Bayes

Beatrice Ngeno - 145614 Allan Aswani - 145613 Joseph Orina Makori - 145111
Chann Isaac - 145110 Cynthia Ombok Juma - 144892 Martin Mwangi - 138536
Gertrude Gichuhi - 106218 Peter Mburu - 145055

November 14, 2022

1.1 Let us consider a medical diagnostic test with 90% sensitivity and 80% specificity for a disease with incidence equal to 1 in 10,000 individuals. Use the Bayes theorem to calculate the probability that an individual has the disease when the test is positive and when the test is negative.

Solution

Given;

i) Sensitivity 90% i.e if one has the disease(D) the test is positive with probability 0.9 otherwise the test will be negative with probability 0.1.

ii) Specificity 80% i.e if one is healthy(H) the test will be negative with probability 0.8 otherwise the test will be positive with probability 0.2.

$$P(D) = 1/10000 P(H) = 1 - P(D) = 0.9999 P(+/D) = 0.9 \text{and} p(-/D) = 0.1 P(+/H) = 0.2 \text{and} P(-/H) = 0.8$$

a)the probability that an individual has the disease when the test is positive

$$p(D/+) = \frac{P(+/D)P(D)}{\sum_{k=1}^{n} P(D) + P(+/H)P(H)} \frac{0.90 * 0.0001}{(0.90 * 0.0001) + (0.2 * 0.9999)} = 0.0004$$

b)the probability that an individual has the disease when the test is negative

$$p(D/-) = \frac{P(-/D)P(D)}{\sum_{k=1}^{n} P(-/D_k)P(D_k)} p(D/-) = \frac{P(-/D)P(D)}{P(-/D)P(D) + P(-/H)P(H)} \frac{0.10 * 0.0001}{(0.10 * 0.0001) + (0.8 * 0.9999)} = 0.0000125$$

1.2 Let us consider the exponential distribution with density function $f(y/\theta) = \theta e^{-\theta y}$ and an i.i.d. sample $Y_i \sim exponential(\theta)$for i = 1, . . . , n.

a) Show that a gamma prior distribution is conjugate for θ.

Solution

i)Finding likelihood for the exponential distribution.

$$\prod_{i=1}^{n} f(y/\theta) = \prod \theta e^{-\theta y} \prod_{i=1}^{n} f(y/\theta) = \theta^n e^{-\theta \sum y_i}$$

ii) The pdf of gamma distribution.

$$\frac{1}{\gamma\alpha\beta^{\alpha}}e^{-\frac{\theta}{\beta}}\theta^{\alpha-1}$$

iii)The posterior distribution

$$\theta^{n}e^{-\theta\sum y}\times\frac{1}{\gamma\alpha\beta^{\alpha}}e^{-\frac{\theta}{\beta}}\theta^{\alpha-1}$$

Collecting the like terms the result is as follows:

$$f(\theta/y)=e^{-n\bar{y}\theta-\frac{\theta}{\beta}}\times\theta^{n+(\alpha-1)}$$

$$f(\theta/y)=\gamma(n+\alpha,n\bar{y}+\beta)$$

b) Calculate the posterior mean and variance under this conjugate prior.

posterior mean

$$n+\alpha\times n\bar{y}+\beta=n+\alpha\times n\bar{y}+\beta$$

posterior variance

$$n+\alpha\times(n\bar{y}+\beta)^{2}=n+\alpha\times(n\bar{y}+\beta)^{2}$$

c) For the following data

0.4 0.0 0.2 0.1 2.1 0.1 0.9 2.4 0.1 0.2

use the exponential distribution and (i) Plot the posterior distribution for gamma prior parameters a = b = 0.001

From the data given we need to obtain the mean so as to calculate the posterior parameters.

