MATH 640 - Assignment 3

Devere Anthony Weaver

2023-02-24

Problem 1

(a) To find the frequentist estimator for p,

$$p \sim Bin(y, n) = \binom{n}{y} p^y (1-p)^{n-y}$$

for $y = 0, 1, \ldots n$ and $0 < p < 1$.

Then,

$$p \approx \binom{116}{11} (0.0948)(.9052)^{105} \approx 0.1255.$$

(b) To compute the posterior distribution using $Beta(1, 10)$,

The posterior distribution is given by

$$f(p|y) = f(p)f(y|p).$$

Then the data distribution is,

$$f(y|p) \propto p^y (1-p)^{n-y},$$

and the prior distribution is,

$$f(p) \propto p^{a-1}(1-p)^{b-1}.$$

Then our posterior is,

$$f(p|y) \propto p^{y+a+1}(1-p)^{n-y+b-1} \sim Beta(12, 115).$$

Therefore $f(p|y) \sim Beta(12, 115)$.

(c) The posterior mean is given by,

$$E(p|y) = \frac{a+y}{a+b+n} = \frac{12}{127} \approx 0.09449.$$

The posterior variance is given by,

$$Var(p|y) = \frac{(a+y)(b+n-y)}{(a+b+n)^2(a+b+n+1)} \approx 8.1788 \times 10^{-5}.$$

The posterior mean is our Bayesian estimator.

(d) 95% credible interval

```
 # Data:
library(xtable)
 n <- 116  # Number of trials
 Y <- 11  # Number of successes
 a = 1
 b = 10

# Equal tail Intervals:
 A    <- Y+a
 B    <- n-Y+b

 ci.level<-0.95
 quantile.low<-(1-ci.level)/2# \alpha/2=(1-ci.level)/2
 quantile.up<-1-quantile.low
 Q05  <- qbeta(quantile.low,A,B)
 Q95  <- qbeta(quantile.up,A,B)


 output <- cbind(Q05, Q95)
 output <- round(output,3)
 print(xtable(output))
```

Yielding the following C.I.

	Q05	Q95
1	0.05	0.15

Problem 2

(a) The prior distribution is given by $\lambda \sim Gamma(a, b)$ with mean 6 and standard deviation 2.

In order to solve for a and b, set up system of equations using $E(\lambda) = \frac{a}{b} = 6$ and $Var(\lambda) = \frac{a}{b^2} = 4$ and solve for a and b. Doing so yields $a = 9$ and $b = \frac{3}{2}$. Hence,

$$\lambda \sim Gamma\left(9, \frac{3}{2}\right).$$

(b) The posterior of the unknown parameter is given by,

$$f(\lambda|y) = f(\lambda)f(y|\lambda)$$

.

Then,

$$f(\lambda|y) = \lambda^{y+a+1}e^{-(b+1)\lambda}$$

.

Hence,

$$f(\lambda|y) \sim Gamma(a + y, b + 1) = Gamma(80, 2.5).$$

(c) The posterior mean is given by

$$E(\lambda|Y) = \frac{a+y}{b+1} = \frac{80}{2.5} = 32.$$

The posterior variance is given by,

$$Var(\lambda|y) = \frac{a}{b^2} = 12.8.$$

(d) 95% credible interval.

```
#library(xtable)
 # Data:
 n <- 100
 Y <- 71  # Number of defects
 a = 80
 b = 2.5

 ci.level<-0.95
 #ci.level= 1-\alpha, \alpha=1-ci.level
 quantile.low<-(1-ci.level)/2# \alpha/2=(1-ci.level)/2
 quantile.up<-1-quantile.low
 Q05  <- qgamma(quantile.low,a,b)
 Q95  <- qgamma(quantile.up,a,b)


 output <- cbind(Q05, Q95)
 output <- round(output,3)
 print(xtable(output, include.rownames=TRUE, include.colnames=TRUE)  )
 print(xtable(output))
```

