Bayesian Inference

Loukia Meligkotsidou

National and Kapodistrian University of Athens

MSc in Statistics and Operational Research,
Department of Mathematics

Conjugate Analysis for the Normal Model

Let $x = (x_1, \ldots, x_n)$ be an i.i.d sample from the $N(\mu, \tau^{-1})$ distribution with both parameters unknown.

The likelihood of the observations is

$$f(x \mid \mu, \tau) \propto \tau^{n/2} \exp\left\{ -\frac{\tau}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\}.$$

The conjugate prior distribution for (μ, τ) is of the form $f(\mu, \tau) = f(\tau) f(\mu \mid \tau)$ where $f(\tau) \equiv Gamma(a, b)$ and $f(\mu \mid \tau) \equiv N\left(\xi, (c\tau)^{-1}\right)$. That is

$$f(\mu, \tau) \propto \tau^{a-1} \exp\{-\tau b\} \times \tau^{1/2} \exp\left\{ -\frac{\tau c}{2} (\mu - \xi)^2 \right\}.$$

Joint and Conditional Posteriors

The joint posterior distribution of (μ, τ), obtained by Bayes theorem, is

$$
\begin{aligned}
f(\mu, \tau \mid x) &\propto f(\mu, \tau) f(x \mid \mu, \tau) \\
&\propto \tau^{\frac{n+1}{2}+a-1} \exp\left\{-\tau\left[\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2+\frac{c}{2}(\mu-\xi)^2+b\right]\right\}.
\end{aligned}
$$

The conditional posterior distribution of τ given μ is given by

$$
\begin{aligned}
f(\tau \mid x, \mu) &\propto f(\mu, \tau \mid x) \qquad \text{(as a function of } \tau\text{)} \\
&\propto \tau^{\frac{n+1}{2}+a-1} \exp\left\{-\tau\left[\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2+\frac{c}{2}(\mu-\xi)^2+b\right]\right\},
\end{aligned}
$$

which is a Gamma(P, Q) distribution with parameters $P = \frac{n+1}{2} + a$ and $Q = \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{c}{2}(\mu-\xi)^2 + b$.

Joint and Conditional Posteriors

The conditional posterior distribution of μ given τ is

$$\begin{aligned}
f(\mu \mid x, \tau) &\propto f(\mu, \tau \mid x) \qquad \text{(as a function of } \mu) \\
&\propto \exp\left\{-\frac{\tau}{2}\left[\sum_{i=1}^{n}(x_i - \mu)^2 + c(\mu - \xi)^2\right]\right\} \\
&\propto \exp\left\{-\frac{\tau}{2}n\mu^2 + \tau\mu\sum_{i=1}^{n}x_i - \frac{\tau c}{2}\mu^2 + \tau c\mu\xi\right\} \\
&\propto \exp\left\{-\frac{\tau(n+c)}{2}\mu^2 + \tau(n\bar{x} + c\xi)\mu\right\} \\
&\propto \exp\left\{-\frac{\tau(n+c)}{2}[\mu^2 + 2\frac{n\bar{x}+c\xi}{n+c}\mu]\right\},
\end{aligned}$$

which is a Normal(B, D^2) distribution with mean $B = \frac{n\bar{x}+c\xi}{n+c}$ and variance $D^2 = \tau^{-1}(n+c)^{-1}$.

Marginal Posterior of τ

Exact Bayesian inference is based on the marginal posteriors.

$$
\begin{aligned}
f(\tau \mid x) &= \int f(\mu, \tau \mid x) d\mu \\
&\propto \int_{-\infty}^{\infty} \tau^{\frac{n+1}{2}+a-1} \exp\{-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{\tau c}{2}(\mu-\xi)^2 - \tau b\} d\mu \\
&= \tau^{\frac{n}{2}+a-1} \exp\{-\tau[\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \frac{c}{2}\xi^2 + b]\} \\
&\times \int_{-\infty}^{\infty} \tau^{1/2} \exp\{-\frac{\tau(n+c)}{2}\mu^2 + \tau(\sum_{i=1}^{n} x_i + c)\mu\} d\mu \\
&\propto \tau^{\frac{n}{2}+a-1} \exp\{-\tau[\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \frac{c\xi^2}{2} + b - \frac{(n\bar{x}+c\xi)^2}{2(n+c)}]\},
\end{aligned}
$$

This is again a Gamma(P', Q') density, with parameters $P' = \frac{n}{2} + a$ and $Q' = \frac{1}{2}\sum_{i=1}^{n} x_i^2 + \frac{c}{2}\xi^2 + b - \frac{(n\bar{x}+c\xi)^2}{2(n+c)}$.

Marginal Posterior of μ

The marginal posterior of μ is obtained by integrating the joint posterior distribution over τ, i.e.

$$\begin{aligned} f(\mu \mid x) &= \int f(\mu, \tau \mid x) d\tau \\ &\propto \int_0^\infty \tau^{\frac{n+1}{2}+a-1} \exp\{-\tau[\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2 + \frac{c}{2}(\mu-\xi)^2 + b]\} d\tau \\ &\propto \left[\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2 + \frac{c}{2}(\mu-\xi)^2 + b\right]^{-(\frac{n+1}{2}+a)}. \end{aligned}$$

By some calculus manipulation, it can be shown that the normalised version of this formula is a non-standardized (three parameter) Student's-t probability density function.

The non-standardized Student's-t distribution

The standard Student's-t distribution can be generalized to a three parameter location-scale family, introducing a location parameter μ and a scale parameter σ , through the relation $X = \mu + \sigma T$, where $T \sim t(\nu)$.

The resulting non-standardized Student's-t distribution has p.d.f.

$$f(x \mid \nu, \mu, \sigma) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}\sigma} \left[1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}}$$

Here, σ does not correspond to a standard deviation. It simply sets the overall scaling of the distribution. The mean and variance of the non-standardized Student's t distribution are, respectively, $E(X) = \mu$, for $\nu > 1$, and $V(X) = \sigma^2 \frac{\nu}{\nu-2}$ for $\nu > 2$.

Marginal Likelihood

$$
\begin{aligned}
f(x) &= \int_\tau \int_\mu f(x \mid \mu, \tau) f(\mu \mid \tau) f(\tau) d\mu d\tau \\
&= \int_\tau \int_\mu (2\pi)^{-\frac{n+1}{2}} \tau^{\frac{n}{2}} \exp\left\{ -\frac{\tau}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\} \\
&\times \frac{b^a}{\Gamma(a)} \tau^{a-1} \exp\{-\tau b\} (2\pi)^{-1/2} (c\tau)^{1/2} \exp\{-\frac{\tau c}{2} (\mu - \xi)^2\} d\mu d\tau
\end{aligned}
$$

Marginal Likelihood

$$
\begin{aligned}
f(x) &= \int_\tau \int_\mu f(x \mid \mu, \tau) f(\mu \mid \tau) f(\tau) d\mu d\tau \\
&= \int_\tau \int_\mu (2\pi)^{-\frac{n+1}{2}} \tau^{\frac{n}{2}} \exp\left\{ -\frac{\tau}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\} \\
&\times \frac{b^a}{\Gamma(a)} \tau^{a-1} \exp\{-\tau b\} (2\pi)^{-1/2} (c\tau)^{1/2} \exp\{-\frac{\tau c}{2}(\mu - \xi)^2\} d\mu d\tau \\
&= (2\pi)^{-\frac{n}{2}} \frac{b^a}{\Gamma(a)} c^{1/2} \int_\tau \tau^{\frac{n}{2}+a-1} \exp\{-\frac{\tau}{2} \sum x_i^2 - \frac{\tau c}{2} \xi^2 - \tau b\} \\
&\times \left[\int_\mu (2\pi)^{-\frac{1}{2}} \tau^{\frac{1}{2}} \exp\{-\frac{\tau(n+c)}{2} \mu^2 + \tau(\sum_{i=1}^{n} x_i + c)\mu\} d\mu \right] d\tau
\end{aligned}
$$

Marginal Likelihood

Therefore,

$$
\begin{aligned}
f(x) &= (2\pi)^{-\frac{n}{2}} \frac{b^a}{\Gamma(a)} c^{\frac{1}{2}} (n+c)^{-\frac{1}{2}} \\
&\times \int_\tau \tau^{\frac{n}{2}+a-1} \exp\{-\tau[\frac{1}{2}\sum_{i=1}^n x_i^2 + \frac{c\xi^2}{2} + b\frac{(n\bar{x}+c\xi)^2}{2(n+c)}]\} d\tau
\end{aligned}
$$

Marginal Likelihood

Therefore,

$$
\begin{aligned}
f(x) &= (2\pi)^{-\frac{n}{2}} \frac{b^a}{\Gamma(a)} c^{\frac{1}{2}} (n+c)^{-\frac{1}{2}} \\
&\times \int_\tau \tau^{\frac{n}{2}+a-1} \exp\{-\tau[\frac{1}{2}\sum_{i=1}^n x_i^2 + \frac{c\xi^2}{2} + b\frac{(n\bar{x}+c\xi)^2}{2(n+c)}]\} d\tau \\
&= (2\pi)^{-\frac{n}{2}} \frac{b^a}{\Gamma(a)} c^{\frac{1}{2}} (n+c)^{-\frac{1}{2}} \\
&\times \Gamma(\frac{n}{2}+a) \left[\frac{1}{2}\sum_{i=1}^n x_i^2 + \frac{c\xi^2}{2} + b - \frac{(n\bar{x}+c\xi)^2}{2(n+c)}\right]^{-(\frac{n}{2}+a)}.
\end{aligned}
$$

Exercise 1.1

The rock strata A and B are difficult to distinguish in the field. Through careful laboratory studies it has been determined that the only characteristic which might be useful in aiding discrimination is the presence or absence of a particular brachipod fossil. The probabilities of fossil presence are found to be as follows.

Stratum	Fossil present	Fossil absent
A	0.9	0.1
B	0.2	0.8

It is also known that rock type A occurs about four times as often as type B. If a sample is taken, and the fossil found to be present, calculate the posterior distribution of rock types.

If the geologist always classifies as A when the fossil is found to be present, and classifies as B when it is absent, what is the probability she will be correct in a future classification?

Solution

Denote
A: rock stratum A, B: rock stratum B and F: fossil is present.

We are given that $\Pr(F \mid A) = 0.9$, $\Pr(F^c \mid A) = 0.1$, $\Pr(F \mid B) = 0.2$, $\Pr(F^c \mid B) = 0.8$ and $\Pr(A) = 4\Pr(B)$.

To obtain the posterior distribution of rock types, after finding the fossil in the sample, we need to calculate $\Pr(A \mid F)$ and $\Pr(B \mid F)$.

Solution

Denote
A: rock stratum A, B: rock stratum B and F: fossil is present.

We are given that $\Pr(F \mid A) = 0.9$, $\Pr(F^c \mid A) = 0.1$, $\Pr(F \mid B) = 0.2$, $\Pr(F^c \mid B) = 0.8$ and $\Pr(A) = 4\Pr(B)$.

To obtain the posterior distribution of rock types, after finding the fossil in the sample, we need to calculate $\Pr(A \mid F)$ and $\Pr(B \mid F)$.

$\Pr(A) + \Pr(B) = 1 \Rightarrow 4\Pr(B) + \Pr(B) = 1 \Rightarrow \Pr(B) = 0.2.$

Solution

Denote
A: rock stratum A, B: rock stratum B and F: fossil is present.

We are given that $\Pr(F \mid A) = 0.9$, $\Pr(F^c \mid A) = 0.1$, $\Pr(F \mid B) = 0.2$, $\Pr(F^c \mid B) = 0.8$ and $\Pr(A) = 4\Pr(B)$.

To obtain the posterior distribution of rock types, after finding the fossil in the sample, we need to calculate $\Pr(A \mid F)$ and $\Pr(B \mid F)$.

$\Pr(A) + \Pr(B) = 1 \Rightarrow 4\Pr(B) + \Pr(B) = 1 \Rightarrow \Pr(B) = 0.2$.

Prior:	$\Pr(A) = 0.8$	$\Pr(B) = 0.2$
Likelihood:	$\Pr(F \mid A) = 0.9$	$\Pr(F \mid B) = 0.2$
Prior x likelihood:		
	$\Pr(A)\Pr(F \mid A) = 0.72$	$\Pr(B)\Pr(F \mid B) = 0.04$

Solution

Law of total probability:
$\Pr(F) = \Pr(A)\Pr(F \mid A) + \Pr(B)\Pr(F \mid B) = 0.76$

Solution

Law of total probability:
$\Pr(F) = \Pr(A)\Pr(F \mid A) + \Pr(B)\Pr(F \mid B) = 0.76$

Posterior: $\Pr(A \mid F) = \frac{\Pr(A)\Pr(F|A)}{\Pr(F)} = \frac{72}{76}$
and $\Pr(B \mid F) = \frac{\Pr(B)\Pr(F|B)}{\Pr(F)} = \frac{4}{76}$.