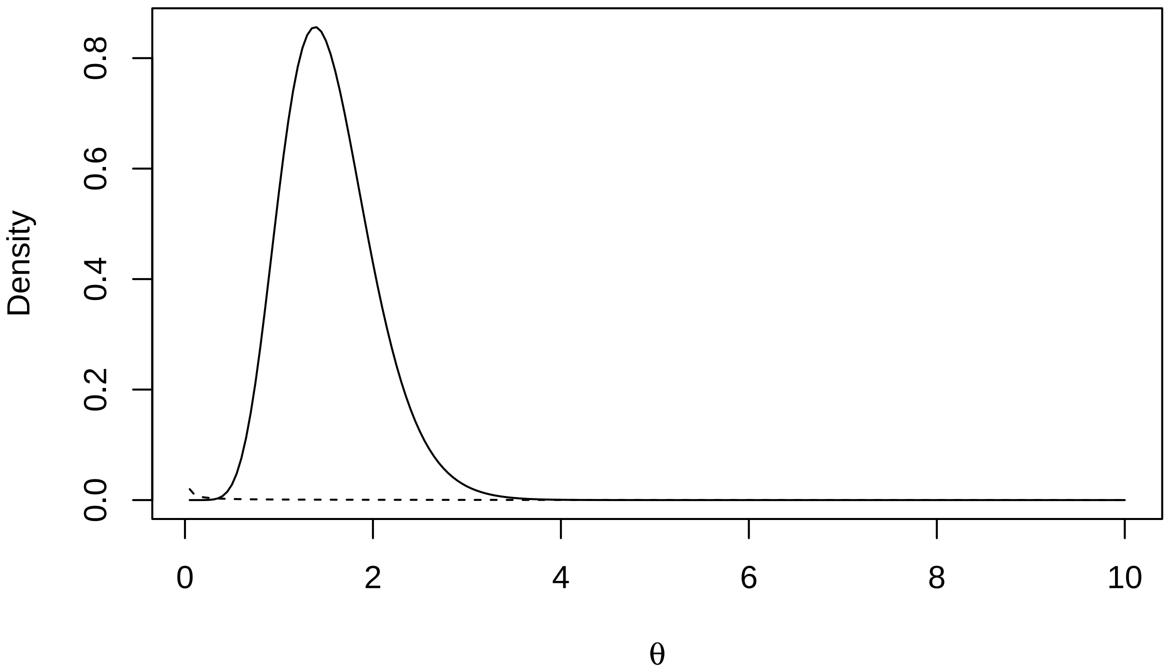
```
data<-c(0.4,0.0,0.2,0.1,2.1,0.1,0.9,2.4,0.1,0.2)
y_bar=mean(data)
y_bar
```

```
## [1] 0.65
```

