Statistical Inference
R: Assessment 2
Solution

I. Kiss and F. Di Lauro

The solution is given in the form of a PDF document as well as an example of R code.

1 Faulty fibres

1. The observations are $x_1, x_2, \ldots, x_{10}$ and the corresponding likelihood is

$$\mathcal{L}(\lambda) = \Pi_{i=1}^{10} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-10\lambda}\lambda^{\sum_{i=1}^{10} x_i}}{\Pi_{i=1}^{10}(x_i!)} \tag{1}$$

The prior is distributed as $\Gamma(2, 12)$ with probability density function

$$\pi(\lambda) = \frac{12^2}{\Gamma(2)}\lambda e^{-12\lambda}. \tag{2}$$

Hence, the posterior is

$$\pi(\lambda|x_1, x_2, \ldots, x_10) \propto \frac{e^{-10\lambda}\lambda^{\sum_{i=1}^{10} x_i}}{\Pi_{i=1}^{10}(x_i!)} \frac{12^2}{\Gamma(2)}\lambda e^{-12\lambda}. \tag{3}$$

Ignoring the constants, using that $\sum_{i=1}^{10} x_i = 8$ and collecting the terms involving λ leads to

$$\pi(\lambda|x_1, x_2, \ldots, x_10) \propto \lambda^9 e^{-22\lambda}, \tag{4}$$

which corresponds to $\Gamma(10, 22)$.

2. The likelihood is

$$\mathcal{L}(x_1, x_2, \ldots, x_{10}|\lambda) = \frac{e^{-10\lambda}\lambda^{\sum_{i=1}^{10} x_i}}{\Pi_{i=1}^{10}(x_i!)}, \tag{5}$$

and the log-likelihood reduces to

$$l(x_1, x_2, \ldots, x_{10}|\lambda) = -10\lambda - \ln(\Pi_{i=1}^{10}(x_i!)) + \ln(\lambda)\sum_{i=1}^{10} x_i. \tag{6}$$

The derivative with respect to λ leads to

$$-10 + \frac{1}{\lambda}\sum_{i=1}^{10} x_i = 0 \Leftrightarrow \lambda_{MLE} = \frac{\sum_{i=1}^{10} x_i}{10} = 8/10 = 0.8. \tag{7}$$

3. From lecture notes or tables of probability distributions, the mean of a random variable distributed as $\Gamma(\alpha, \beta)$ is α/β and the mode (where the probability distribution attains its highest value) of it is $(\alpha - 1)/\beta$ (This can be obtained easily by differentiating the pdf of the distribution and finding the extremal). This leads to

$$\underbrace{\lambda_{MAP} = 9/22}_{\text{mode of the posterior}} < \underbrace{\lambda_{CM} = 10/22}_{\text{mean of the posterior}} < \underbrace{\lambda_{MLE} = 0.8}_{\text{larger than any Bayesian estimates}}$$

The mean of the prior is $2/12 \simeq 0.166$ and this really pulls the estimates towards lower values since the posterior is a combination of prior belief and data.

4. Using R, type the following
shape ← 10
rate ← 22
z← pgamma(0.646, shape, rate, log = FALSE)
one gets $z \simeq 0.9$.

5. There are multiple ways to solve this but one ends up with the figure below and with the left and right limits being $x_l \simeq 0.22$ and $x_r \simeq 0.68$,

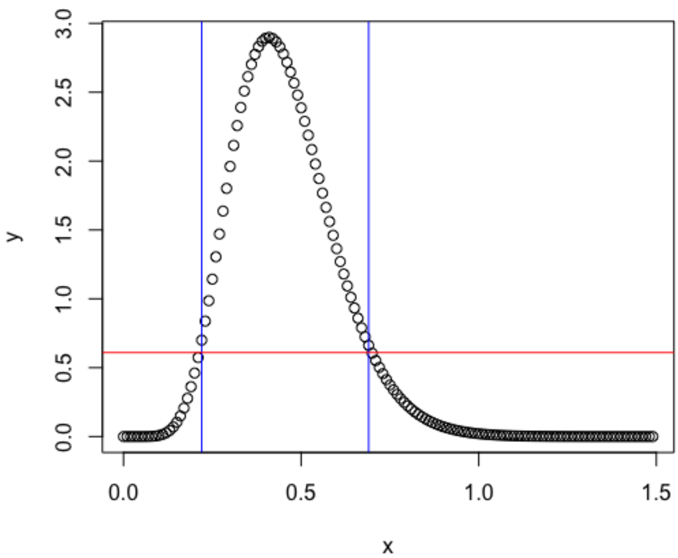

Figure 1: Posterior with the high probability density credible region.

respectively.

6. The probability of no error is

$$\mathcal{P}(\text{no error}) = \int_0^\infty \frac{e^{-\lambda}\lambda^0}{0!} \frac{22^{10}}{\Gamma(10)} \lambda^9 e^{-22\lambda} d\lambda = \left(\frac{22}{23}\right)^{10} \int_0^\infty \frac{23^{10}}{\Gamma(10)} \lambda^9 e^{-23\lambda} d\lambda = \left(\frac{22}{23}\right)^{10} = 0.64.$$

2 Model Selection

In the following, $\{x_1, \ldots, x_n\}$ denote the values of the dataset.