Solution

Law of total probability:
$\Pr(F) = \Pr(A)\Pr(F \mid A) + \Pr(B)\Pr(F \mid B) = 0.76$

Posterior: $\Pr(A \mid F) = \frac{\Pr(A)\Pr(F|A)}{\Pr(F)} = \frac{72}{76}$
and $\Pr(B \mid F) = \frac{\Pr(B)\Pr(F|B)}{\Pr(F)} = \frac{4}{76}$.

Probability of correct classification:

$$\begin{aligned}\Pr(\text{correct}) &= \Pr(A, F) + \Pr(B, F^c) \\ &= \Pr(A)\Pr(F \mid A) + \Pr(B)\Pr(F^c \mid B) \\ &= 0.72 + 0.16 = 0.88.\end{aligned}$$

Exercise 1.4

A seed collector who has acquired a small number of seeds from a plant, has a prior belief that the probability θ of germination of each seed is uniform over the range $0 \leq \theta \leq 1$. She experiments by sowing two seeds and finds that they both germinate.

i. Write down the likelihood function for θ deriving from this observation, and obtain the collector's posterior distribution of θ

ii. Compute the posterior probability that θ is less than one half.

Exercise 1.4

A seed collector who has acquired a small number of seeds from a plant, has a prior belief that the probability θ of germination of each seed is uniform over the range $0 \leq \theta \leq 1$. She experiments by sowing two seeds and finds that they both germinate.

i. Write down the likelihood function for θ deriving from this observation, and obtain the collector's posterior distribution of θ

ii. Compute the posterior probability that θ is less than one half.

Solution.

(i.) The likelihood function is

$$f(X = 2 \mid \theta) = \binom{2}{2} \theta^2 (1 - \theta)^{2-2} = \theta^2$$

Exercise 1.4-Solution

Using the uniform prior $f(\theta) = 1$, $0 \leq \theta \leq 1$ we obtain the posterior distribution $f(\theta \mid x) = \frac{f(\theta) f(x|\theta)}{f(x)} = 3\theta^2$.
Note that $f(x) = \int_0^1 \theta^2 d\theta = \frac{1}{3}$.

Exercise 1.4-Solution

Using the uniform prior $f(\theta) = 1$, $0 \leq \theta \leq 1$ we obtain the posterior distribution $f(\theta \mid x) = \frac{f(\theta) f(x|\theta)}{f(x)} = 3\theta^2$.
Note that $f(x) = \int_0^1 \theta^2 d\theta = \frac{1}{3}$.

(ii.) The prior probability is $P(\theta < 1/2) = \int_0^{0.5} 1 d\theta = 0.5$.
The posterior probability is
$P(\theta < 1/2 \mid X = 2) = \int_0^{0.5} 3\theta^2 d\theta = 0.8$.

Exercise 1.5

A posterior distribution is calculated up to a normalizing constant as

$$f(\theta \mid x) \propto \theta^{-3},$$

for $\theta > 1$. Calculate the normalizing constant of this posterior and the posterior probability of $\theta < 2$.

Exercise 1.5

A posterior distribution is calculated up to a normalizing constant as

$$f(\theta \mid x) \propto \theta^{-3},$$

for $\theta > 1$. Calculate the normalizing constant of this posterior and the posterior probability of $\theta < 2$.

Solution.

Since $\int f(\theta \mid x)\, d\theta = 1$, thus,

$\int_1^\infty c \frac{1}{\theta^3} d\theta = 1 \Leftrightarrow \left[\frac{c\theta^{-2}}{-2}\right]_1^\infty = 1 \Leftrightarrow c = 2.$

Exercise 1.5

A posterior distribution is calculated up to a normalizing constant as

$$f(\theta \mid x) \propto \theta^{-3},$$

for $\theta > 1$. Calculate the normalizing constant of this posterior and the posterior probability of $\theta < 2$.

Solution.

Since $\int f(\theta \mid x)\, d\theta = 1$, thus,

$\int_1^\infty c \frac{1}{\theta^3} d\theta = 1 \Leftrightarrow \left[\frac{c\theta^{-2}}{-2}\right]_1^\infty = 1 \Leftrightarrow c = 2.$

The posterior probability is:

$P(\theta < 2 \mid x) = \int_1^2 f(\theta \mid x)\, d\theta = \int_1^2 2\theta^{-3} d\theta = 3/4.$

Exercise

a. Show that Jeffrey's prior is consistent across 1-1 parameter transformations.

b. Suppose that $X \mid \theta \sim \textit{Binomial}\,(n, \theta)$. Find Jeffrey's prior for the corresponding posterior distribution of θ

c. Now suppose that $\phi = \frac{1}{\theta}$. What is the Jeffrey's prior for ϕ?

Exercise

a. Show that Jeffrey's prior is consistent across 1-1 parameter transformations.

b. Suppose that $X \mid \theta \sim \textit{Binomial}\,(n, \theta)$. Find Jeffrey's prior for the corresponding posterior distribution of θ

c. Now suppose that $\phi = \frac{1}{\theta}$. What is the Jeffrey's prior for ϕ?

Solution.

a. We need to show that $J_\Phi \phi = J_\Theta \theta \left| \frac{d\theta}{d\phi} \right|$. It is sufficient to show that $I_\Phi \phi = I_\Theta \theta \left| \frac{d\theta}{d\phi} \right|^2$

We have that, $\frac{d \log f(x|\Phi)}{d\phi} = \frac{d \log f(x|\theta(\phi))}{d\theta} \frac{d\theta(\phi)}{d\phi}$. Hence,

$$I\,(\phi) = E\left\{ \left(\frac{d \log f(x|\phi)}{d\phi} \right)^2 \right\} = E\left\{ \left(\frac{d \log f(x|\theta(\phi))}{d\theta} \frac{d\theta(\phi)}{d\phi} \right)^2 \right\} =$$

$$E\left\{ \left(\frac{d \log f(x|\theta)}{d\theta} \right)^2 \right\} \left(\frac{d\theta}{d\phi} \right)^2 = I(\theta) \left| \frac{d\theta}{d\phi} \right|^2$$

Solution

b.

$$f(x \mid \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

and $L(\theta) = \log f(x \mid \theta) = x \log(\theta) + (n-x)\log(1-\theta) + c$

$$\frac{dL(\theta)}{d\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta}$$

$$\frac{d^2L(\theta)}{d\theta^2} = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$$

Since $E(x) = n\theta \Rightarrow I(\theta) = -\frac{n\theta}{\theta^2} - \frac{n-n\theta}{(1-\theta)^2} = n\left(\frac{1-\theta+\theta}{\theta(1-\theta)}\right) =$
$n\theta^{-1}\left(1-\theta^{-1}\right) \Rightarrow J(\theta) \propto \theta^{-1/2}(1-\theta)^{-1/2}$

Solution

b.

$$f(x \mid \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

and $L(\theta) = \log f(x \mid \theta) = x \log(\theta) + (n-x)\log(1-\theta) + c$

$$\frac{dL(\theta)}{d\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta}$$
$$\frac{d^2L(\theta)}{d\theta^2} = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$$

Since $E(x) = n\theta \Rightarrow I(\theta) = -\frac{n\theta}{\theta^2} - \frac{n-n\theta}{(1-\theta)^2} = n\left(\frac{1-\theta+\theta}{\theta(1-\theta)}\right) =$ $n\theta^{-1}\left(1-\theta^{-1}\right) \Rightarrow J(\theta) \propto \theta^{-1/2}(1-\theta)^{-1/2}$

c.

$$f(x \mid \phi) = \binom{n}{x} \phi^{-x} \left(1 - \frac{1}{\phi}\right)^{n-x}$$

with $E(X) = \frac{n}{\phi}$

Solution

$L(\phi) = -x \log(\phi) + (n-x) \log\left(\frac{\phi-1}{\phi}\right) + c =$
$-n \log(\phi) + (n-x) \log(\phi - 1) + c$

$$\frac{dL(\phi)}{d\phi} = \frac{x-x}{\phi-1} - \frac{n}{\phi}$$

$$\frac{d^2L(\phi)}{d\phi^2} = -\frac{n-x}{(\phi-1)^2} + \frac{n}{\phi^2}$$

$$I(\phi) = -E\left\{\frac{d^2L(\phi)}{d\phi^2}\right\} = \frac{n-n/\phi}{(\phi-1)^2} - \frac{n}{\phi^2} = \frac{n\phi - n(\phi-1)}{\phi^2(\phi-1)} \propto \left(\phi^3 - \phi^2\right)^{-1}$$

$$\Rightarrow J(\phi) \propto \left(\phi^3 - \phi^2\right)^{-1/2}$$

Solution

$L(\phi) = -x \log(\phi) + (n-x) \log\left(\frac{\phi-1}{\phi}\right) + c =$
$-n \log(\phi) + (n-x) \log(\phi - 1) + c$

$$\frac{dL(\phi)}{d\phi} = \frac{x-x}{\phi-1} - \frac{n}{\phi}$$

$$\frac{d^2L(\phi)}{d\phi^2} = -\frac{n-x}{(\phi-1)^2} + \frac{n}{\phi^2}$$

$$I(\phi) = -E\left\{\frac{d^2L(\phi)}{d\phi^2}\right\} = \frac{n-n/\phi}{(\phi-1)^2} - \frac{n}{\phi^2} = \frac{n\phi - n(\phi-1)}{\phi^2(\phi-1)} \propto \left(\phi^3 - \phi^2\right)^{-1}$$

$$\Rightarrow J(\phi) \propto \left(\phi^3 - \phi^2\right)^{-1/2}$$

- Using the result from a. we can find the Jeffrey's prior for ϕ as:

$J(\phi) = J(\theta)\left|\frac{d\theta}{d\phi}\right| = \frac{1}{B(1/2,1/2)}\theta^{-1/2}(1-\theta)^{-1/2}\frac{1}{\phi^2} \propto$
$\theta^{-1/2}(1-\theta)^{-1/2}\frac{1}{\phi^2} = \phi^{1/2}\left(1-\frac{1}{\phi}\right)^{-1/2}\frac{1}{\phi^2} = \phi^{-3/2}\frac{(\phi-1)^{-1/2}}{\phi^{-1/2}} =$
$\phi^{-1}(\phi-1)^{-1/2} = \left(\phi^3-\phi^2\right)^{-1/2}$

Exercise 2.1

In each of the following cases, derive the posterior distribution:

a. $x_1, x_2, \ldots x_n$ are a random sample from the distribution with probability function

$$f(x \mid \theta)\, \theta^{x-1}(1-\theta)\,;\ x = 1, 2, \ldots$$

with the $Beta\,(p, q)$ prior distribution

$$f(\theta) = \frac{\theta^{p-1}(1-\theta)^{q-1}}{B(p,q)},\ 0 \leq \theta \leq 1.$$

b. $x_1, x_2, \ldots x_n$ are a random sample from the distribution with probability density function

$$f(x \mid \theta) = \frac{e^{-\theta}\theta^{x}}{x!},\ x = 0, 1, \ldots$$

with the prior distribution

$$f(\theta) = e^{-\theta},\ \theta \geq 0.$$

Exercise 2.1-Solution

a. We know that $f(\theta \mid x) \propto f(\theta) f(x \mid \theta)$.
The likelihood function is:
$f(x \mid \theta) = \prod_{i=1}^{n} \theta^{x_i - 1}(1-\theta) = (1-\theta)^n \theta^{\sum_{i=1}^{n} x_i - n}$.

Exercise 2.1-Solution

a. We know that $f(\theta \mid x) \propto f(\theta) f(x \mid \theta)$.
The likelihood function is:
$f(x \mid \theta) = \prod_{i=1}^{n} \theta^{x_i - 1} (1 - \theta) = (1 - \theta)^n \theta^{\sum_{i=1}^{n} x_i - n}$.

The posterior distribution is:

$$\begin{aligned} f(\theta|x) &\propto \theta^{p-1} (1-\theta)^{q-1} (1-\theta)^n \theta^{-n} \theta^{\sum_{i=1}^{n} x_i} \\ &\propto \theta^{\sum_{i=1}^{n} x_i + p - n - 1} (1-\theta)^{q+n-1} \equiv Beta(P, Q) \end{aligned}$$

Exercise 2.1-Solution

a. We know that $f(\theta \mid x) \propto f(\theta) f(x \mid \theta)$.
The likelihood function is:
$f(x \mid \theta) = \prod_{i=1}^{n} \theta^{x_i - 1}(1-\theta) = (1-\theta)^n \theta^{\sum_{i=1}^{n} x_i - n}$.