$$\text{Posterior alpha}=n+\alpha\text{posterior beta}=n\bar{y}+\beta$$

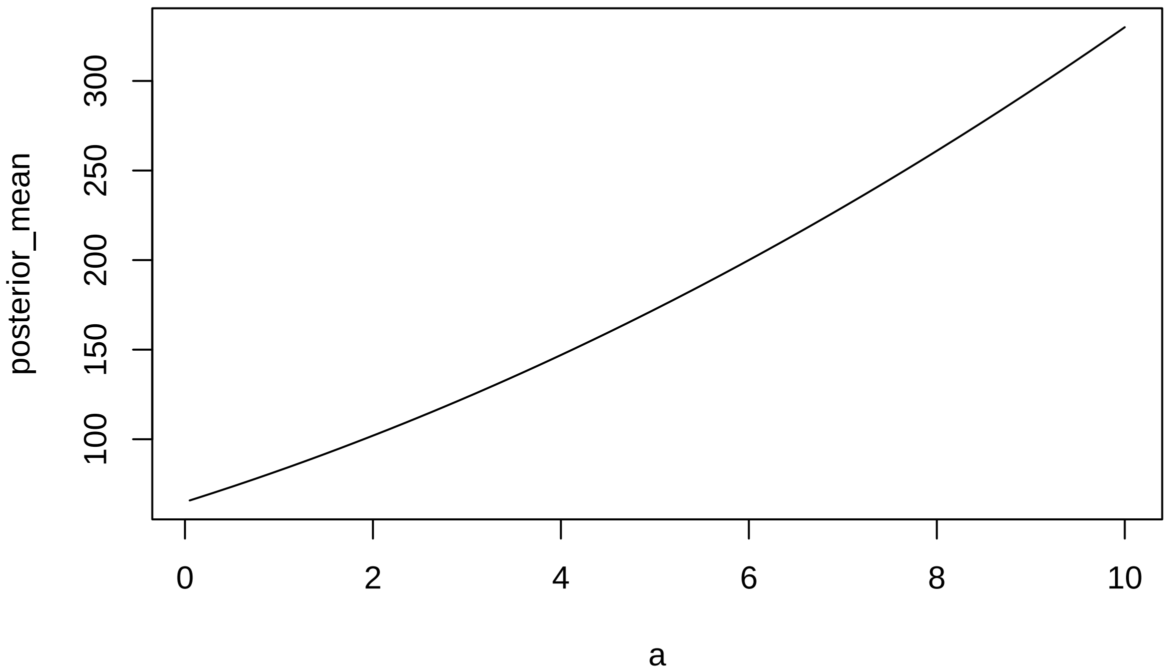
$$\text{Posterior alpha}=10+0.001=10.001\text{posterior beta}=10*0.65+0.001=6.501$$

```
theta<-seq(0.05,10.0,0.05)
prior<-dgamma(theta,0.001,0.001)
posterior<-dgamma(theta,10.001,6.501)
plot(theta,posterior,xlab=expression(theta),ylab="Density",type="l")
lines(theta,prior,lty=2)
```