Yielding the following C.I.,

	Q05	Q95
1	25.37	39.38

Problem 3

(a) Write the function.

```
beta_binom <- function(Y, N, a = 1, b = 1, prob = .95,
                       title = "Prior and Posterior Distribution")
{
  # beta_binom() - returns a list of posterior mean and variance values where
  # the prior distribution follows Beta and the posterior is Binomial

  # Arguments:
  # Y = number of success for Binomial
  # N = number of trials for Binomial
  # a, b are size and shape parameters for prior Beta distribution

  theta <- seq(0,1,length=10000)    # values of theta following beta distribution

  # only plotting one prior
  prior = dbeta(theta, a, b)
  posterior = dbeta(theta, Y + a, Y + b)

  # plot the actual priors
  label = c("Prior", "Posterior")
  plot.new()
  plot(theta,prior,
       type="l",
       col = 4,
       xlab=expression(theta),
       ylab="Prior",
       ylim = c(0, max(posterior))
       )
  lines(theta,prior,col=1)
  lines(theta, posterior, col = 2)
  title(main = title)
  legend("topright",label,col=c(1,2), lty=1, cex=.9)

  # compute posterior mean and posterior standard deviation
  post_mean <- (Y + a) / (N + a + b)
  post_var <- ((Y + a)*(N - Y + b)) / ((a + N + b)^2 * (a + N + b + 1))
  post_sd <- sqrt(post_var)
  post_mode <- (a - 1) / (a + b - 2)
  post_median <- (a - (1/3)) / (a + b - (2/3))

  # compute credible interval tails
  A <- Y + a
  B <- N - Y + b
  quantile.low<-(1-prob)/2# \alpha/2=(1-ci.level)/2
  quantile.up<-1-quantile.low
  Q05  <- qbeta(quantile.low, A, B)
  Q95  <- qbeta(quantile.up,A,B)
  output <- cbind(Q05, Q95)
  output <- round(output,3)

  # return posterior mean and variance
```

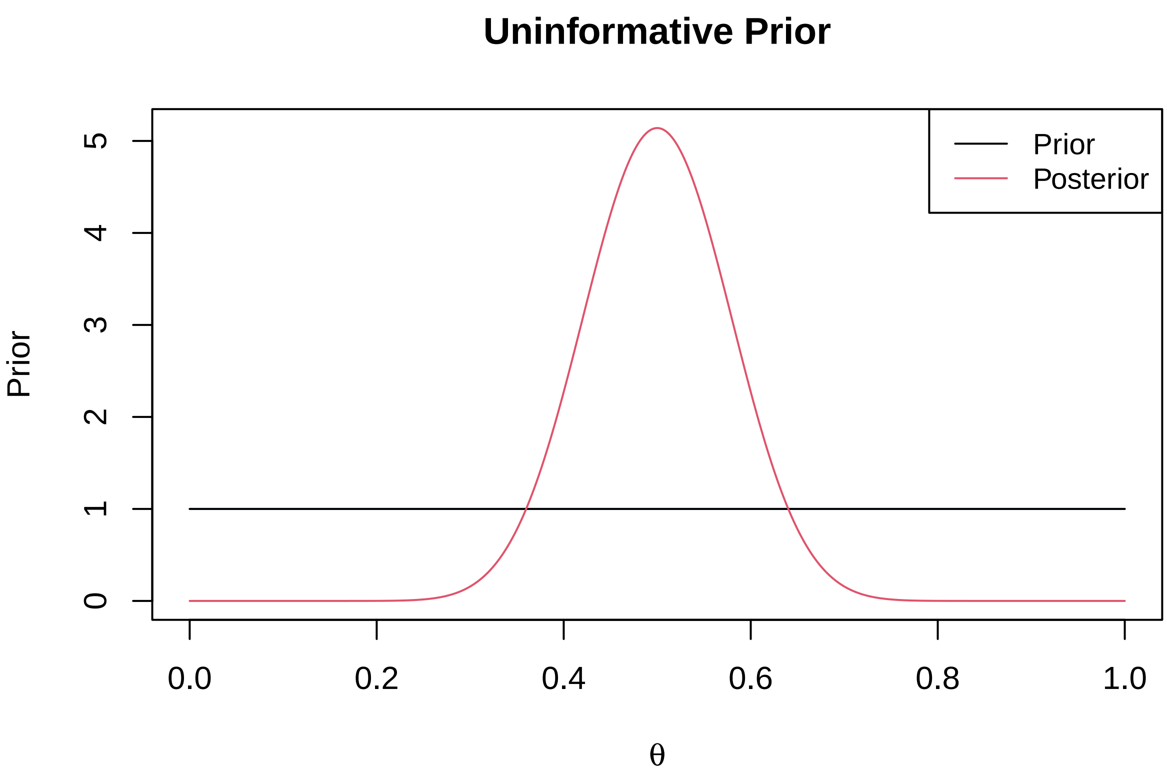
```
  mylist <- list("posterior mean" = post_mean, "posterior standard deviation" = post_sd,
                 "posterior mode" = post_mode, "posterior median" = post_median,
                 "endpoint" = output)
  return(mylist)
}
```