The posterior distribution is:

$$
\begin{aligned}
f(\theta|x) &\propto \theta^{p-1}(1-\theta)^{q-1}(1-\theta)^n \theta^{-n}\theta^{\sum_{i=1}^{n} x_i} \\
&\propto \theta^{\sum_{i=1}^{n} x_i + p - n - 1}(1-\theta)^{q+n-1} \equiv Beta(P, Q)
\end{aligned}
$$

b. Likewise,
$f(x \mid \theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$ and
$f(\theta|x) \propto e^{-n\theta}\theta^{\sum_{i=1}^{n} xi} e^{-\theta} = e^{-(n+1)\theta}\theta^{\sum_{i=1}^{n} x_i} \equiv Gamma(p, q)$.

Note: $Y \sim Beta(p, q) \Leftrightarrow f(y) = \frac{1}{B(p,q)} y^{p-1}(1-\theta)^{q-1}$ and
$Y \sim Gamma(p, q) \Leftrightarrow f(y) = \frac{p^q}{\Gamma(p)} e^{-qy} y^{p-1}$.

Exercise 2.2

The proportion, θ, of defective items in a large shipment is unknown, but the expert assessment assigns θ of the $Beta(2, 200)$ prior distribution. If 100 items are selected at random from the shipment, and 3 are found to be defective, what is the posterior distribution of θ?

If another statistician, having observed the 3 defectives, calculated the posterior distribution as being a beta distribution with mean $4/102$ and variance 0.0003658, then what prior distribution had she used?

Exercise 2.2-Solution

Prior distribution: $f(\theta) \propto \theta (1-\theta)^{199}, \; 0 \leq \theta \leq 1.$

Exercise 2.2-Solution

Prior distribution: $f(\theta) \propto \theta (1-\theta)^{199}, \ 0 \leq \theta \leq 1.$

Likelihood:
$f(x \mid \theta) = \binom{100}{3} \theta^3 (1-\theta)^{97} \propto \theta^3 (1-\theta)^{97}, \ 0 \leq \theta \leq 1.$

Exercise 2.2-Solution

Prior distribution: $f(\theta) \propto \theta (1-\theta)^{199}, \ 0 \leq \theta \leq 1.$

Likelihood:
$f(x \mid \theta) = \binom{100}{3} \theta^3 (1-\theta)^{97} \propto \theta^3 (1-\theta)^{97}, \ 0 \leq \theta \leq 1.$

Posterior: $f(\theta \mid x) \propto \theta^3 (1-\theta)^{97} \theta (1-\theta)^{199} = \theta^4 (1-\theta)^{296} \equiv$ $Beta(5, 297)$.

Exercise 2.2-Solution

Prior distribution: $f(\theta) \propto \theta (1-\theta)^{199}$, $0 \leq \theta \leq 1$.

Likelihood:
$f(x \mid \theta) = \binom{100}{3} \theta^3 (1-\theta)^{97} \propto \theta^3 (1-\theta)^{97}$, $0 \leq \theta \leq 1$.

Posterior: $f(\theta \mid x) \propto \theta^3 (1-\theta)^{97} \theta (1-\theta)^{199} = \theta^4 (1-\theta)^{296} \equiv Beta(5, 297)$.

Let $\theta \mid x \sim Beta(P, Q) \Leftrightarrow E(\theta \mid x) = \frac{P}{P+Q} = 4/102$ and $V(\theta \mid x) = \frac{PQ}{(P+Q)^2(P+Q+1)} = 0.0003658$

After some straight forward algebra: $P = 4$ and $Q = 98$.
Also, $P = p + x \Longrightarrow 4 = p + 3 \Longrightarrow p = 1$ and
$Q = q + n - x \Longrightarrow 98 = q + 3 - 100 \Longrightarrow q = 1$.
$\Longrightarrow$ The statistician used $Beta(1, 1) \equiv U(0, 1)$ prior distribution.

Exercise 2.3

The diameter of a component from a long production run varies according to a *Normal* $(\theta, 1)$ distribution. An engineer specifies that the prior distribution for θ is *Normal* $(10, 0.25)$. In one production run 12 components are sampled and found to have a sample mean diameter of $31/3$. Use this information to find the posterior distribution of mean component diameter. Hence calculate the probability that this is more than 10 units.

Exercise 2.3

The diameter of a component from a long production run varies according to a *Normal* $(\theta, 1)$ distribution. An engineer specifies that the prior distribution for θ is *Normal* $(10, 0.25)$. In one production run 12 components are sampled and found to have a sample mean diameter of $31/3$. Use this information to find the posterior distribution of mean component diameter. Hence calculate the probability that this is more than 10 units.

Solution.

For known σ: $f(x_i \mid \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(x_i-\theta)^2}{2\sigma^2}\} \propto$ $\exp\{-\frac{1}{2\sigma^2}x_i^2 + \frac{x_i\theta}{\sigma^2} - \frac{\theta^2}{2\sigma^2}\} \propto \exp\{-\frac{1}{2\sigma^2}\theta^2 + \frac{1}{\sigma^2}x_i\theta\}$

Likelihood: $f(x \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) \propto$ $\prod_{i=1}^{n} \exp\{-\frac{\theta^2}{2\sigma^2} + \frac{1}{\sigma^2}x_i\theta\} = \exp\{-\frac{n}{2\sigma^2}\theta^2 + \frac{\theta}{\sigma^2}\sum_{i=1}^{n} x_i\}$

Exercise 2.3-Solution

Prior distribution:

$f(\theta) \propto \exp\{-\frac{1}{2d^2}(\theta - b)^2\} \propto \exp\{-\frac{1}{2d^2}\theta^2 + \frac{1}{d^2}b\theta\}$

Exercise 2.3-Solution

Prior distribution:

$f(\theta) \propto \exp\{-\frac{1}{2d^2}(\theta - b)^2\} \propto \exp\{-\frac{1}{2d^2}\theta^2 + \frac{1}{d^2}b\theta\}$

Posterior distribution:

$f(\theta \mid x) \propto f(\theta) f(x \mid \theta) \propto$
$\exp\left\{\left(-\frac{n}{2\sigma^2} - \frac{1}{2d^2}\right)\theta^2 + \left(\frac{1}{d^2}b + \frac{1}{\sigma^2}\sum_{i=1}^{n}\right)\theta\right\} =$
$\exp\left\{-\frac{1}{2D^2}\theta^2 + \frac{B}{D^2}\theta\right\}$

Exercise 2.3-Solution

Prior distribution:

$f(\theta) \propto \exp\{-\frac{1}{2d^2}(\theta - b)^2\} \propto \exp\{-\frac{1}{2d^2}\theta^2 + \frac{1}{d^2}b\theta\}$

Posterior distribution:

$f(\theta \mid x) \propto f(\theta) f(x \mid \theta) \propto$
$\exp\left\{\left(-\frac{n}{2\sigma^2} - \frac{1}{2d^2}\right)\theta^2 + \left(\frac{1}{d^2}b + \frac{1}{\sigma^2}\sum_{i=1}^{n}\right)\theta\right\} =$
$\exp\left\{-\frac{1}{2D^2}\theta^2 + \frac{B}{D^2}\theta\right\}$

In our problem:

$d^2 = 0.25, b = 10, \frac{\sum_{i=1}^{n} x_i}{n} = \frac{31}{3}$
Substituting, $B = 10.25$ and $D = 16^{-1} = 0.0625$

Calculate $P(\theta > 10) = \int_{10}^{\infty} f(\theta \mid x)\, d\theta.$

Exercise 2.4

The number of defects in a single roll of magnetic tape has a *Poisson* (θ) distribution. The prior distribution for θ is $\Gamma(3,1)$. When 5 rolls of this tape are selected at random, the number of defects found on each are 2,2,6,0 and 3 respectively. Determine the posterior distribution of θ.

Solution.

Prior:

$f(\theta) = \frac{q^p}{\Gamma(p)} \theta^{p-1} e^{-q\theta}$

Exercise 2.4

The number of defects in a single roll of magnetic tape has a *Poisson* (θ) distribution. The prior distribution for θ is $\Gamma(3, 1)$. When 5 rolls of this tape are selected at random, the number of defects found on each are 2,2,6,0 and 3 respectively. Determine the posterior distribution of θ.

Solution.
Prior:
$f(\theta) = \frac{q^p}{\Gamma(p)} \theta^{p-1} e^{-q\theta}$
Likelihood:
$f(x \mid \theta) = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!}$

Exercise 2.4

The number of defects in a single roll of magnetic tape has a *Poisson* (θ) distribution. The prior distribution for θ is $\Gamma(3,1)$. When 5 rolls of this tape are selected at random, the number of defects found on each are 2,2,6,0 and 3 respectively. Determine the posterior distribution of θ.

Solution.
Prior:
$f(\theta) = \frac{q^p}{\Gamma(p)} \theta^{p-1} e^{-q\theta}$
Likelihood:
$f(x \mid \theta) = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!}$
Posterior:
$f(\theta \mid x) \propto \theta^{p-1} e^{-q\theta} e^{-n\theta} \theta^{\sum_{i=1}^{n} x_i} = e^{-(q+n)\theta} \theta^{\sum_{i=1}^{n} x_i + p - 1} \equiv$
$Gamma\left(p + \sum_{i=1}^{n} x_i, q + n\right) = Gamma(16, 6)$

Exercise 2.8

(i) Observations $y_1, y_2, \ldots, y_n$ are obtained from independent random variables which are normally distributed, each with the same (known) variance σ^2 but with respective means $x_1\theta, x_2\theta, \ldots, x_n\theta$. The values of $x_1, x_2, \ldots, x_n$ are known but θ is unknown. Show that the likelihood, given a single observation y_i is of the form

$$f(y_i \mid \theta) \propto \exp\left(-\frac{1}{2\sigma^2}x_i^2\theta^2 + \frac{1}{\sigma^2}y_i x_i \theta\right)$$

(ii) Given the prior distribution for the unknown coefficient θ may be described as normal with mean b and variance σ^2/α^2, show that the posterior distribution of θ is proportional to

$$\exp\left\{-\frac{1}{2}\left[\left(a^2 + \sum_{i=1}^{n} x_i^2\right)/\sigma^2\right]\theta^2 + \left[\left(a^2 b + \sum_{i=1}^{n} y_i x_i\right)/\sigma^2\right]\theta\right\}$$

Exercise 2.8

(iii) Use this to write down the posterior mean of θ. Show that it may be written as

$$\hat{\theta} = wb + (1-w)\frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}$$

and obtain an expression for w.

Exercise 2.8

(iii) Use this to write down the posterior mean of θ. Show that it may be written as

$$\hat{\theta} = wb + (1-w)\frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}$$

and obtain an expression for w.

Solution.

(i) $f(y_i \mid \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - x_i\theta)^2\right\} \propto$
$\exp\left\{-\frac{1}{2\sigma^2}\left(y_i^2 + x_i^2\theta^2 - 2x_i y_i\theta\right)\right\} \propto \exp\left\{-\frac{1}{2\sigma^2}\left(x_i^2\theta^2 - 2x_i y_i\theta\right)\right\}$

(ii)Prior:

$f(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{a^2}{2\sigma^2}(\theta - b)^2\right\} \propto \exp\left\{-\frac{a^2}{2\sigma^2}\theta^2 + \frac{a^2}{\sigma^2}b\theta\right\}$

Exercise 2.8-Solution

Likelihood:

$f(y \mid \theta) = \prod_{i=1}^{n} f(y_i \mid \theta) \propto \prod_{i=1}^{n} \exp\left\{-\frac{1}{2\sigma^2} x_i^2 \theta^2 + \frac{1}{\sigma^2} y_i x_i \theta\right\} =$
$\exp\left\{-\frac{\theta^2}{2\sigma^2} \sum_{i=1}^{n} x_i^2 + \frac{\theta}{\sigma} \sum_{i=1}^{n} x_i y_i\right\}$

Posterior:

$f(\theta \mid y) \propto \exp\left\{-\frac{1}{2}\left(\frac{a^2}{\sigma^2} + \frac{\sum_{i=1}^{n} x_i^2}{\sigma^2}\right)\theta^2 + \left(\frac{a^2}{\sigma^2} b + \frac{\sum_{i=1}^{n} x_i y_i}{\sigma^2}\right)\theta\right\} \equiv$
$N(B, D)$ where $D = \left(\frac{a^2}{\sigma^2} + \frac{\sum_{i=1}^{n} x_i^2}{\sigma^2}\right)^{-1}$ and
$B = \left(\frac{a^2}{\sigma^2} + \frac{\sum_{i=1}^{n} x_i^2}{\sigma^2}\right)^{-1}\left(\frac{a^2}{\sigma^2} b + \frac{\sum_{i=1}^{n} x_i y_i}{\sigma^2}\right) = \frac{a^2 b + \sum_{i=1}^{n} x_i y_i}{a^2 + \sum_{i=1}^{n} x_i^2} =$
$wb + (1-w)\frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}$
where $w = \frac{a^2}{a^2 + \sum_{i=1}^{n} x_i^2}$

Exercise 3.1

Which of the following densities belong to the exponential family

$$f_1(x \mid \theta) = \theta 2^{\theta} x^{-(\theta+1)} \quad \text{for} \quad x > 2$$

$$f_2(x \mid x) = \theta x^{\theta-1} \exp\{-x^{\theta}\} \quad \text{for} \quad x > 0$$

In each case calculate the conjugate prior if the density belongs to the exponential family.