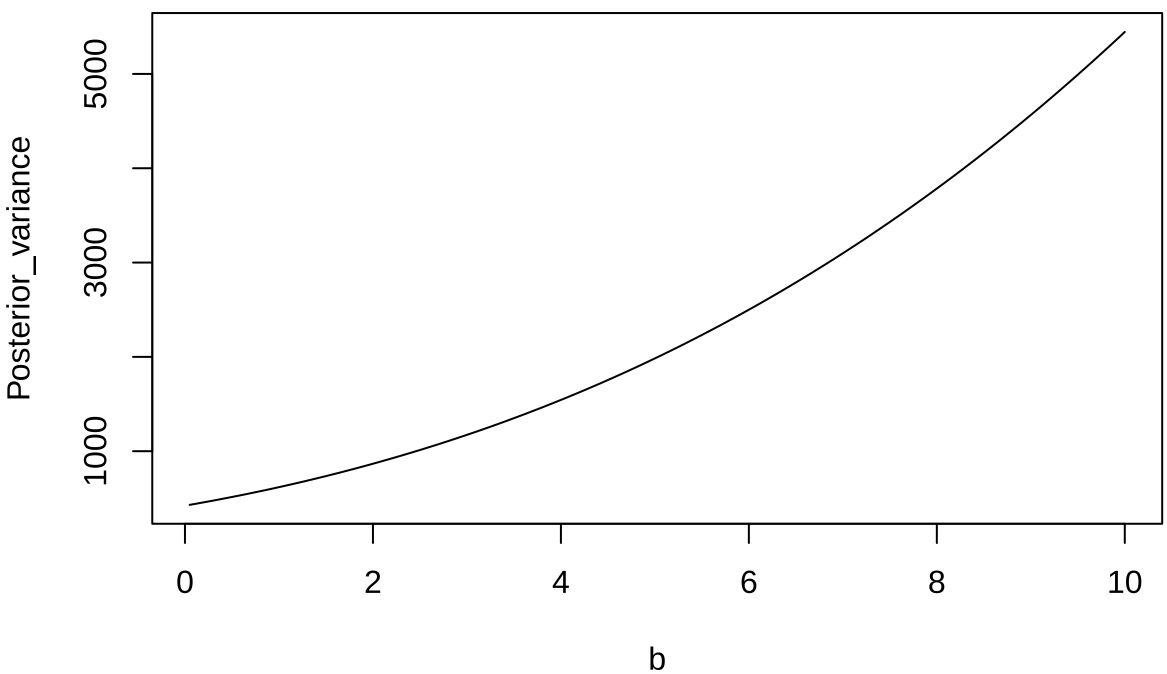


(ii) Perform sensitivity analysis for various values of a and b. Produce related plots depicting changes on the posterior mean and variance.

```
#prior parameters
set.seed(2455)
a=seq(0.05,10,0.05)
b=seq(0.05,10,0.05)
n<-10
y_bar=0.65
#Posterior mean
prior_mean=a*b
posterior_mean=(n+a)*(n*y_bar+b)
plot(posterior_mean~a,type="l")# similar to b since the sequence is the same.
```




```r
#Posterior variance
Posterior_variance=(n+a)*(n*y_bar+b)^2
prior_variance=a*b^2
plot(Posterior_variance~b,type="l")
```



1.3

In the exponential distribution, consider directly the mean parameter $\mu = \frac{1}{\theta}$

In the exponential distribution i.e

$$f(y|\theta) = \theta e^{-\theta y}$$

We consider directly that $\mu = \frac{1}{\theta}$

a) What is the conjugate prior for μ?

$$\mu = \frac{1}{\theta}$$

Consider μ for Gamma distribution as the prior for μ in exponential distribution

$$\mu^{Gamma} = \alpha\beta$$

$$\mu Exp = \frac{1}{\theta}$$

The posterior ;

$$\alpha\beta \times \frac{1}{\theta} \equiv \alpha\beta$$

$$\mu^{\text{Gamma pos}} = \alpha\beta$$

Therefore conjugate prior for μ is μ Gamma distribution

b) What is the posterior distribution for μ and θ under this setup?

The posterior distribution for μ and θ = Gamma

c) Is the posterior analysis under this approach equivalent to the corresponding one in Problem 1.2?

Posterior mean in Q1.2 = $(n + \alpha)(n\bar{y} = \beta)$ whereas in this case posterior mean = $\alpha\beta$

The analysis under this approach is not equivalent to that in Q1.2

1.4 For $Y_i \sim$ gamma(v, θ) assuming that V is known

a) Prove that the gamma distribution is a conjugate prior for θ

Solution

$$p(Y_i|\theta) = \frac{\beta^v \theta^{v-1} e^{-\beta\theta}}{\Gamma v}$$

$$\theta \sim \text{gamma}(v, \theta) = \frac{\beta^v \theta^{v-1} e^{-\beta\theta}}{\Gamma v}$$

Posterior $\propto$ likelihood $\times$ prior

Likelihood

$$\prod \frac{\beta^v \theta^{v-1} e^{-\beta\theta}}{\Gamma v} \;\propto\; \theta^{n(v-1)} e^{-n\beta\theta}$$

$$\propto \theta^{n(v-1)} e^{-n\beta\theta}\; \theta^{v-1} e{-}\beta\theta$$

collecting similar terms

$$\propto \theta^{nv-n+v-1} e^{-n\beta\theta-\beta\theta}$$

$$\propto \theta^{nv-n+v-1} e^{-\beta\theta(n+1)}$$

$$\sim gamma(nv - n + v, \theta(n+1))$$

b) Find the posterior mean and variance for θ

Mean

$$= \frac{nv - n + v}{\theta(n+a)}$$

Variance

$$= \frac{nv - n + v}{(\theta(n+1))^2}$$

c) Examine the effect on the posterior density of θ

(i) Of the known parameter v

$$= \frac{nv - n + v - v}{\theta} = \frac{n(v-1)}{\theta}$$

(ii) Of the sample size n

$$= \frac{\theta(n+1) - \theta}{\sqrt{\frac{(n-1)\theta + (n-1)\theta(n+1)}{n+n-2}}}$$

