MTH5131 Actuarial Statistics

Coursework 3 — Solutions

Exercise 1. 1. Let X be the number of claims received in a week. To determine the posterior distribution of μ, we must calculate the conditional probabilities $P(\mu = 8|X = 7)$, $P(\mu = 10|X = 7)$, and $P(\mu = 12|X = 7)$,. The first of these is

$$P(\mu = 8|X = 7) = \frac{P(\mu = 8, X = 7)}{P(X = 7)} = \frac{P(X = 7|\mu = 8)P(\mu = 8)}{P(X = 7)}$$

Since $X \sim \text{Poisson}(\mu)$,

$$P(X = 7|\mu = 8) = \frac{e^{-8}8^7}{7!}$$

and since the prior distribution is uniform on the integers 8, 10 and 12:

$$P(\mu = 8) = \frac{1}{3}.$$

So

$$P(\mu = 8|X = 7) = \frac{\frac{e^{-8}8^7}{7!} \times \frac{1}{3}}{P(X = 7)} = \frac{0.04653}{P(X = 7)}$$

Similarly,

$$P(\mu = 10|X = 7) = \frac{P(X = 7|\mu = 10)P(\mu = 10)}{P(X = 7)} = \frac{\frac{e^{-10}10^7}{7!} \times \frac{1}{3}}{P(X = 7)} = \frac{0.03003}{P(X = 7)}$$

and

$$P(\mu = 12|X = 7) = \frac{P(X = 7|\mu = 12)P(\mu = 12)}{P(X = 7)} = \frac{\frac{e^{-12}12^7}{7!} \times \frac{1}{3}}{P(X = 7)} = \frac{0.01456}{P(X = 7)}$$

Since these conditional probabilities must sum to 1, the denominator must be the sum of the numerators, so

$$P(X = 7) = 0.04653 + 0.03003 + 0.01456 = 0.09112$$

The posterior probabilities are:

$$P(\mu = 8|X = 7) = \frac{0.04653}{0.09112} = 0.51066$$

$$P(\mu = 10|X = 7) = \frac{0.03003}{0.09112} = 0.32954$$

$$P(\mu = 12|X = 7) = \frac{0.01456}{0.09112} = 0.15980$$

2. The Bayesian estimate under squared error loss is the mean of the posterior distribution:

$$8 \times 0.51066 + 10 \times 0.32954 + 12 \times 0.15980 = 9.29830$$

Exercise 2. 1. Since the prior distribution of p is Beta(4,4):

$$f(p) \propto p^3(1-p)^3$$

Now let X denote the number of successes from a sample of size n. Then $X \sim \mathrm{Binomial}(n, p)$. Since x successes have been observed, the likelihood function is:

$$L(p) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \propto p^x (1-p)^{n-x}$$

Combining the prior PDF with the likelihood function gives:

$$f(p|x) \propto p^3(1-p)^3 \times p^x(1-p)^{n-x} = p^{x+3}(1-p)^{n-x+3}$$

Comparing this with the PDF of the Beta distribution, we see that the posterior distribution of p is Beta$(x+4, n-x+4)$.

The Bayesian estimate under all-or-nothing loss is the mode of the posterior distribution, ie the value of p that maximises the posterior PDF. To find the mode, we need to differentiate the PDF (or equivalently differentiate the log of the PDF) and equate it to zero.

Given that $x = 10$ and $n = 25$, the posterior of p is Beta (15,18) and:

$$f(p|x) = Cp^{14}(1-p)^{17}.$$

Taking logs (to make the differentiation easier):

$$\ln f(p|x) = \ln C + 14 \ln p + 17(1-p)$$

Differentiating gives

$$\frac{d}{dp} \ln f(p|x) = \frac{14}{p} - \frac{17}{1-p}$$

The derivative is equal to 0 when

$$14(1-p) = 17p$$

or

$$p = \frac{14}{31}$$

Differentiating again gives:

$$\frac{d^2}{dp^2} \ln f(p|x) = -\frac{14}{p^2} - \frac{17}{(1-p)^2} < 0$$

and therefore we have found a maximum.