Solution.
A density belongs to Exponential Family of distributions if it can be written in the form of $f(x \mid \theta) = h(x)\, g(\theta) \exp\{t(x)c(\theta)\}$

Exercise 3.1

Which of the following densities belong to the exponential family

$$f_1(x \mid \theta) = \theta 2^\theta x^{-(\theta+1)} \quad \text{for} \quad x > 2$$

$$f_2(x \mid x) = \theta x^{\theta-1} \exp\{-x^\theta\} \quad \text{for} \quad x > 0$$

In each case calculate the conjugate prior if the density belongs to the exponential family.

Solution.

A density belongs to Exponential Family of distributions if it can be written in the form of $f(x \mid \theta) = h(x)\, g(\theta) \exp\{t(x)c(\theta)\}$

$f_1(x \mid \theta) = \theta 2^\theta \exp\{-(\theta+1)\log x\}$

Exponential family with:

$h(x) = 1$

$g(\theta) = \theta 2^\theta$

$t(x) = \log x$

$c(\theta) = -(\theta+1)$

Exercise 3.1-Solution

In this case, we choose prior of the form
$f(\theta) \propto (g(\theta))^d \exp\{bc(\theta)\}$, hence
$f(\theta) \propto \theta^d 2^{d\theta} \exp\{-(\theta+1)b\} = \theta^d \exp\{d\theta \log 2 - (\theta+1)b\} =$
$\theta^d \exp\{(d\theta \log 2 - b)\theta\} \equiv Gamma(\alpha, \beta)$
with $\alpha = d+1$ and $\beta = b - d\log 2$

Exercise 3.1-Solution

In this case, we choose prior of the form
$f(\theta) \propto (g(\theta))^d \exp\{bc(\theta)\}$, hence
$f(\theta) \propto \theta^d 2^{d\theta} \exp\{-(\theta+1)b\} = \theta^d \exp\{d\theta \log 2 - (\theta+1)b\} =$
$\theta^d \exp\{(d\theta \log 2 - b)\theta\} \equiv Gamma(\alpha, \beta)$
with $\alpha = d+1$ and $\beta = b - d\log 2$

$f_2(x \mid x) = \theta \exp\left\{(\theta-1)\log x - x^{\theta}\right\}$
$\Downarrow$
Does not belong to the Exponential family.

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f(x \mid \theta) = (1-\theta)^{x-1}\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f\left(x \mid \theta\right) = \left(1-\theta\right)^{x-1}\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Solution.
$J\left(\theta\right) \propto \left|I(\theta)\right|^{1/2}$,
where $I\left(\theta\right) = -E\left(\frac{d^2 L(\theta)}{d\theta^2}\right) = E\left\{\left(\frac{dL(\theta)}{d\theta}\right)^2\right\}$

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f(x \mid \theta) = (1-\theta)^{x-1}\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Solution.
$J(\theta) \propto |I(\theta)|^{1/2}$,
where $I(\theta) = -E\left(\frac{d^2L(\theta)}{d\theta^2}\right) = E\left\{\left(\frac{dL(\theta)}{d\theta}\right)^2\right\}$ and $L(\theta) =$
$\log f(x \mid \theta) = \log\left[(1-\theta)^{x-1}\theta\right] = (x-1)\log(1-\theta) + \log\theta$

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f(x \mid \theta) = (1-\theta)^{x-1}\,\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Solution.
$J(\theta) \propto |I(\theta)|^{1/2}$,
where $I(\theta) = -E\left(\frac{d^2L(\theta)}{d\theta^2}\right) = E\left\{\left(\frac{dL(\theta)}{d\theta}\right)^2\right\}$ and $L(\theta) =$
$\log f(x \mid \theta) = \log\left[(1-\theta)^{x-1}\,\theta\right] = (x-1)\log(1-\theta) + \log\theta$
$\frac{dL(\theta)}{d\theta} = \frac{1}{\theta} - \frac{(x-1)}{(1-\theta)}$

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f(x \mid \theta) = (1-\theta)^{x-1}\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Solution.

$J(\theta) \propto |I(\theta)|^{1/2}$,

where $I(\theta) = -E\left(\frac{d^2 L(\theta)}{d\theta^2}\right) = E\left\{\left(\frac{dL(\theta)}{d\theta}\right)^2\right\}$ and $L(\theta) =$

$\log f(x \mid \theta) = \log\left[(1-\theta)^{x-1}\theta\right] = (x-1)\log(1-\theta) + \log\theta$

$\frac{dL(\theta)}{d\theta} = \frac{1}{\theta} - \frac{(x-1)}{(1-\theta)}$

$\frac{d^2 L(\theta)}{d\theta^2} = -\frac{1}{\theta^2} - \frac{(x-1)}{(1-\theta)^2}$

Exercise 3.2

Find the Jeffreys prior for θ in the geometric model:

$$f(x \mid \theta) = (1-\theta)^{x-1}\,\theta \quad x = 1, 2, \ldots$$

Note: $E(X) = 1/\theta$.

Solution.

$J(\theta) \propto |I(\theta)|^{1/2}$,

where $I(\theta) = -E\left(\frac{d^2 L(\theta)}{d\theta^2}\right) = E\left\{\left(\frac{dL(\theta)}{d\theta}\right)^2\right\}$ and $L(\theta) =$

$\log f(x \mid \theta) = \log\left[(1-\theta)^{x-1}\,\theta\right] = (x-1)\log(1-\theta) + \log\theta$

$\frac{dL(\theta)}{d\theta} = \frac{1}{\theta} - \frac{(x-1)}{(1-\theta)}$

$\frac{d^2 L(\theta)}{d\theta^2} = -\frac{1}{\theta^2} - \frac{(x-1)}{(1-\theta)^2}$

$I(\theta) = E\left(\frac{1}{\theta^2} + \frac{(x-1)}{(1-\theta)^2}\right) = \frac{1}{\theta^2} + \frac{\frac{1}{\theta}-1}{(1-\theta)^2} = \frac{1}{\theta^2(1-\theta)} \Rightarrow$ Jeffrey's prior

$\propto \frac{1}{\theta(1-\theta)^{1/2}} = (1-\theta)^{-1/2}\,\theta^{-1}$.

Exercise 3.3

Suppose x has the Pareto distribution $Pareto(a, b)$, where a is known but b is unknown. So,

$$f(x|b) = ba^b x^{-b-1}; \qquad (x > a, \quad b > 0).$$

Find the Jeffreys prior and the corresponding posterior distribution for b.

Exercise 3.3

Suppose x has the Pareto distribution $Pareto(a, b)$, where a is known but b is unknown. So,

$$f(x|b) = ba^b x^{-b-1}; \qquad (x > a, \quad b > 0).$$

Find the Jeffreys prior and the corresponding posterior distribution for b.

Solution.

log-likelihood: $\ell(b) = \log b + b \log a - (b+1) \log x$

$$\frac{d\ell}{db} = \frac{1}{b} + \log a - \log x$$

Solution

$$\frac{d\ell}{db} = \frac{1}{b} + \log a - \log x$$

$$\frac{d^2\ell}{db^2} = -\frac{1}{b^2} \Rightarrow E\left(-\frac{d^2\ell}{db^2}\right) = \frac{1}{b^2}$$

Solution

$$\frac{d\ell}{db} = \frac{1}{b} + \log a - \log x$$

$$\frac{d^2\ell}{db^2} = -\frac{1}{b^2} \Rightarrow E\left(-\frac{d^2\ell}{db^2}\right) = \frac{1}{b^2}$$

Jeffreys' prior: $J(b) \propto |E\left(-\frac{d^2\ell}{db^2}\right)|^{1/2} = \frac{1}{b}$

Solution

$$\frac{d\ell}{db} = \frac{1}{b} + \log a - \log x$$

$$\frac{d^2\ell}{db^2} = -\frac{1}{b^2} \Rightarrow E\left(-\frac{d^2\ell}{db^2}\right) = \frac{1}{b^2}$$

Jeffreys' prior: $J(b) \propto |E\left(-\frac{d^2\ell}{db^2}\right)|^{1/2} = \frac{1}{b}$

Posterior:

$$\begin{aligned} f(b \mid x) &\propto f(x|b)J(b) \propto ba^b x^{-b-1}\frac{1}{b} = a^b x^{-b-1} \\ &\propto a^b x^{-b} = \left(\frac{a}{x}\right)^b = \exp\{-b\log(\frac{x}{a})\}. \end{aligned}$$

The posterior distribution is exponential with rate $\log(\frac{x}{a})$.

Exercise 3.5

You are interested in estimating θ, the probability that a drawing pin will land point up. Your prior belief can be described by a mixture of Beta distributions:

$$f(\theta) = \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1} + \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)}\theta^{p-1}(1-\theta)^{q-1}.$$

You throw a drawing pin n independent times, and observe x occasions on which the pin lands point up. Calculate the posterior distribution for θ.

Solution

Let X denote the number of times that the drawing pin lands point up.

$$X \sim Binomial(n, \theta)$$

Likelihood: $f(x \mid \theta) \propto \theta^x (1-\theta)^{n-x}$

Solution

Let X denote the number of times that the drawing pin lands point up.

$$X \sim Binomial(n, \theta)$$

Likelihood: $f(x \mid \theta) \propto \theta^x (1-\theta)^{n-x}$

Posterior: $f(\theta \mid x) \propto f(x \mid \theta) f(\theta)$

$$\propto \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)} \theta^{x+a-1}(1-\theta)^{n-x+b-1} + \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)} \theta^{x+p-1}(1-\theta)^{n-x+q-1}$$

Solution

Let X denote the number of times that the drawing pin lands point up.

$$X \sim Binomial(n, \theta)$$

Likelihood: $f(x \mid \theta) \propto \theta^x (1-\theta)^{n-x}$

Posterior: $f(\theta \mid x) \propto f(x \mid \theta) f(\theta)$

$$\propto \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)} \theta^{x+a-1}(1-\theta)^{n-x+b-1} + \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)} \theta^{x+p-1}(1-\theta)^{n-x+q-1}$$

$$= \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n-x+b)}{\Gamma(n+a+b)} \left\{ \frac{\Gamma(n+a+b)}{\Gamma(x+a)\Gamma(n-x+b)} \theta^{x+a-1}(1-\theta)^{n-x+b-1} \right\}$$

$$+ \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)} \frac{\Gamma(x+p)\Gamma(n-x+q)}{\Gamma(n+p+q)} \left\{ \frac{\Gamma(n+p+q)}{\Gamma(x+p)\Gamma(n-x+q)} \theta^{x+p-1}(1-\theta)^{n-x+q-1} \right\}$$

Solution

Let $\alpha = \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n-x+b)}{\Gamma(n+a+b)}$ and $\beta = \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)} \frac{\Gamma(x+p)\Gamma(n-x+q)}{\Gamma(n+p+q)}$.

Solution

Let $\alpha = \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n-x+b)}{\Gamma(n+a+b)}$ and $\beta = \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)} \frac{\Gamma(x+p)\Gamma(n-x+q)}{\Gamma(n+p+q)}$.

Posterior:

$$\begin{aligned} f(\theta \mid x) &= \left(\frac{\alpha}{\alpha+\beta}\right) \frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)} \theta^{A-1}(1-\theta)^{B-1} \\ &+ \left(\frac{\beta}{\alpha+\beta}\right) \frac{\Gamma(P+Q)}{\Gamma(P)\Gamma(Q)} \theta^{P-1}(1-\theta)^{Q-1}, \end{aligned}$$

where $A = x + a$, $B = b + n - x$, $P = p + x$ and $Q = q + n - x$.

Solution

Let $\alpha = \frac{\Gamma(a+b)}{2\Gamma(a)\Gamma(b)}\frac{\Gamma(x+a)\Gamma(n-x+b)}{\Gamma(n+a+b)}$ and $\beta = \frac{\Gamma(p+q)}{2\Gamma(p)\Gamma(q)}\frac{\Gamma(x+p)\Gamma(n-x+q)}{\Gamma(n+p+q)}$.