$$= \frac{\theta(n+1-1)}{(\frac{\theta(n-1)(1+n+1)}{2n-2})^{\frac{1}{2}}}$$

$$= \frac{\theta n}{(\frac{\theta(n+2)}{2})^{\frac{1}{2}}}$$

(iii) Of the prior parameters

v proportion

$$\frac{nv - n + v}{V}$$

θ proportion

$$\frac{\theta(n+1)}{\theta} = n + 1$$

1.5. Let us consider Yi (for i = 1, . . . , n) be an i.i.d. sample of categorical variables with n categories.

a. Show that a Dirichlet prior is conjugate for the probability of each category.

Dirichlet - Categorical Model

$$Y_i......Yn \sim Cat(\theta)$$

$$\theta \sim Dir(\alpha)$$

$$P(\theta|\alpha) \propto \prod_{j=1}^{m} \theta_j^{\alpha_j - 1}$$

let data be D $= (y_1 \dots y_n) \in \{1 \dots m\}$

$$P(D|\theta) = \prod_{i=1}^{n} P(Y_i = y_i|\theta) = \prod_{i=1}^{n} \theta_{y_i} = \prod_{i=1}^{n} \prod_{j=1}^{m} \theta_j^{I(y_i=j)}$$

$$P(D|\theta) = \prod_{j=1}^{m} \theta_j^{\sum I(y_i=j)}$$

Let the exponent be c

$$c = \sum I(y_i = j)$$

$$P(D|\theta) = \prod_{j=1}^{m} \theta_j^{c_j}$$

$$\text{Likelyhood } P(D|\theta) = \prod_{j=1}^{m} \theta_j^{c_j}$$

$$\text{Prior } P(\theta) \propto \prod_{j=1}^{m} \theta_j^{\alpha_j - 1}$$

$$\text{Posterior distribution } P(D|\theta) \propto P(D|\theta) * (P(\theta)$$

$$= \prod_{j=1}^{m} \theta_j^{c_j} * \prod_{j=1}^{m} \theta_j^{\alpha_j - 1}$$

$$= \prod_{j=1}^{m} \theta_j^{c_j + \alpha_j - 1} \propto Dir(\theta|c + \alpha)$$

This indicates that Dir is a conjugate Prior for categorical distribution

Pridictive distribution

$$P(y|D) = \int P(y|\theta, D) * P(\theta) d\theta = \int \theta_y P(\theta|D) = E(\theta_y|D)$$

$$= \frac{c_y + \alpha_y}{n + \alpha_0} \text{ where } \alpha_0 = \sum_{x=1}^{m} \alpha_y$$

b. Calculate the posterior mean and variance for the probability of each category.

In general

$$Y \sim Dir(\beta) \text{ where } \beta > 0$$

$$E(Y_i) = \frac{\beta_i}{\beta_0}$$

$$\text{From the posterior distribution } P(\theta|D) \propto Dir(\theta|c + \alpha)$$

$$\beta_i = c_y + \alpha_y$$

$$\therefore E(\theta_y|D) = \frac{c_y + \alpha_y}{n + \alpha_0}$$

$$\sigma^2(Y_i) = \frac{\beta_i(\beta_0 + \beta_i)}{\beta_0^2(\beta_0 + 1)}$$

$$= \frac{(c_y + \alpha_y)(n + \alpha_0 - (c_y + \alpha_y))}{(n + \alpha_0)^2(n + \alpha_0 + 1)}$$

c. Calculate the posterior correlations between the probabilities of two different categories

Covariance of β_j and β_k where $j \neq k$

$$Cov(\beta_j, \beta_k) = -\frac{\beta_j \beta_k}{\beta_0^2(\beta_0 + 1)}$$

$$= -\frac{(c_y + \alpha_y)^2}{(n + \alpha)^2(n + \alpha_0 + 1)}$$