 0.5 = 0.1$.

Type II risks can produce an aggregate amount of 2 if there is:

(a) a single claim for 2 units, or

(b) 2 claims each for 1 unit.

The probability of this happening is:

$$0.1 \times 0.6 \times 0.6 + 0.3 \times 0.4 = 0.156$$

So the probability of this happening is 10% for Type I and 15.6% for Type II.

Solution 10.19

We use Bayes' Theorem to turn the probabilities round. So:

$$\Pr(I \mid 2) = \frac{\Pr(2 \mid I)\Pr(I)}{\Pr(2 \mid I)\Pr(I) + \Pr(2 \mid II)\Pr(II)}$$

where the prior probabilities are $\Pr(I) = 0.4$ and $\Pr(II) = 0.6$.

So:

$$\Pr(I \mid 2) = \frac{\Pr(2 \mid I)\Pr(I)}{\Pr(2 \mid I)\Pr(I) + \Pr(2 \mid II)\Pr(II)} = \frac{0.1 \times 0.4}{0.1 \times 0.4 + 0.156 \times 0.6} = 0.2994$$

So the posterior probability that the risk is Type I is 0.2994 (and that it is Type II is 0.7006).

Solution 10.20

We now need the expected aggregate loss for each type of risk. For Type I risks the aggregate loss will be:

0 with probability 0.8

1 with probability $0.2 \times 0.5 = 0.1$

2 with probability 0.1 (as before)

This gives an expected aggregate loss for Type I risks of 0.3.

For Type II risks the aggregate loss will be:

0 with probability 0.6

1 with probability $0.3 \times 0.6 = 0.18$

2 with probability 0.156 (as before)

3 with probability $0.1 \times 0.6 \times 0.4 \times 2 = 0.048$ (NB either order is possible)

4 with probability $0.1 \times 0.4 \times 0.4 = 0.016$

This gives an expected aggregate loss for Type II risks of 0.7.

So using the posterior distribution, the expected aggregate loss from this risk for the coming year will be:

$$0.3 \times 0.2994 + 0.7 \times 0.7006 = 0.58024$$