Posterior:

$$
\begin{aligned}
f(\theta \mid x) &= \left(\frac{\alpha}{\alpha+\beta}\right)\frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)}\theta^{A-1}(1-\theta)^{B-1} \\
&+ \left(\frac{\beta}{\alpha+\beta}\right)\frac{\Gamma(P+Q)}{\Gamma(P)\Gamma(Q)}\theta^{P-1}(1-\theta)^{Q-1},
\end{aligned}
$$

where $A = x + a$, $B = b + n - x$, $P = p + x$ and $Q = q + n - x$.

That is the posterior distribution of θ is a mixture of Beta distributions with updated parameters and updated mixing proportions.

Exercise 4.3

(a) Observations x_1 and x_2 are obtained of random variables X_1 and X_2, having Poisson distributions with respective means θ and $\phi\theta$, where ϕ is a known positive coefficient. Show, by evaluating the posterior density of θ, that the *Gamma* (p, q) family of prior distributions of θ is a conjugate or this data mode.

(b) Now suppose that ϕ is also an unknown parameter with prior density $f(\phi) = 1/(1+\phi)^2$, and independent of θ. Obtain the joint posterior distribution of θ and ϕ and show that the marginal posterior distribution of ϕ is proportional to

$$\frac{\phi^{x_2}}{(1+\phi)^2 (1+\phi+q)^{x_1+x_2+p}}$$

Exercise 4.3-Solution

(a) $f(x_1, x_2 \mid \theta) = \frac{e^{-\theta}\theta^{x_1}}{x_1!}\frac{e^{-\theta\phi}(\theta\phi)^{x_2}}{x_2!} \propto e^{-\theta(1+\phi)}\theta^{x_1+x_2}$

Exercise 4.3-Solution

(a) $f(x_1, x_2 \mid \theta) = \frac{e^{-\theta}\theta^{x_1}}{x_1!} \frac{e^{-\theta\phi}(\theta\phi)^{x_2}}{x_2!} \propto e^{-\theta(1+\phi)}\theta^{x_1+x_2}$

Prior:
$f(\theta) = \frac{q^p}{\Gamma(p)}\theta^{p-1}e^{-q\theta} \propto \theta^{p-1}e^{-q\theta} \equiv Gamma(p, q)$

Exercise 4.3-Solution

(a) $f(x_1, x_2 \mid \theta) = \frac{e^{-\theta}\theta^{x_1}}{x_1!}\frac{e^{-\theta\phi}(\theta\phi)^{x_2}}{x_2!} \propto e^{-\theta(1+\phi)}\theta^{x_1+x_2}$

Prior:
$f(\theta) = \frac{q^p}{\Gamma(p)}\theta^{p-1}e^{-q\theta} \propto \theta^{p-1}e^{-q\theta} \equiv Gamma(p, q)$

Posterior:
$f(\theta \mid x_1, x_2) = \theta^{x_1+x_2+p-1}e^{-(1+\phi+q)\theta} \equiv$
$Gamma(x_1 + x_2 + p, 1 + \phi + q) \Longrightarrow$ Gamma family is conjugate for this model.
Posterior Mean $= \frac{x_1+x_2+p}{1+\phi+q}$

Exercise 4.3-Solution

(a) $f(x_1, x_2 \mid \theta) = \frac{e^{-\theta}\theta^{x_1}}{x_1!}\frac{e^{-\theta\phi}(\theta\phi)^{x_2}}{x_2!} \propto e^{-\theta(1+\phi)}\theta^{x_1+x_2}$

Prior:
$f(\theta) = \frac{q^p}{\Gamma(p)}\theta^{p-1}e^{-q\theta} \propto \theta^{p-1}e^{-q\theta} \equiv Gamma(p, q)$

Posterior:
$f(\theta \mid x_1, x_2) = \theta^{x_1+x_2+p-1}e^{-(1+\phi+q)\theta} \equiv$
$Gamma(x_1 + x_2 + p, 1 + \phi + q) \Longrightarrow$ Gamma family is conjugate for this model.
Posterior Mean $= \frac{x_1+x_2+p}{1+\phi+q}$

(b) Joint Posterior:
$f(\phi, \theta \mid x_1, x_2) \propto f(\phi) f(\theta) f(x_1, x_2 \mid \phi, \theta) \propto$
$\frac{1}{(1+\phi)^2}\theta^{p-1}e^{-q\theta}e^{-\theta(1+\phi)}\theta^{x_1}\theta^{x_2}\phi^{x_2} =$
$\frac{1}{(1+\phi)^2}\theta^{p+x_1+x_2-1}\phi^{x_2}e^{-(1+\phi+1)\theta}$

Exercise 4.3-Solution

Marginal Posterior:

$$
\begin{aligned}
f(\phi, x_1, x_2) &= \int f(\phi, \theta \mid x_1, x_2)\, d\theta \\
&\propto \int_0^\infty \frac{\phi^{x_2}}{(1+\phi)^2} \theta^{p+x_1+x_2-1} e^{-(1+\phi+1)\theta} d\theta \\
&= \frac{\phi^{x_2}}{(1+\phi)^2} \frac{\Gamma(p+x_1+x_2)}{(1+\phi+q)^{x_1+x_2+p}} \\
&\propto \frac{\phi^{x_2}}{(1+\phi)^2} \frac{1}{(1+\phi+q)^{x_1+x_2+p}}.
\end{aligned}
$$

Exercise 5.1

I have been offered an objet d'art at what seems a bargain price of £100. If it is not genuine, then it is worth nothing. If it is genuine I believe I can sell it immediately for £300. I believe there is a 0.5 chance that the object is genuine. Should I buy the object?

An art 'expert' has undergone a test of her reliability in which she has separately pronounced judgement — 'genuine' or 'counterfeit' — on a large number of art subjects of know origin. From these it appears that she has probability 0.8 of detecting a counterfeit and probability 0.7 of recognising a genuine object. The expert charges £30 for her services. Is it to my advantage to pay her for an assessment?

Solution

1. The parameter space is $\Theta = \{\theta_1, \theta_2\}$, where θ_1 and θ_2 correspond to the object being genuine and counterfeit respectively;
2. The set of actions is $A = \{a_1, a_2\}$ where a_1 and a_2 correspond to buying and not buying the object respectively;
3. The loss function is

$L(\theta, a)$	θ_1	θ_2
a_1	-200	100
a_2	0	0

Solution

1. The parameter space is $\Theta = \{\theta_1, \theta_2\}$, where θ_1 and θ_2 correspond to the object being genuine and counterfeit respectively;
2. The set of actions is $A = \{a_1, a_2\}$ where a_1 and a_2 correspond to buying and not buying the object respectively;
3. The loss function is

$L(\theta, a)$	θ_1	θ_2
a_1	-200	100
a_2	0	0

The **decision strategy** is to evaluate the expected loss for each action and choose the action which has the minimum expected loss.

Solution

The expected loss is first calculated based on the prior distribution of θ: $f(\theta_1) = 0.5$ and $f(\theta_2) = 0.5$.

Solution

The expected loss is first calculated based on the prior distribution of θ: $f(\theta_1) = 0.5$ and $f(\theta_2) = 0.5$.

$f(\theta)$	0.5	0.5	
$L(\theta, a)$	θ_1	θ_2	$E[L(\theta, a)]$
a_1	-200	100	$0.5 \times (-200) + 0.5 \times 100 = -50$
a_2	0	0	$0.5 \times 0 + 0.5 \times 0 = 0$

Solution

The expected loss is first calculated based on the prior distribution of θ: $f(\theta_1) = 0.5$ and $f(\theta_2) = 0.5$.

$f(\theta)$	0.5	0.5	
$L(\theta, a)$	θ_1	θ_2	$E[L(\theta, a)]$
a_1	-200	100	$0.5 \times (-200) + 0.5 \times 100 = -50$
a_2	0	0	$0.5 \times 0 + 0.5 \times 0 = 0$

So, under the prior, the best action is to buy (according to minimisation of expected loss - here maximization of expected profit). The expected profit of this action is £50.

Solution

The expected loss is first calculated based on the prior distribution of θ: $f(\theta_1) = 0.5$ and $f(\theta_2) = 0.5$.

$f(\theta)$	0.5	0.5	
$L(\theta, a)$	θ_1	θ_2	$E[L(\theta, a)]$
a_1	-200	100	$0.5 \times (-200) + 0.5 \times 100 = -50$
a_2	0	0	$0.5 \times 0 + 0.5 \times 0 = 0$

So, under the prior, the best action is to buy (according to minimisation of expected loss - here maximization of expected profit). The expected profit of this action is £50.

Suppose I pay the expert. Her possible conclusions about the object (observations) are $x_1 =$ 'says genuine' and $x_2 =$ 'says counterfeit'.

Solution

		θ_1	θ_2	
	$f(x_1\|\theta)$	0.70	0.20	
Likelihoods	$f(x_2\|\theta)$	0.30	0.80	

Solution

		θ_1	θ_2	
	$f(x_1\|\theta)$	0.70	0.20	
Likelihoods	$f(x_2\|\theta)$	0.30	0.80	
Prior	$f(\theta)$	0.50	0.50	

Solution

		θ_1	θ_2		
	$f(x_1\|\theta)$	0.70	0.20		
Likelihoods	$f(x_2\|\theta)$	0.30	0.80		
Prior	$f(\theta)$	0.50	0.50		
	$f(x_1, \theta)$	0.35	0.10	0.45	$f(x_1)$
Joints	$f(x_2, \theta)$	0.15	0.40	0.55	$f(x_2)$

Solution

		θ_1	θ_2			
	$f(x_1\|\theta)$	0.70	0.20			
Likelihoods	$f(x_2\|\theta)$	0.30	0.80			
Prior	$f(\theta)$	0.50	0.50			
	$f(x_1, \theta)$	0.35	0.10	0.45	$f(x_1)$	
Joints	$f(x_2, \theta)$	0.15	0.40	0.55	$f(x_2)$	
				a_1	a_2	
	$f(\theta\|x_1)$	7/9	2/9	-1200/9	0	Expected
Posteriors	$f(\theta\|x_2)$	3/11	8/11	200/11	0	Losses

$E(L(\theta, a_1) \mid x_1) = f(\theta_1 \mid x_1)L(\theta_1, a_1) + f(\theta_2 \mid x_1)L(\theta_2, a_1) =$
$\frac{7}{9} \times (-200) + \frac{2}{9} \times 100 = -\frac{1200}{9}$
$E(L(\theta, a_1) \mid x_2) = f(\theta_1 \mid x_2)L(\theta_1, a_1) + f(\theta_2 \mid x_2)L(\theta_2, a_1) =$
$\frac{3}{11} \times (-200) + \frac{8}{11} \times 100 = \frac{200}{11}$
$E(L(\theta, a_2) \mid x_1) = E(L(\theta, a_2) \mid x_2) = 0$

Solution

Bayes Decision Rule: $d(x_1) = a_1$, $d(x_2) = a_2$.

Solution

Bayes Decision Rule: $d(x_1) = a_1$, $d(x_2) = a_2$.

Bayes Risk:

$$BR(d) = \rho(d(x_1), x_1)f(x_1) + \rho(d(x_2), x_2)f(x_2) = -\frac{1200}{9} \times 0.45 + 0 \times 0.55 = -60$$

Solution

Bayes Decision Rule: $d(x_1) = a_1$, $d(x_2) = a_2$.

Bayes Risk:

$$BR(d) = \rho(d(x_1), x_1)f(x_1) + \rho(d(x_2), x_2)f(x_2) = -\frac{1200}{9} \times 0.45 + 0 \times 0.55 = -60$$

That is the profit associated with our decision rule is £60. The gain of £10 does not worth the £30 cost of the expert's services.

Exercise 5.2

Consider a decision problem with two actions, α_1 ana α_2 and a loss function which depends on a parameter θ, with $0 \leq \theta \leq 1$. The loss function is

$$L(\theta, \alpha) = \begin{cases} 0 & \alpha = \alpha_1. \\ 2 - 3\theta & \alpha = \alpha_2. \end{cases}$$

Assume a $Beta(1, 1)$ prior for θ, and an observation $X \sim Binomial(n, \theta)$. The posterior distribution is $Beta(x + 1, n - x + 1)$. Calculate the expected loss of each action and the Bayes rule.

Exercise 5.2

Consider a decision problem with two actions, α_1 ana α_2 and a loss function which depends on a parameter θ, with $0 \leq \theta \leq 1$. The loss function is

$$L(\theta, \alpha) = \begin{cases} 0 & \alpha = \alpha_1. \\ 2 - 3\theta & \alpha = \alpha_2. \end{cases}$$

Assume a $Beta(1,1)$ prior for θ, and an observation $X \sim Binomial(n, \theta)$. The posterior distribution is $Beta(x+1, n-x+1)$. Calculate the expected loss of each action and the Bayes rule.