So the Bayesian estimate of p under all-or-nothing loss is $\frac{14}{P31}$ or 0.45161

Exercise 3. 1. The likelihood function is:

$$L(\mu) = \frac{1}{\mu} e^{-\frac{x_1}{\mu}} \times \cdots \times \frac{1}{\mu} e^{-\frac{x_n}{\mu}} = \frac{e^{\frac{-\sum x_i}{\mu}}}{\mu^n}$$

2. The posterior distribution is given by:

$$f(\mu|\underline{x}) \propto \frac{e^{-\frac{\theta}{\mu}}}{\mu^{\alpha+1}} \times \frac{e^{-\frac{\sum x_i}{\mu}}}{\mu^n} = \frac{e^{-\frac{(\theta+\sum x_i)}{\mu}}}{\mu^{n+\alpha+1}}$$

We see that this is the same form as the prior distribution but with different paramters.
So we have the same distribution as before but with parameters:

$$\theta^* = \theta + \sum x_i$$

and

$$\alpha^* = n + \alpha$$

We can now use the formula for the mean of the distribution given in the question:

$$E(\mu) = \frac{\theta^*}{\alpha^* - 1} = \frac{\theta + \sum x_i}{n + \alpha - 1}$$

This is the Bayesian estimate for μ under squared error loss.

3. **Credibility estimate**

$$\hat{\mu} = \frac{\theta + \sum x_i}{n + \alpha - 1} = \frac{\theta}{n + \alpha - 1} + \frac{\sum x_i}{n + \alpha - 1}$$
$$= \frac{\theta}{\alpha - 1} \times \frac{\alpha - 1}{n + \alpha - 1} + \frac{\sum x_i}{n} \times \frac{n}{n + \alpha - 1}$$

This is in the form of a credibility estimate with:

$$Z = \frac{n}{n + \alpha - 1}$$

4. **Posterior estimate**
We now have:

$$\hat{\mu} = \frac{\theta + \sum x_i}{n + \alpha - 1} = \frac{40 + 9826}{100 + 1.5 - 1} = 98.1692$$

The value of the credibility factor is:

$$Z = \frac{n}{n + \alpha - 1} = \frac{100}{100 + 1.5 - 1} = 0.9950$$

5. **Comments**
We see that the value of Z is very close to 1.
This means we are placing almost full weight on our sample mean and take little account of our prior mean.
This is because n is much bigger than α.
We would use a prior like this if we are not very sure initially about the true value of μ.
We have chosen a prior distribution with a large variance, to reflect our high initial degree of uncertainty about μ.

Exercise 4. The prior is $\propto 1$ and the likelihood is $\propto \exp\left(-\frac{(1-\theta)^2}{2}\right)$, so the posterior is $\propto \exp\left(-\frac{(1-\theta)^2}{2}\right)$. We conclude that the posterior is $\theta \sim \mathrm{N}(1, 1)$. Therefore $P(\theta > 0) = 0.8413$ or about 84%. But the observation is consistent with random noise, e.g. $\mathrm{N}(0, 1)$. This is an argument against uninformative priors.

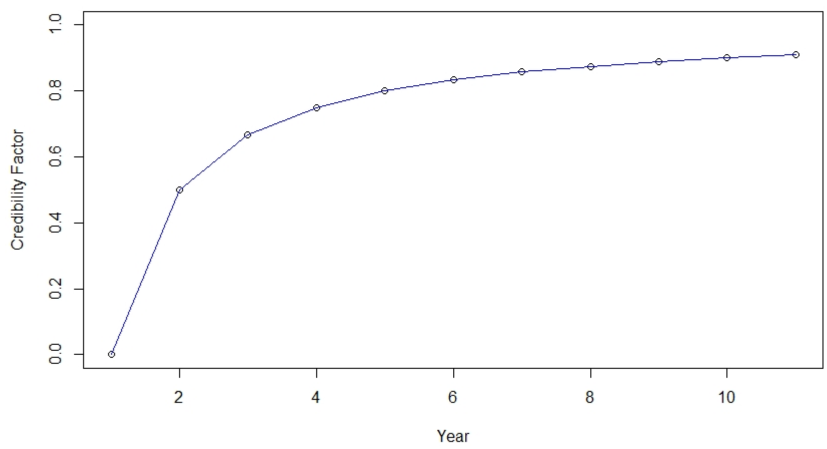
Exercise 5. 1. Assuming $\alpha = 100$, $\beta = 1$:

