Parametric Statistics-Recitation 6 (Solutions)

Exercise 1.

Suppose that the proportion θ of defective items in a large manufactured lot is known to be either 0.1 or 0.2 , and the prior p.f. of θ is as follows:

$$\xi(0.1) = 0.7 \quad \text{and} \quad \xi(0.2) = 0.3$$

Suppose also that when eight items are selected at random from the lot, it is found that exactly two of them are defective. Determine the posterior p.f. of θ.

Solution.

We know that $f(\mathbf{x}|\theta) = \theta^2(1-\theta)^6$. Therefore,

$$\xi(0.1|\mathbf{x}) = \frac{\xi(0.1)f(\mathbf{x}|0.1)}{\xi(0.1)f(\mathbf{x}|0.1) + \xi(0.2)f(\mathbf{x}|0.2)} = 0.5418$$

It follows that $\xi(0.2|\mathbf{x}) = 1 - \xi(0.1|\mathbf{x}) = 0.4582$.

Exercise 2.

Suppose that the prior distribution of some parameter θ is a gamma distribution for which the mean is 10 and the variance is 5 . Determine the prior p.d.f. of θ.

Solution.

If α and β denote the parameters of the gamma distribution, then we have that

$$\frac{\alpha}{\beta} = 10 \text{ and } \frac{\alpha}{\beta^2} = 5$$

Therefore $\alpha = 20$ and $\beta = 2$. Hence the prior p.d.f of θ is as follows for $\theta > 0$:

$$\xi(\theta) = \frac{2^{20}}{\Gamma(20)}\theta^{19}e^{-2\theta}$$

Exercise 3.

Suppose that the proportion θ of defective items in a large manufactured lot is unknown, and the prior distribution of θ is the uniform distribution on the interval $[0,1]$. When eight items are selected at random from the lot, it is found that exactly three of them are defective. Determine the posterior distribution of θ.

Solution.

Here $f(\mathbf{x}|\theta) = \theta^y(1-\theta)^{8-y}$ where $y = \sum_{i=1}^{8} x_i$ and $x_i = 0$ or 1, therefore $y = 3$. For the posterior $\xi(\theta|\mathbf{x})$ it stands that

$$\xi(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)\xi(\theta) = \theta^3(1-\theta)^5$$

So $\theta|\mathbf{x}$ is a beta distribution with parameters $\alpha = 4$ and $\beta = 6$.

Exercise 4.

Suppose that the number of defects in a 1200 -foot roll of magnetic recording tape has a Poisson distribution for which the value of the mean θ is unknown and that the prior distribution of θ is the gamma distribution with parameters $\alpha = 3$ and $\beta = 1$. When five rolls of this tape are selected at random and inspected, the numbers of defects found on the rolls are 2,2,6, 0, and 3. Determine the posterior distribution of θ.

Solution.

In this case it is known that we have a prior conjugate distribution and so the posterior $\theta|\mathbf{x}$ will be a gamma distribution with parameters $\alpha = 3 + \sum_{i=1}^{5} x_i = 3 + 13 = 16$ and $\beta = 1 + n = 1 + 5 = 6$.

Exercise 5.

Suppose that the time in minutes required to serve a customer at a certain facility has an exponential distribution for which the value of the parameter θ is unknown and that the prior distribution of θ is a gamma distribution for which the mean is 0.2 and the standard deviation is 1 . If the average time required to serve a random sample of 20 customers is observed to be 3.8 minutes, what is the posterior distribution of θ ?

Solution.

Let α and β denote the parameters of the gamma distribution, then we have that

$$\frac{\alpha}{\beta} = 0.2 \text{ and } \frac{\alpha}{\beta^2} = 1$$

Therefore $\alpha = 0.04$ and $\beta = 0.2$. Furthermore, the total time required to serve the sample of 20 customers is $\sum_{i=1}^{20} x_i = 20 \cdot 3.8 = 76$. Again we have a prior conjugate distribution and so the posterior $\theta|\mathbf{x}$ will be a gamma distribution with parameters $\alpha = 0.04 + 20 = 20.04$ and $\beta = 0.2 + \sum_{i=1}^{20} x_i = 76.2$.

Exercise 6.

Let $\xi(\theta)$ be a p.d.f. that is defined as follows for constants $\alpha > 0$ and $\beta > 0$:

$$\xi(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} & \text{for } \theta > 0, \\ 0 & \text{for } \theta \leq 0. \end{cases}$$

A distribution with this p.d.f. is called an inverse gamma distribution.
a. Verify that $\xi(\theta)$ is actually a p.d.f. by verifying that $\int_0^\infty \xi(\theta) d\theta = 1$.
b. Consider the family of probability distributions that can be represented by a p.d.f. $\xi(\theta)$ having the given form for all possible pairs of constants $\alpha > 0$ and $\beta > 0$. Show that this family is a conjugate family of prior distributions for samples from a normal distribution with a known value of the mean μ and an unknown value of the variance θ.

Solution.