Exercise 5.2-Solution

Prior:
$f(\theta) = 1, \quad 0 \leq \theta \leq 1$
Likelihood:
$f(x \mid \theta) = \binom{n}{\theta} \theta^x (1-\theta)^{n-x}$

Exercise 5.2-Solution

Prior:
$f(\theta) = 1, \quad 0 \leq \theta \leq 1$
Likelihood:
$f(x \mid \theta) = \binom{n}{\theta} \theta^x (1-\theta)^{n-x}$

Posterior:
$f(\theta \mid x) \propto \theta^x (1-\theta)^{n-x} \equiv Beta(x+1, n-x+1)$

Exercise 5.2-Solution

Prior:
$f(\theta) = 1, \quad 0 \leq \theta \leq 1$
Likelihood:
$f(x \mid \theta) = \binom{n}{\theta} \theta^{x} (1-\theta)^{n-x}$

Posterior:
$f(\theta \mid x) \propto \theta^{x} (1-\theta)^{n-x} \equiv Beta(x+1, n-x+1)$

Expected loss under the two actions:
$E(L(\theta, \alpha_1) \mid x) = \int_0^1 0 d\theta = 0$

Exercise 5.2-Solution

Prior:
$f\left(\theta\right)=1,\quad 0\leq\theta\leq 1$
Likelihood:
$f(x\mid\theta)=\binom{n}{\theta}\theta^{x}\left(1-\theta\right)^{n-x}$

Posterior:
$f(\theta\mid x)\propto\theta^{x}\left(1-\theta\right)^{n-x}\equiv Beta(x+1,n-x+1)$

Expected loss under the two actions:
$E\left(L\left(\theta,\alpha_1\right)\mid x\right)=\int_0^1 0d\theta=0$
$E\left(L\left(\theta,\alpha_2\right)\mid x\right)=\int_0^1\left(2-3\theta\right)f(\theta\mid x)d\theta=2-3E\left(\theta\mid x\right)=$
$2-\frac{3(x+1)}{x+1+n-x+1}=2-\frac{3(x+1)}{n+2}$

Exercise 5.2-Solution

Prior:
$f(\theta) = 1, \quad 0 \leq \theta \leq 1$
Likelihood:
$f(x \mid \theta) = \binom{n}{\theta} \theta^x (1-\theta)^{n-x}$

Posterior:
$f(\theta \mid x) \propto \theta^x (1-\theta)^{n-x} \equiv Beta(x+1, n-x+1)$

Expected loss under the two actions:
$E(L(\theta, \alpha_1) \mid x) = \int_0^1 0 d\theta = 0$
$E(L(\theta, \alpha_2) \mid x) = \int_0^1 (2-3\theta) f(\theta \mid x) d\theta = 2 - 3E(\theta \mid x) =$
$2 - \frac{3(x+1)}{x+1+n-x+1} = 2 - \frac{3(x+1)}{n+2}$

We prefer the action α_2 if $E(L(\theta, \alpha_2) \mid x) < E(L(\theta, \alpha_1) \mid x) \Rightarrow$
$2 - \frac{3(x+1)}{n+2} < 0 \Rightarrow \cdots \Rightarrow x \geq \frac{2n+1}{3}$

Exercise 5.3

(a) For a parameter θ with a posterior distribution described by the $Beta(P, Q)$ distribution, find the posterior mode in terms of P and Q and compare it with the posterior mean.

Solution.

$f(\theta \mid x) \propto \theta^{P-1}(1-\theta)^{Q-1}$
$\log f(\theta \mid x) = (P-1)\log\theta + (Q-1)\log(1-\theta) + c$

Exercise 5.3

(a) For a parameter θ with a posterior distribution described by the $Beta(P, Q)$ distribution, find the posterior mode in terms of P and Q and compare it with the posterior mean.

Solution.

$f(\theta \mid x) \propto \theta^{P-1}(1-\theta)^{Q-1}$
$\log f(\theta \mid x) = (P-1)\log\theta + (Q-1)\log(1-\theta) + c$
$\frac{d \log f(\theta|x)}{d\theta} = 0 \Rightarrow \frac{P-1}{\theta} - \frac{Q-1}{1-\theta} = 0 \Rightarrow$
$\frac{P-1}{\theta} = \frac{Q-1}{1-\theta} \Rightarrow P-1-(P-1)\theta = (Q-1)\theta \Rightarrow$
$(P+Q-2)\theta = P-1 \Rightarrow \theta = \frac{P-1}{P+Q-2}$ [posterior mode]

Exercise 5.3

(a) For a parameter θ with a posterior distribution described by the $Beta(P, Q)$ distribution, find the posterior mode in terms of P and Q and compare it with the posterior mean.

Solution.

$$f(\theta \mid x) \propto \theta^{P-1}(1-\theta)^{Q-1}$$
$$\log f(\theta \mid x) = (P-1)\log\theta + (Q-1)\log(1-\theta) + c$$
$$\frac{d \log f(\theta|x)}{d\theta} = 0 \Rightarrow \frac{P-1}{\theta} - \frac{Q-1}{1-\theta} = 0 \Rightarrow$$
$$\frac{P-1}{\theta} = \frac{Q-1}{1-\theta} \Rightarrow P-1-(P-1)\theta = (Q-1)\theta \Rightarrow$$
$$(P+Q-2)\theta = P-1 \Rightarrow \theta = \frac{P-1}{P+Q-2} \text{ [posterior mode]}$$

Posterior mean: $E(\theta \mid x) = \frac{P}{P+Q}$ [closer to 1/2 than mode]

Exercise 5.4

The parameter θ has a $\mathit{Beta}(3, 2)$ posterior density. Show that the interval $[5/21, 20/21]$ is a 94.3% highest posterior density region for θ.

Solution.

Exercise 5.4

The parameter θ has a $Beta(3,2)$ posterior density. Show that the interval $[5/21, 20/21]$ is a 94.3% highest posterior density region for θ.

Solution.
A region $C_\alpha(x)$ is a $100(1-\alpha)\%$ credible region for θ if $\int_{C_\alpha(x)} f(\theta \mid x)\, d\theta = 1-\alpha$. It is an HPD region if $C_\alpha(x) = \{\theta : f(\theta \mid x) \geq \gamma\}$.

Exercise 5.4

The parameter θ has a $Beta(3,2)$ posterior density. Show that the interval $[5/21, 20/21]$ is a 94.3% highest posterior density region for θ.

Solution.
A region $C_\alpha(x)$ is a $100(1-\alpha)\%$ credible region for θ if $\int_{C_\alpha(x)} f(\theta \mid x)\, d\theta = 1-\alpha$. It is an HPD region if $C_\alpha(x) = \{\theta : f(\theta \mid x) \geq \gamma\}$.

In our case, $f(\theta \mid x) = \frac{1}{B(3,2)}\theta^2(1-\theta) = 12\theta^2(1-\theta), \quad \theta \in [0,1]$.

$\int_a^b 12\theta^2(1-\theta)\, d\theta = 0.943 \Leftrightarrow 12\left[\frac{\theta^3}{3} - \frac{\theta^4}{4}\right]_a^b = 0.943 \Leftrightarrow$
$4b^3 - 3b^4 - 4a^3 + 3a^4 = 0.943.$

Exercise 5.4

The parameter θ has a $Beta(3,2)$ posterior density. Show that the interval $[5/21, 20/21]$ is a 94.3% highest posterior density region for θ.

Solution.

A region $C_\alpha(x)$ is a $100(1-\alpha)\%$ credible region for θ if $\int_{C_\alpha(x)} f(\theta \mid x)\, d\theta = 1-\alpha$. It is an HPD region if $C_\alpha(x) = \{\theta : f(\theta \mid x) \geq \gamma\}$.

In our case, $f(\theta \mid x) = \frac{1}{B(3,2)}\theta^2(1-\theta) = 12\theta^2(1-\theta), \quad \theta \in [0,1]$.

$\int_a^b 12\theta^2(1-\theta)\, d\theta = 0.943 \Leftrightarrow 12\left[\frac{\theta^3}{3} - \frac{\theta^4}{4}\right]_a^b = 0.943 \Leftrightarrow$

$4b^3 - 3b^4 - 4a^3 + 3a^4 = 0.943.$

Moreover, $12a^2(1-a) = 12b^2(1-b) = \gamma$.

Exercise 5.4

The parameter θ has a $Beta(3,2)$ posterior density. Show that the interval $[5/21, 20/21]$ is a 94.3% highest posterior density region for θ.

Solution.
A region $C_\alpha(x)$ is a $100(1-\alpha)\%$ credible region for θ if $\int_{C_\alpha(x)} f(\theta \mid x)\, d\theta = 1-\alpha$. It is an HPD region if $C_\alpha(x) = \{\theta : f(\theta \mid x) \geq \gamma\}$.

In our case, $f(\theta \mid x) = \frac{1}{B(3,2)}\theta^2 (1-\theta) = 12\theta^2(1-\theta)\,, \quad \theta \in [0,1]\,.$

$\int_a^b 12\theta^2(1-\theta)\, d\theta = 0.943 \Leftrightarrow 12\left[\frac{\theta^3}{3} - \frac{\theta^4}{4}\right]_a^b = 0.943 \Leftrightarrow$

$4b^3 - 3b^4 - 4a^3 + 3a^4 = 0.943.$

Moreover, $12a^2(1-a) = 12b^2(1-b) = \gamma.$

Solving the two-equation system, we derive $a = 5/21$ and $b = 20/21$.

Exercise 5.6

A parameter θ has a posterior density that is *Gamma*$(1, 1)$. Calculate the 95% HPD region for θ. Now consider the transformation $\phi = \sqrt{2\theta}$. Obtain the posterior density of ϕ and explain why the highest posterior density region for ϕ is not obtained by transforming the interval for θ in the same way.

Exercise 5.6

A parameter θ has a posterior density that is *Gamma*$(1, 1)$. Calculate the 95% HPD region for θ. Now consider the transformation $\phi = \sqrt{2\theta}$. Obtain the posterior density of ϕ and explain why the highest posterior density region for ϕ is not obtained by transforming the interval for θ in the same way.

Solution.

The posterior density of θ is decreasing (Exponential(1)). Thus, the 95% HPD region is an interval of the form [0,b] satisfying $\int_0^b e^{-\theta} = 0.95 \Leftrightarrow 1 - e^{-b} = 0.95 \Longrightarrow C_{0.05}(x) = [0, 3]$.

Exercise 5.6

A parameter θ has a posterior density that is *Gamma*$(1, 1)$. Calculate the 95% HPD region for θ. Now consider the transformation $\phi = \sqrt{2\theta}$. Obtain the posterior density of ϕ and explain why the highest posterior density region for ϕ is not obtained by transforming the interval for θ in the same way.

Solution.

The posterior density of θ is decreasing (Exponential(1)). Thus, the 95% HPD region is an interval of the form [0,b] satisfying $\int_0^b e^{-\theta} = 0.95 \Leftrightarrow 1 - e^{-b} = 0.95 \Longrightarrow C_{0.05}(x) = [0, 3]$.
Now, we have that $\phi = \sqrt{2\theta} \Leftrightarrow \theta = \frac{\phi^2}{2}$. Posterior density of ϕ:

$$f_\phi(\phi \mid x) = f_\theta(\phi) \left| \frac{d\theta}{d\phi} \right| = \phi e^{-\frac{1}{2}\phi^2}.$$

Exercise 5.6

A parameter θ has a posterior density that is *Gamma*$(1, 1)$. Calculate the 95% HPD region for θ. Now consider the transformation $\phi = \sqrt{2\theta}$. Obtain the posterior density of ϕ and explain why the highest posterior density region for ϕ is not obtained by transforming the interval for θ in the same way.

Solution.

The posterior density of θ is decreasing (Exponential(1)). Thus, the 95% HPD region is an interval of the form [0,b] satisfying $\int_0^b e^{-\theta} = 0.95 \Leftrightarrow 1 - e^{-b} = 0.95 \Longrightarrow C_{0.05}(x) = [0, 3]$.

Now, we have that $\phi = \sqrt{2\theta} \Leftrightarrow \theta = \frac{\phi^2}{2}$. Posterior density of ϕ:

$$f_\phi(\phi \mid x) = f_\theta(\phi) \left| \frac{d\theta}{d\phi} \right| = \phi e^{-\frac{1}{2}\phi^2}.$$

Since the posterior density of ϕ is not monotonic, the credibility interval will be of the form $[a, b]$ with $a \neq 0$. Hence, it is not a transformation of the credibility interval for θ.