(a) -

Year	Number of Claims	Z at the start of year	$\overline{X}$ =average number of claims based on number of years of past data available at the start of year	At the start of the year, the credibility estimate of the number of claims in the coming year $= Z\overline{X} + (1 - Z)\mu$
1	144	0.000	0.000	100.000
2	144	0.500	144.000	122.000
3	174	0.667	144.000	129.333
4	148	0.750	154.000	140.500
5	151	0.800	152.500	142.000
6	156	0.833	152.200	143.500
7	168	0.857	152.833	145.286
8	147	0.875	155.000	148.125
9	140	0.899	154.000	148.000
10	161	0.900	152.444	147.200

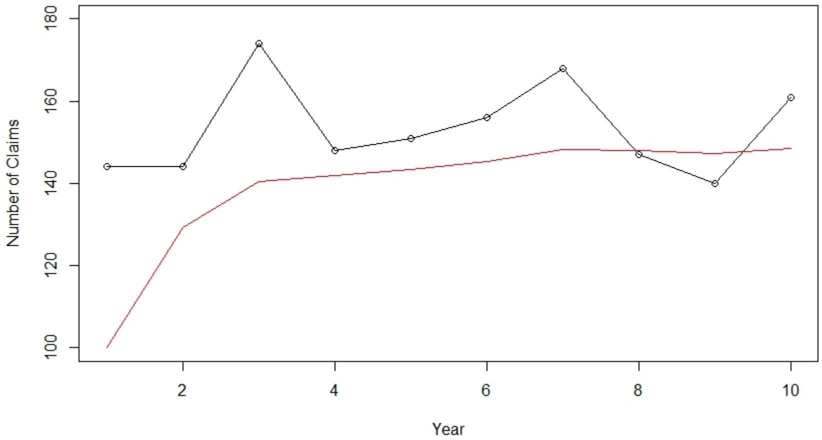
Hint: For the Poisson/Gamma model, in the formula for the credibility factor $Z = n/(n + \beta)$, n should be taken as the number of years of past data.

(b) -



(c) We can see from the graph that the credibility factor increases with time. This fits with our previous comments that as time goes by and more data is collected from the risk itself, then the higher should be the credibility factor. This allows for the increasing reliability of the data in estimating the true but unknown expected number of claims for the risk.

(d) -



(e) The initial estimate of claims is 100, this being the mean of the prior distribution of λ.

However, this turns out to be a very poor estimate since all of the actual claim numbers are around 150 and none, in the first ten years, is lower than 140.

We can see the estimated number of claims increasing with time until it reaches the level of the actual claim numbers after 8 years.

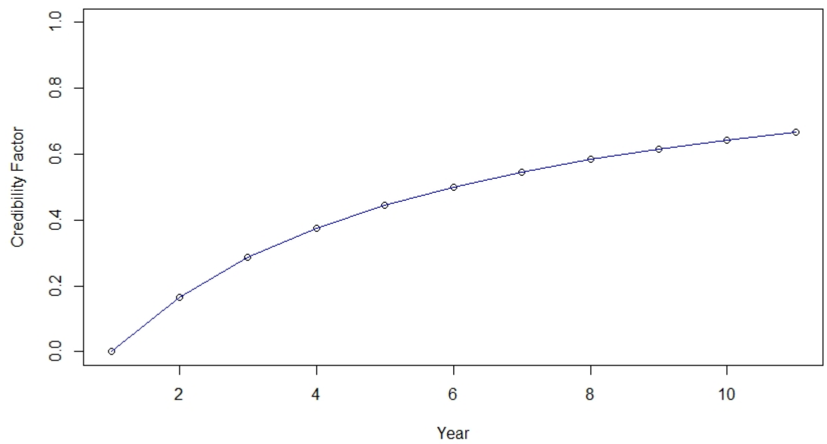
This increase is due to progressively more weight (credibility) being given to the data from the risk itself and correspondingly less weight being given to the collateral data (the prior distribution of λ).