1. The prior is $\pi_1(\alpha, \beta) = \Gamma(\alpha, \beta)$, the likelihood is $\mathcal{L}(\{x_1, \ldots, x_n\}) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$, therefore the posterior distribution is

$$\pi_1(\lambda|(x_1, \ldots, x_n)) \propto \lambda^{\alpha-1} e^{-\beta\lambda} \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \lambda^{\alpha-1+\sum x_i} e^{-(\beta+n)\lambda},$$

   which is $\Gamma(\sum_i^n x_i + \alpha, n + \beta)$.

2. The prior is $\pi_2(\alpha, \beta) = p^\alpha (1-p)^\beta$, the likelihood is $\mathcal{L}(\{x_1, \ldots, x_n\}) = \prod_{i=1}^n p(1-p)^{x_i}$, therefore

$$\pi_2(p|(x_1, \ldots, x_n)) \propto p^\alpha (1-p)^\beta \prod_{i=1}^n p(1-p)^{x_i} = p^{\alpha+n}(1-p)^{\beta+\sum x_i},$$

   which is $\beta(\alpha + n, \beta + \sum x_i)$.

3. See figures 2, 3.

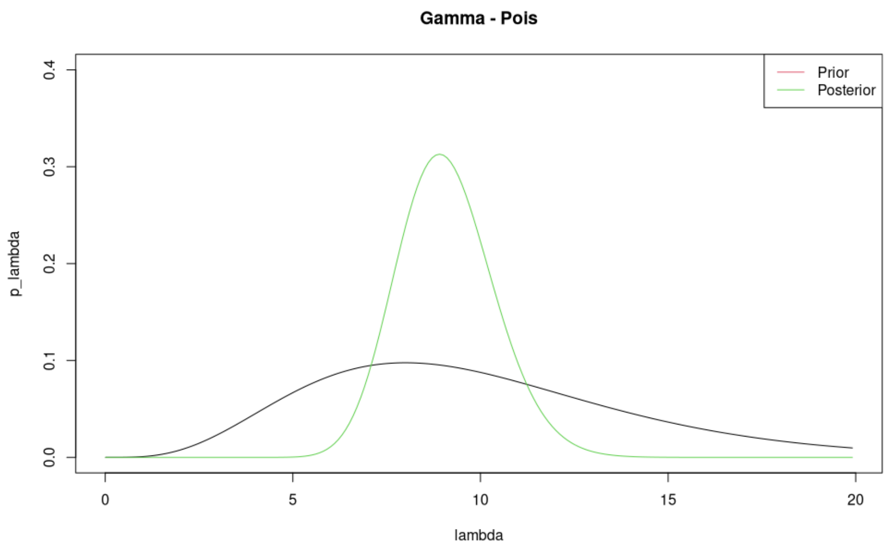

Figure 2: Gamma Poisson prior-posterior distributions

4. The values for the MAP and the CM can be computed as in the code. For the Gamma distribution, is MAP = 8.91, CM = 9.09. For the Beta distribution instead we get MAP = 0.143 and CM = 0.153.

5. the Bayes factor - when the two models have the same a priori probability - is defined as the ratio of the marginals, as shown in the lecture notes

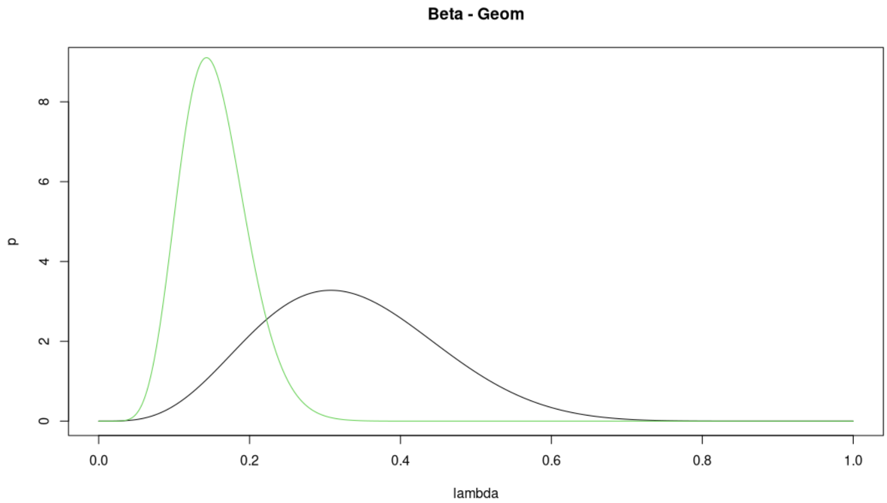

Figure 3: Beta-Geometric prior-posterior distributions

(see eq. (3.114) at page 58), that is

$$\mathcal{B} = \frac{P(\{x_1, \ldots, x_n\}|M_1)}{P(\{x_1, \ldots, x_n\}|M_2)} = \frac{\int_0^\infty \pi_1(\lambda|x)d\lambda}{\int_0^1 \pi_2(p|x)dp},$$

where M_1 and M_2 are the two models (Poisson and Geometric). The evaluations of the integrals is done in the code, and should give back $\mathcal{B} = 12655678$. This means that the Poisson distribution is greatly favoured over the geometric one.

6. With the new prior hyperparameters, we can repeat the calculation in the point (6) and get $\mathcal{B} \sim 0.1383$. In this case neither of the two models is supported by the data.