Exercise 5.8

Consider a sample $x_1, \ldots, x_n$ consisting of independent draws from a Poisson random variable with mean θ. Consider the hypothesis test, with Null hypothesis

$$\mathrm{H}_0 : \theta = 1$$

against an alternative hypothesis

$$\mathrm{H}_1 : \theta \neq 1$$

Assume a prior probability of 0.95 for H_0 and a Gamma prior

$$f(\theta) = \frac{q^p}{\Gamma(p)} \theta^{p-1} \exp\{-q\theta\},$$

under H_1.

(a) Calculate the posterior probability of H_0.

Solution

Likelihood: $f(x \mid \theta) = \frac{1}{\prod x_i!} \theta^{\sum x_i} e^{-n\theta}$

Solution

Likelihood: $f(x \mid \theta) = \frac{1}{\prod x_i!} \theta^{\sum x_i} e^{-n\theta}$

$$
\begin{aligned}
H_0: \quad P(H_0 \mid x) &\propto P(H_0)P(x \mid H_0) \\
&= P(H_0) f(x \mid \theta = 1) = 0.95 \frac{1}{\prod x_i!} e^{-n}
\end{aligned}
$$

Solution

Likelihood: $f(x \mid \theta) = \frac{1}{\prod x_i!}\theta^{\sum x_i} e^{-n\theta}$

$$
\begin{aligned}
H_0: \quad P(H_0 \mid x) &\propto P(H_0)P(x \mid H_0) \\
&= P(H_0) f(x \mid \theta = 1) = 0.95 \frac{1}{\prod x_i!} e^{-n}
\end{aligned}
$$

$$
\begin{aligned}
H_1: \quad P(H_1 \mid x) &\propto P(H_1)P(x \mid H_1) \\
&= P(H_1) \int f(x \mid \theta) f(\theta \mid H_1) d\theta \\
&= 0.05 \int_0^\infty \frac{1}{\prod x_i!} \theta^{\sum x_i} e^{-n\theta} \frac{q^p}{\Gamma(p)} \theta^{p-1} e^{-q\theta} d\theta \\
&= 0.05 \frac{1}{\prod x_i!} \frac{q^p}{\Gamma(p)} \int_0^\infty \theta^{\sum x_i + p - 1} e^{-(n+q)\theta} d\theta \\
&= 0.05 \frac{1}{\prod x_i!} \frac{q^p}{\Gamma(p)} \frac{\Gamma(\sum x_i + p)}{(n+q)^{\sum x_i + p}}
\end{aligned}
$$

Exercise 5.8 - Solution

Let $\alpha = 0.95e^{-n}$ and $\beta = 0.05\frac{q^p}{\Gamma(p)}\frac{\Gamma(\sum x_i+p)}{(n+q)^{\sum x_i+p}}$.

Then $P(H_0 \mid x) = \frac{\alpha}{\alpha+\beta}$.

Exercise 5.8 - Solution

Let $\alpha = 0.95e^{-n}$ and $\beta = 0.05\frac{q^p}{\Gamma(p)}\frac{\Gamma(\sum x_i + p)}{(n+q)^{\sum x_i + p}}$.

Then $P(H_0 \mid x) = \frac{\alpha}{\alpha+\beta}$.

(b) Assume $n = 10$, $\sum_{i=1}^{n} x_i = 20$, and $p = 2q$. What is the posterior probability of H_0 for each of $p = 2, 1, 0.5, 0.1$. What happens to this posterior probability as $p \to 0$?

Exercise 5.8 - Solution

Let $\alpha = 0.95e^{-n}$ and $\beta = 0.05\frac{q^p}{\Gamma(p)}\frac{\Gamma(\sum x_i+p)}{(n+q)^{\sum x_i+p}}$.

Then $P(H_0 \mid x) = \frac{\alpha}{\alpha+\beta}$.

(b) Assume $n = 10$, $\sum_{i=1}^n x_i = 20$, and $p = 2q$. What is the posterior probability of H_0 for each of $p = 2, 1, 0.5, 0.1$. What happens to this posterior probability as $p \to 0$?

For $n = 10$, $\sum_{i=1}^n x_i = 20$, and $p = 2q$:
$\alpha = 0.95e^{-10}$ and $\beta = 0.05\frac{(p/2)^p}{\Gamma(p)}\frac{\Gamma(20+p)}{(10+p/2)^{20+p}}$.

Exercise 6.1

A random sample $x_1, \ldots, x_n$ is observed from a *Poisson*(θ) distribution.The prior on θ is a *Gamma*(g, h). Show that the predictive distribution for a future observation y from this *Poisson*(θ) distribution is

$$f(y \mid x) = \binom{y + G - 1}{G - 1} \left(\frac{1}{1+H}\right)^y \left(1 - \frac{1}{1+H}\right)^G \quad y = 0, 1, \ldots$$

What is this distribution?

Exercise 6.1

A random sample $x_1, \ldots, x_n$ is observed from a $Poisson(\theta)$ distribution.The prior on θ is a $Gamma(g, h)$. Show that the predictive distribution for a future observation y from this $Poisson(\theta)$ distribution is

$$f(y \mid x) = \binom{y+G-1}{G-1} \left(\frac{1}{1+H}\right)^{y} \left(1 - \frac{1}{1+H}\right)^{G} \quad y = 0, 1, \ldots$$

What is this distribution?

Solution

$f\left(y \mid x\right) = \int f\left(y \mid \theta\right) f\left(x \mid \theta\right) d\theta$

Exercise 6.1

A random sample $x_1, \ldots, x_n$ is observed from a *Poisson*(θ) distribution.The prior on θ is a *Gamma*(g, h). Show that the predictive distribution for a future observation y from this *Poisson*(θ) distribution is

$$f(y \mid x) = \binom{y+G-1}{G-1} \left(\frac{1}{1+H}\right)^y \left(1 - \frac{1}{1+H}\right)^G \quad y = 0, 1, \ldots$$

What is this distribution?

Solution

$f(y \mid x) = \int f(y \mid \theta) f(x \mid \theta) d\theta$

Posterior: $f(\theta \mid x) \propto \prod_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{x_i!} \frac{h^g}{\Gamma(g)} \theta^{g-1} \exp\{-h\theta\}$

Exercise 6.1

A random sample $x_1, \ldots, x_n$ is observed from a $Poisson(\theta)$ distribution.The prior on θ is a $Gamma(g, h)$. Show that the predictive distribution for a future observation y from this $Poisson(\theta)$ distribution is

$$f(y \mid x) = \binom{y+G-1}{G-1} \left(\frac{1}{1+H}\right)^y \left(1 - \frac{1}{1+H}\right)^G \quad y = 0, 1, \ldots$$

What is this distribution?

Solution

$f\left(y \mid x\right) = \int f\left(y \mid \theta\right) f\left(x \mid \theta\right) d\theta$

Posterior: $f\left(\theta \mid x\right) \propto \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \frac{h^g}{\Gamma(g)} \theta^{g-1} \exp\left\{-h\theta\right\}$

Mariginal likelihood:

$f\left(x\right) = \int_0^\infty \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \frac{h^g}{\Gamma(g)} \theta^{g-1} \exp\left\{-h\theta\right\} d\theta =$

$\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}}}{\prod_{i=1}^{n} x_i! \Gamma(g)} \int_0^\infty \exp\left\{-\left(n+h\right)\theta\right\} \theta^{\sum_{i=1}^{n} x_i + g - 1} d\theta$

Exercise 6.1- Solution

$$= \frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}}}{\prod_{i=1}^{n} x_i! \Gamma(g)} \frac{\Gamma \sum_{i=1}^{n} x_i + g}{(n+h)^{\sum_{i=1}^{n} x_i + g}}$$

$$\Rightarrow f(\theta \mid x) = \frac{\prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \frac{h^g}{\Gamma(g)} \theta^{g-1} \exp\{-h\theta\}}{\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}}}{\Gamma(g)\prod_{i=1}^{n} x_i!} \frac{\Gamma\left(\sum_{i=1}^{n} x_i + g\right)}{(n+h)^{\sum_{i=1}^{n} x_i + g}}}$$

$$= \frac{(n+h)^{\sum_{i=1}^{n} x_i + g}}{\Gamma\left(\sum_{i=1}^{n} x_i + g\right)} \theta^{\sum_{i=1}^{n} x_i + g - 1} \exp\{-(h+n)\theta\}$$

Exercise 6.1- Solution

Predictive distribution

$f(y \mid x) = \int f(y \mid \theta) f(\theta \mid x) d\theta =$

$\frac{(n+h)^{\sum_{i=1}^{n} x_i + g}}{\Gamma\left(\sum_{i=1}^{n} x_i + g\right)} \int_0^\infty \frac{e^{-\theta}\theta^y}{y!} \theta^{\sum_{i=1}^{n} x_i + g - 1} \exp\left\{-\left(h+n\right)\theta\right\} =$

Exercise 6.1- Solution

Predictive distribution

$f(y \mid x) = \int f(y \mid \theta) f(\theta \mid x) d\theta =$

$\frac{(n+h)^{\sum_{i=1}^n x_i+g}}{\Gamma\left(\sum_{i=1}^n x_i+g\right)} \int_0^\infty \frac{e^{-\theta}\theta^y}{y!} \theta^{\sum_{i=1}^n x_i+g-1} \exp\left\{-\left(h+n\right)\theta\right\} =$

$\frac{(n+h)^{\sum_{i=1}^n x_i+g}}{\Gamma\left(\sum_{i=1}^n x_i+g\right)y!} \int_0^\infty \theta^{\sum_{i=1}^n x_i+g+y-1} \exp\left\{-\left(h+n+1\right)\theta\right\} =$

Exercise 6.1- Solution

Predictive distribution

$$f(y \mid x) = \int f(y \mid \theta) f(\theta \mid x) d\theta =$$

$$\frac{(n+h)^{\sum_{i=1}^n x_i+g}}{\Gamma\left(\sum_{i=1}^n x_i+g\right)} \int_0^\infty \frac{e^{-\theta}\theta^y}{y!} \theta^{\sum_{i=1}^n x_i+g-1} \exp\left\{-\left(h+n\right)\theta\right\} =$$

$$\frac{(n+h)^{\sum_{i=1}^n x_i+g}}{\Gamma\left(\sum_{i=1}^n x_i+g\right)y!} \int_0^\infty \theta^{\sum_{i=1}^n x_i+g+y-1} \exp\left\{-\left(h+n+1\right)\theta\right\} =$$

$$\frac{(n+h)^{\sum_{i=1}^n x_i+g}}{\Gamma\left(\sum_{i=1}^n x_i+g\right)y!} \frac{\Gamma\left(\sum_{i=1}^n x_i+g+y\right)}{(h+n+1)^{\sum_{i=1}^n x_i+g+y}} =$$

$$\frac{\left(\sum_{i=1}^n x_i+g+y-1\right)!}{\left(\sum_{i=1}^n x_i+g-1\right)y!} \left(\frac{1}{h+n+1}\right)^y \left(\frac{n+h}{n+h+1}\right)^{\sum_{i=1}^n x_i+g} =$$

$$\binom{G-1+y}{G-1} \left(\frac{1}{1+H}\right)^y \left(1-\frac{1}{H+1}\right)^G \equiv NegativeBinomial\left(\frac{H}{1+H}, G\right)$$

Exercise 6.3

A random sample $x_1, \ldots, x_n$ is observed from a $N(\theta, \sigma^2)$ distribution with σ^2 known, and a normal prior for θ is assumed, leading to a posterior distribution $N(B, D^2)$ for θ. Show that the predictive distribution for a further observation, y, from the $N(\theta, \sigma^2)$ distribution, is $N(B, D^2 + \sigma^2)$.

Exercise 6.3

A random sample $x_1, \ldots, x_n$ is observed from a $N(\theta, \sigma^2)$ distribution with σ^2 known, and a normal prior for θ is assumed, leading to a posterior distribution $N(B, D^2)$ for θ. Show that the predictive distribution for a further observation, y, from the $N(\theta, \sigma^2)$ distribution, is $N(B, D^2 + \sigma^2)$.