2. Assuming $\alpha = 500$, $\beta = 5$:

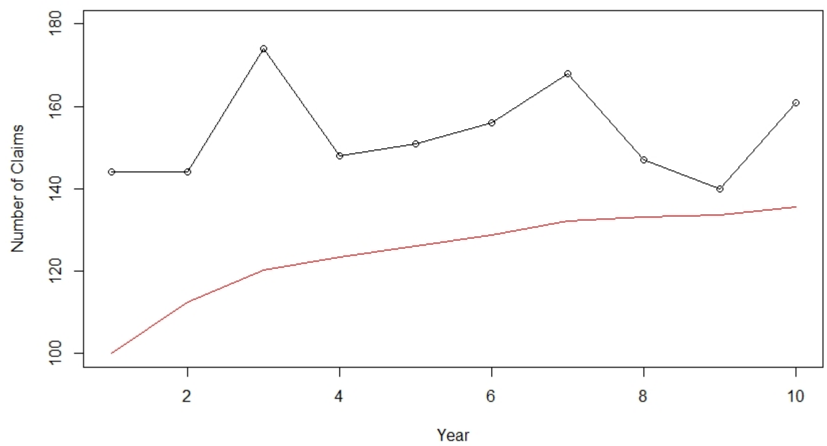
(a) -

Year	Number of Claims	Z at the start of year	$\overline{X}$ =average number of claims based on number of years of past data available at the start of year	At the start of the year, the credibility estimate of the number of claims in the coming year $= Z\overline{X} + (1 - Z)\mu$
1	144	0.000	0.000	100.000
2	144	0.167	144.000	107.333
3	174	0.286	144.000	112.571
4	148	0.375	154.000	120.250
5	151	0.444	152.500	123.333
6	156	0.500	152.200	126.100
7	168	0.545	152.833	128.818
8	147	0.583	155.000	132.083
9	140	0.615	154.000	133.231
10	161	0.643	152.444	133.714

(b) -



(c) -



(d) We can see that both prior distributions Gamma(100,1) and Gamma(500,5) give an increasing credibility factor and the same general features.

However, the most obvious difference is that for the Gamma(500,5) prior the credibility factor increases more slowly.

The mean of both distributions is the same, 100.

This represents the initial credibility estimate of the number of claims for both prior distributions.

However, the standard deviation is lower for the Gamma(500,5) prior at $(500/5^2)^{0.5} = 4.472$ than for the Gamma(100,1) prior at $(100/1^2)^{0.5} = 10$.

The size of the standard deviation of the prior distribution can be interpreted as an indication of how much confidence is placed in the initial estimate of the number of claims.

The smaller the standard deviation of the prior distribution, the more reliable this initial estimate is believed to be.

Since, in Bayesian credibility, the prior distribution represents the collateral data, the above statement can be expressed as "the smaller the standard deviation of the prior distribution, the more relevant the collateral data are considered to be".

Given this interpretation, a smaller standard deviation for the prior distribution would be expected to result in a lower credibility factor.

Exercise 6. The prior is

$$f(\theta) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu}{\sigma_2}\right)^2\right] \propto \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu}{\sigma_2}\right)^2\right]$$

The likelihood is

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma_1}\right)^2\right] \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_1}\right)^2\right]$$

Therefore, the posterior is proportional to

$$\begin{aligned} f(\theta|\underline{y}) &= \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu}{\sigma_2}\right)^2\right] \times \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_1}\right)^2\right] \\ &= \exp\left[-\frac{1}{2}\left[\left(\frac{\theta-\mu}{\sigma_2}\right)^2 + \sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_1}\right)^2\right]\right] \end{aligned}$$

We will prove that

$$f(\theta|\underline{y}) \propto \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu_*}{\sigma_*}\right)^2\right]$$

for some μ_*, σ_* to be determined. This will imply that the posterior is $\mathsf{N}(\mu_*, \sigma_*^2)$. Because of the $\propto$ sign, terms not involving θ are unimportant. We have

$$\left(\frac{\theta-\mu_*}{\sigma_*}\right)^2 = \left(\frac{\theta-\mu}{\sigma_2}\right)^2 + \sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_1}\right)^2 + C$$

for a constant C. Equating the coefficients of θ^2 on both sides gives

$$\frac{1}{\sigma_*^2} = \frac{1}{\sigma_2^2} + \frac{n}{\sigma_1^2} \tag{1}$$

and equating coefficients of θ gives

$$\frac{\mu_*}{\sigma_*^2} = \frac{\mu}{\sigma_2^2} + \frac{\sum_{i=1}^n x_i}{\sigma_1^2} = \frac{\mu}{\sigma_2^2} + \frac{n\overline{x}}{\sigma_1^2} \tag{2}$$

It follows from (1) that

$$\sigma_*^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + n\sigma_2^2} \tag{3}$$

Multiplying (2) by (3) produces

$$\mu_* = \frac{\mu\sigma_1^2 + n\overline{x}\sigma_2^2}{\sigma_1^2 + n\sigma_2^2}$$