## 3.1 Exercises

1. Assume the number of infections $X_i$ (per 100 000) in age groups $i = 1, \dots, n$ are reported and $X_i \sim \text{Poisson}(\lambda)$. What is the posterior distribution of the mean incidence $\lambda$? Assume $\text{Gamma}(\alpha, \beta)$ prior. Interpret the prior as 'prior data'.

2. Observed bacterial counts (cfu /10 grams, cfu = colony forming units) in 17 samples were as follows:

```
X <- c(0, 0, 0, 0, 5, 3, 0, 0, 70, 0, 0, 0, 8, 0, 0, 3, 0)
```

Assume Poisson model $X_i \sim \text{Poisson}(\lambda)$. Use $\text{Gamma}(\alpha, \beta)$ prior for $\lambda$. Choose a prior, trying to be uninformative, and compute posterior mean, mode and standard deviation of $\lambda$.

3. The observed life times were

```
X<-c(1.54, 0.70, 1.23, 0.82, 0.99, 1.33, 0.38, 0.99, 1.97, 1.10, 0.40)
```

and there were 4 censored observations at time $T = 2$. Assume $X_i \sim \text{Exp}(\theta)$ and prior $\theta \sim \text{Gamma}(2, 1)$. Compute posterior mean $E(\theta \mid \text{data})$.

4. Assume $X \sim \text{N}(\mu, \sigma^2)$. Derive the posterior of the unknown mean $\mu$, assuming known variance $\sigma^2$ and single observation $x$. Prior is $\text{N}(\mu_0, \sigma_0^2)$. Hint: completion of a square.

5. Using the table of distributions, verify by calculating the means that the mean of $Y$ is the same as the mean of $Y'$ when

$$Y \sim \text{Inv Gamma}(\alpha, \beta), \text{ that is: } Y = 1/X \qquad X \sim \text{Gamma}(\alpha, \beta) \qquad \alpha = \nu/2, \beta = vs^2/2$$

$$Y' \sim \text{Scaled Inv } \chi^2(\nu, s^2), \text{ that is: } Y' = \nu s^2/Z \qquad Z \sim \chi^2(\nu)$$

6. Using R, write a function for calculating the Inverse Gamma density. Use this to plot the Scaled Inverse $\chi^2$ prior density for $\sigma^2$ with parameters $\nu_0 = 2, \sigma_0^2 = 8$. Plot the posterior of $\sigma^2$ assuming the data consist of $N = 20$ measurements with $\frac{1}{20}\sum_{i=1}^{20}(X_i - \mu)^2 = 1.5$. Plot the posterior with larger values of $N$.

7. Janne Ahonen and Jakub Janda shared the Four Hills Tournament (Vierschanzentournee, Keski-Euroopan mäkiviikot) championship in 2006. Both scored a total of 1081.5 points from four competitions. Before the tournament, both took part in four other competitions. Their scores from all eight competitions were

```
ahonen <-c(299.7, 255.2, 281.7, 238.0,   270.9, 262.2, 255.4, 293.0)
janda <-c(238.7, 285.6, 287.1, 252.2,   262.6, 264.7, 263.2, 291.0)
```

Assuming a normal model $\text{N}(\mu_i, \sigma_i^2)$ for both jumpers ($i = 1, 2$) and the uninformative prior $\pi(\mu_i, \sigma_i^2) \propto 1/\sigma_i^2$, a posterior density can be obtained. If both mean and variance are unknown, and the prior is uninformative, can we have realistic estimates of $(\mu_1, \mu_2)$ with these data only? Note: a 95% posterior interval of $\mu_i$ is obtained from $t_{n-1}$ distribution as $\bar{X}_i \pm 1.997 s_i/\sqrt{n}$. Compare the two jumpers using such posterior intervals.