Solution

Posterior: $f(\theta \mid x) = \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta - B)^2\right\}$

Predictive: $f(y \mid x) = \int f(y \mid \theta) f(\theta \mid x) d\theta$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y - \theta)^2\right\} \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta - B)^2\right\} d\theta$$

Exercise 6.3 - Solution

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta)^2\right\} \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta-B)^2\right\} d\theta$$

Exercise 6.3 - Solution

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta)^2\right\} \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta-B)^2\right\} d\theta$$

$$= \frac{1}{2\pi\sigma D} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{y^2}{\sigma^2} - 2\frac{y\theta}{\sigma^2} + \frac{\theta^2}{\sigma^2} + \frac{\theta^2}{D^2} - 2\frac{\theta B}{D^2} + \frac{B^2}{D^2}\right]\right\} d\theta$$

Exercise 6.3 - Solution

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta)^2\right\} \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta - B)^2\right\} d\theta$$

$$= \frac{1}{2\pi\sigma D} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}[\frac{y^2}{\sigma^2} - 2\frac{y\theta}{\sigma^2} + \frac{\theta^2}{\sigma^2} + \frac{\theta^2}{D^2} - 2\frac{\theta B}{D^2} + \frac{B^2}{D^2}]\right\} d\theta$$

$$= \int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^{-2}+D^{-2}}{2}[\theta^2 - 2\theta\frac{y\sigma^{-2}+BD^{-2}}{\sigma^{-2}+D^{-2}} + (\frac{y\sigma^{-2}+BD^{-2}}{\sigma^{-2}+D^{-2}})^2]\right\} d\theta$$

$$\times \frac{1}{2\pi\sigma D} \exp\left\{-\frac{y^2}{2\sigma^2} - \frac{B^2}{2D^2} + \frac{(\sigma^{-2}y+BD^{-2})^2}{2(\sigma^{-2}+D^{-2})}\right\}$$

Exercise 6.3 - Solution

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta)^2\right\} \frac{1}{\sqrt{2\pi D^2}} \exp\left\{-\frac{1}{2D^2}(\theta-B)^2\right\} d\theta$$

$$= \frac{1}{2\pi\sigma D} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{y^2}{\sigma^2} - 2\frac{y\theta}{\sigma^2} + \frac{\theta^2}{\sigma^2} + \frac{\theta^2}{D^2} - 2\frac{\theta B}{D^2} + \frac{B^2}{D^2}\right]\right\} d\theta$$

$$= \int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^{-2}+D^{-2}}{2}\left[\theta^2 - 2\theta\frac{y\sigma^{-2}+BD^{-2}}{\sigma^{-2}+D^{-2}} + \left(\frac{y\sigma^{-2}+BD^{-2}}{\sigma^{-2}+D^{-2}}\right)^2\right]\right\} d\theta$$

$$\times \frac{1}{2\pi\sigma D} \exp\left\{-\frac{y^2}{2\sigma^2} - \frac{B^2}{2D^2} + \frac{(\sigma^{-2}y+BD^{-2})^2}{2(\sigma^{-2}+D^{-2})}\right\}$$

$$= \frac{\sqrt{2\pi}}{\sqrt{\sigma^{-2}+D^{-2}}} \times \frac{1}{2\pi\sigma D} \times$$

$$\exp\left\{-\frac{1}{2\sigma^2 D^2(\sigma^{-2}+D^{-2})}\left[\frac{y^2 D^2}{\sigma^2} + y^2 + \frac{B^2\sigma^2}{D^2} + B^2 - \sigma^2 D^2(\sigma^{-2}y + BD^{-2})^2\right]\right\}$$

Exercise 6.3 - Solution

$$= \frac{1}{\sqrt{2\pi(D^2+\sigma^2)}} \exp\{-\frac{1}{2(D^2+\sigma^2)}[\frac{y^2D^2}{\sigma^2} + y^2 + \frac{B^2\sigma^2}{D^2} + B^2$$

$$-\sigma^2D^2\sigma^{-4}y^2 - \sigma^2D^2B^2D^{-4} - 2\sigma^2D^2\sigma^{-2}yBD^{-2}]\}$$

Exercise 6.3 - Solution

$$= \frac{1}{\sqrt{2\pi(D^2+\sigma^2)}} \exp\{-\frac{1}{2(D^2+\sigma^2)}[\frac{y^2D^2}{\sigma^2} + y^2 + \frac{B^2\sigma^2}{D^2} + B^2$$

$$-\sigma^2D^2\sigma^{-4}y^2 - \sigma^2D^2B^2D^{-4} - 2\sigma^2D^2\sigma^{-2}yBD^{-2}]\}$$

$$= \frac{1}{\sqrt{2\pi(D^2+\sigma^2)}} \exp\{-\frac{1}{2(D^2+\sigma^2)}(y^2 + B^2 - 2yB)\}$$

$$= \frac{1}{\sqrt{2\pi(D^2+\sigma^2)}} \exp\{-\frac{1}{2(D^2+\sigma^2)}(y - B)^2\}$$

Therefore $y \mid x \sim N(B, D^2 + \sigma^2)$

Exercise 6.5

Observations $x = (x_1, x_2, \ldots, x_n)$ are made of independent random variables $X = (X_1, X_2, \ldots, X_n)$ with X_i having uniform distribution

$$f(x_i|\theta) = \frac{1}{\theta}; \quad 0 \leq x_i \leq \theta.$$

Assume that θ has an improper prior distribution

$$f(\theta) = \frac{1}{\theta}; \quad \theta \geq 0.$$

(a) Show that the posterior distribution of θ is given by

$$f(\theta|x) = \frac{nM^n}{\theta^{n+1}}; \quad \theta \geq M,$$

where $M = \max(x_1, x_2, \ldots, x_n)$.

(b) Show that θ has posterior expectation

$$E(\theta|x) = \frac{n}{n-1}M.$$

Solution

(a) Likelihood:
$f(x \mid \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(0 \leq x_i \leq \theta) = \frac{1}{\theta^n} I(0 \leq M \leq \theta)$,
where $M = max(x_1, \ldots, x_n)$
Prior: $f(\theta) = \frac{1}{\theta}$

Posterior: $f(\theta \mid x) = cf(x \mid \theta) f(\theta) = c \frac{1}{\theta^{n+1}} I(\theta \geq M)$

Solution

(a) Likelihood:
$f(x \mid \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(0 \leq x_i \leq \theta) = \frac{1}{\theta^n} I(0 \leq M \leq \theta)$,
where $M = max(x_1, \ldots, x_n)$
Prior: $f(\theta) = \frac{1}{\theta}$

Posterior: $f(\theta \mid x) = cf(x \mid \theta) f(\theta) = c\frac{1}{\theta^{n+1}} I(\theta \geq M)$

$$\int f(\theta \mid x) d\theta = 1 \Rightarrow c \int_M^{\infty} \frac{1}{\theta^{n+1}} d\theta = 1 \Rightarrow c \left[\frac{\theta^{-n}}{-n}\right]_M^{\infty} = 1 \Rightarrow$$

Solution

(a) Likelihood:
$f(x \mid \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(0 \leq x_i \leq \theta) = \frac{1}{\theta^n} I(0 \leq M \leq \theta)$,
where $M = max(x_1, \ldots, x_n)$
Prior: $f(\theta) = \frac{1}{\theta}$

Posterior: $f(\theta \mid x) = cf(x \mid \theta)f(\theta) = c\frac{1}{\theta^{n+1}} I(\theta \geq M)$

$$\int f(\theta \mid x) d\theta = 1 \Rightarrow c \int_M^\infty \frac{1}{\theta^{n+1}} d\theta = 1 \Rightarrow c \left[\frac{\theta^{-n}}{-n}\right]_M^\infty = 1 \Rightarrow$$

$$c\frac{M^{-n}}{n} = 1 \Rightarrow c = nM^n \Rightarrow f(\theta \mid x) = \frac{nM^n}{\theta^{n+1}}, \quad \theta \geq M.$$

Solution

(a) Likelihood:
$f(x \mid \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(0 \le x_i \le \theta) = \frac{1}{\theta^n} I(0 \le M \le \theta)$,
where $M = max(x_1, \ldots, x_n)$
Prior: $f(\theta) = \frac{1}{\theta}$

Posterior: $f(\theta \mid x) = cf(x \mid \theta) f(\theta) = c\frac{1}{\theta^{n+1}} I(\theta \ge M)$

$$\int f(\theta \mid x) d\theta = 1 \Rightarrow c \int_M^\infty \frac{1}{\theta^{n+1}} d\theta = 1 \Rightarrow c \left[\frac{\theta^{-n}}{-n}\right]_M^\infty = 1 \Rightarrow$$

$$c\frac{M^{-n}}{n} = 1 \Rightarrow c = nM^n \Rightarrow f(\theta \mid x) = \frac{nM^n}{\theta^{n+1}}, \quad \theta \ge M.$$

(b) Posterior Expectation:
$$E(\theta \mid x) = \int_M^\infty \theta f(\theta \mid x) d\theta = \int_M^\infty \frac{nM^n}{\theta^n} d\theta = nM^n \left[\frac{\theta^{-n+1}}{-n+1}\right]_M^\infty = \frac{nM}{n-1}$$

Exercise 6.5

(c) Verify the posterior probability:

$$\Pr(\theta > t\,M|x) = \frac{1}{t^n} \text{ for any } t \geq 1.$$

(d) A further, independent, observation Y is made from the same distribution as X. Show that the predictive distribution of Y has density

$$f(y|x) = \frac{1}{M}\left(\frac{n}{n+1}\right)\frac{1}{[\max(1, y/M)]^{n+1}}; \quad y \geq 0.$$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

(d) Likelihood of future observation: $f(y|\theta) = \frac{1}{\theta}; \;\; 0 \leq y \leq \theta.$

$$f(y \mid x) \;\; = \;\; \int f(y \mid \theta) f(\theta \mid x) d\theta = \int \frac{1}{\theta} I(\theta \geq y) \frac{nM^n}{\theta^{n+1}} I(\theta \geq M) d\theta$$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

(d) Likelihood of future observation: $f(y|\theta) = \frac{1}{\theta}; \;\; 0 \leq y \leq \theta.$

$$
\begin{aligned}
f(y \mid x) &= \int f(y \mid \theta) f(\theta \mid x) d\theta = \int \frac{1}{\theta} I(\theta \geq y) \frac{nM^n}{\theta^{n+1}} I(\theta \geq M) d\theta \\
&= \int_{max(M,y)}^{\infty} \frac{nM^n}{\theta^{n+2}} d\theta = \left[\frac{nM^n}{-n-1} \theta^{-n-1}\right]_{max(M,y)}^{\infty}
\end{aligned}
$$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

(d) Likelihood of future observation: $f(y|\theta) = \frac{1}{\theta}; \;\; 0 \leq y \leq \theta.$

$$\begin{aligned} f(y \mid x) &= \int f(y \mid \theta) f(\theta \mid x) d\theta = \int \frac{1}{\theta} I(\theta \geq y) \frac{nM^n}{\theta^{n+1}} I(\theta \geq M) d\theta \\ &= \int_{max(M,y)}^{\infty} \frac{nM^n}{\theta^{n+2}} d\theta = \left[\frac{nM^n}{-n-1}\theta^{-n-1}\right]_{max(M,y)}^{\infty} \\ &= \frac{nM^n}{n+1} \frac{1}{[max(M,y)]^{n+1}} \end{aligned}$$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

(d) Likelihood of future observation: $f(y|\theta) = \frac{1}{\theta}; \;\; 0 \leq y \leq \theta.$

$$
\begin{aligned}
f(y \mid x) &= \int f(y \mid \theta) f(\theta \mid x) d\theta = \int \frac{1}{\theta} I(\theta \geq y) \frac{nM^n}{\theta^{n+1}} I(\theta \geq M) d\theta \\
&= \int_{max(M,y)}^{\infty} \frac{nM^n}{\theta^{n+2}} d\theta = \left[\frac{nM^n}{-n-1}\theta^{-n-1}\right]_{max(M,y)}^{\infty} \\
&= \frac{nM^n}{n+1} \frac{1}{[max(M,y)]^{n+1}} \\
&= \frac{nM^n}{n+1} \frac{1}{M^{n+1}[max(M/M, y/M)]^{n+1}}
\end{aligned}
$$

Solution

(c) $\Pr(\theta > tM \mid x) = \int_{tM}^{\infty} \frac{nM^n}{\theta^{n+1}} d\theta = nM^n \left[\frac{\theta^{-n}}{-n}\right]_{tM}^{\infty} = \frac{1}{t^n}, \quad t \geq 1$

(d) Likelihood of future observation: $f(y|\theta) = \frac{1}{\theta}; \;\; 0 \leq y \leq \theta.$

$$
\begin{aligned}
f(y \mid x) &= \int f(y \mid \theta) f(\theta \mid x) d\theta = \int \frac{1}{\theta} I(\theta \geq y) \frac{nM^n}{\theta^{n+1}} I(\theta \geq M) d\theta \\
&= \int_{max(M,y)}^{\infty} \frac{nM^n}{\theta^{n+2}} d\theta = \left[\frac{nM^n}{-n-1}\theta^{-n-1}\right]_{max(M,y)}^{\infty} \\
&= \frac{nM^n}{n+1} \frac{1}{[max(M,y)]^{n+1}} \\
&= \frac{nM^n}{n+1} \frac{1}{M^{n+1}[max(M/M, y/M)]^{n+1}} \\
&= \frac{1}{M}\left(\frac{n}{n+1}\right) \frac{1}{[max(1, y/M)]^{n+1}}, \quad y \geq 0.
\end{aligned}
$$