(a) Let $y = \frac{1}{\theta}$ and $\theta = \frac{1}{y}$ and $d\theta = -\frac{dy}{y^2}$. Also when θ goes between 0 and ∞ then y goes between ∞ and 0. Therefore

$$\int_0^\infty \xi(\theta) d\theta = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} y^{a-1} e^{-by} dy = 1$$

(b) If an observation X has a normal distribution with a known mean μ and an unknown variance θ, then the p.d.f of X has the form

$$f(x|\theta) \propto \theta^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\theta}}$$

Also, the prior p.d.f of θ has the form

$$\xi(\theta) \propto \theta^{-(\alpha+1)} e^{-\beta/\theta}$$

Therefore, the posterior p.d.f of $\xi(\theta|x)$ has the form

$$\xi(\theta|x) \propto \xi(\theta) f(x|\theta) \propto \theta^{-(\alpha+\frac{3}{2})} exp\left(-\left[\beta + \frac{1}{2}(x-\mu)^2\right] \cdot \frac{1}{\theta}\right)$$

Hence, the posterior p.d.f of $\xi(\theta|x)$ has the same form as $\xi(\theta)$ with parameters $\alpha' = \alpha + \frac{1}{2}$ and $\beta' = \beta + \frac{1}{2}(x-\mu)^2$.

Exercise 7.

Suppose that a random sample of size n is taken from a Poisson distribution for which the value of the mean θ is unknown, and the prior distribution of θ is a gamma distribution for which the mean is μ_0. Show that the mean of the posterior distribution of θ will be a weighted average having the form $\gamma_n \bar{X}_n + (1-\gamma_n)\,\mu_0$, and show that $\gamma_n \to 1$ as $n \to \infty$.

Solution.

Suppose that the parameters of the prior gamma distribution of θ are α and β. Then $\mu_0 = \frac{\alpha}{\beta}$. We know that the posterior distribution of θ is also a gamma distribution with parameters

$$\alpha' = \alpha + \sum_{i=1}^{n} x_i \text{ and } \beta' = \beta + n$$

So the mean of this posterior distribution is

$$\mu = \frac{\alpha + \sum_{i=1}^{n} x_i}{\beta + n} = \frac{\beta}{\beta + n}\mu_0 + \frac{n}{\beta + n}\bar{X}_n$$

Hence, $\gamma_n = \frac{n}{\beta+n}$ and $\gamma_n \to 1$ as $n \to \infty$.

Exercise 8.

Consider again the conditions of Exercise 7, and suppose that the value of θ must be estimated by using the squared error loss function. Show that the Bayes estimators, for $n = 1, 2, \ldots$, form a consistent sequence of estimators of θ.

Solution.

The Bayes estimator is the mean of the posterior distribution of θ, as given in Exercise 7. Since θ is the mean of the Poisson distribution, it follows from the law of large numbers that $\bar{X}_n$ converges to θ in probability as $n \to \infty$. It now follows from Exercise 7 that, since $\gamma_n \to 1$, the Bayes estimators will also converge to θ in probability as $n \to \infty$. Hence the Bayes estimators form a consistent sequence of estimators of θ.

Exercise 9.

It is not known what proportion p of the purchases of a certain brand of breakfast cereal are made by women and what proportion are made by men. In a random sample of 70 purchases of this cereal, it was found that 58 were made by women and 12 were made by men. Find the M.L.E. of p.

Solution.

It is known that the M.L.E in this "Bernoulli" case is $\bar{X}_n = \frac{58}{70}$.

Exercise 10.

Suppose that $X_1, \ldots, X_n$ form a random sample from the Bernoulli distribution with parameter θ, which is unknown, but it is known that θ lies in the open interval $0 < \theta < 1$. Show that the M.L.E. of θ does not exist if every observed value is 0 or if every observed value is 1 .

Solution.

Let y denote the sum of the observations in the sample. Then the likelihood function is $p^y(1-p)^{n-y}$. If $y = 0$, this function is a decreasing function of p. Since $p = 0$ is not a value in the parameter space, there is no M.L.E. Similarly, if $y = n$, then the likelihood function is an increasing function of p. Since $p = 1$ is not a value in the parameter space, there is no M.L.E.

Exercise 11.

Suppose that $X_1, \ldots, X_n$ form a random sample from a Poisson distribution for which the mean θ is unknown, $(\theta > 0)$.
a. Determine the M.L.E. of θ, assuming that at least one of the observed values is different from 0 .
b. Show that the M.L.E. of θ does not exist if every observed value is 0

Solution.

Let y denote the sum of the observed values $x_1, ..., x_n$. Then the likelihood function is

$$f(\mathbf{x}|\theta) = \frac{e^{-n\theta}\theta^y}{\prod_{i=1}^n x_i!}$$

(a) If $y > 0$ and we let $l(\theta) = log f(\mathbf{x}|\theta)$, then

$$\frac{\delta}{\delta\theta} l(\theta) = -n + \frac{y}{\theta}$$

The maximum of $l(\theta)$ will be attained at the value of θ for which this derivative is equal to 0.
So $\theta_{MLE} = \frac{y}{n} = \bar{X}_n$

(b) If $y = 0$, then $f(\mathbf{x}|\theta)$ is a decreasing function of θ. Since $\theta = 0$ is not a value in the parameter space, there is no M.L.E.

Exercise 12.

Suppose that $X_1, \ldots, X_n$ form a random sample from a normal distribution for which the mean μ is known, but the variance σ^2 is unknown. Find the M.L.E. of σ^2.

Solution.

Let $\theta = \sigma^2$. Then the likelihood function is

$$f(\mathbf{x}|\theta) = \frac{1}{(2\pi\theta)^{n/2}} exp\left[-\frac{1}{2\theta}\sum_{i=1}^n (x_i - \mu)^2\right]$$

If we let $l(\theta) = log f(\mathbf{x}|\theta)$, then

$$\frac{\delta}{\delta\theta} l(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n (x_i - \mu)^2$$

The maximum of $l(\theta)$ will be attained at a value of θ for which this derivative is equal to 0. In this way, we find that

$$\theta_{MLE} = \frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2$$