(b) The values of a and b that give little information about the unknown parameter are 1 and 1 respectively. This is simply a uniform distribution.

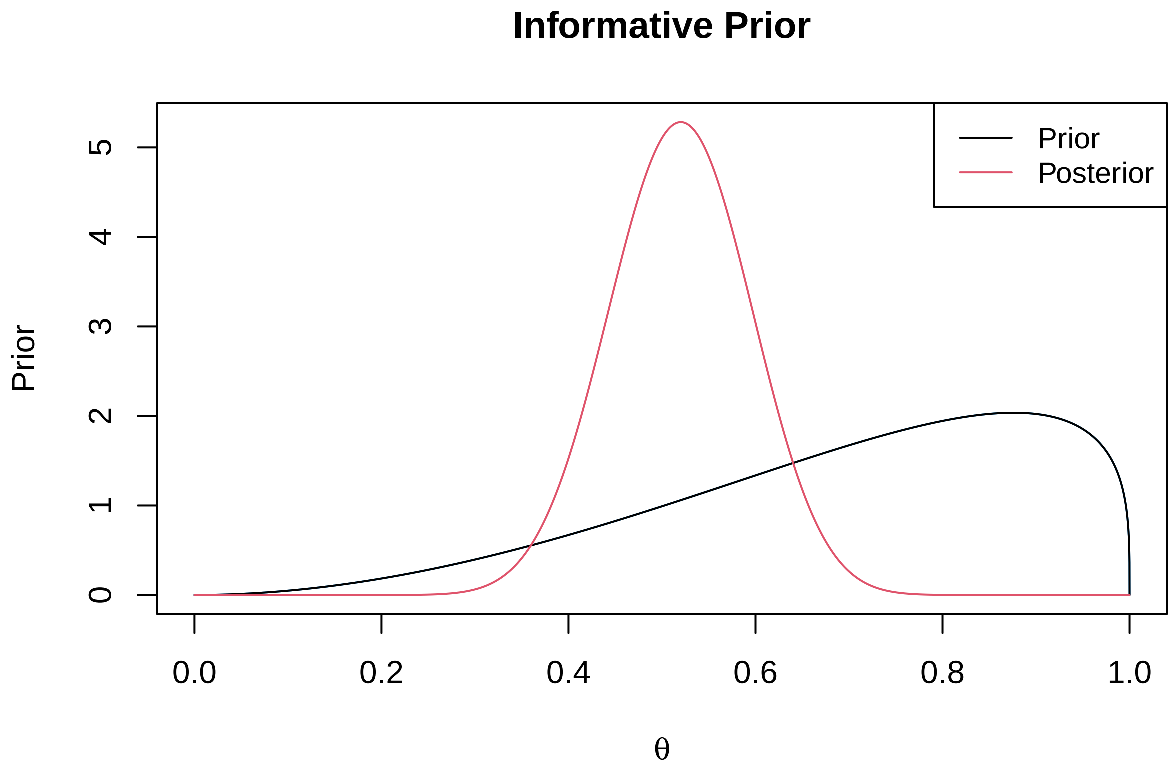
(c) The values of a and b that given prior mean 0.7 and prior standard deviation 0.2 are $a = 2.98$ and $b = 1.28$.

(d) Perform analysis with uninformative and informative prior.

```
# uninformative prior with default parameters
a <- beta_binom(20, 30, title = "Uninformative Prior")
```




```
# informative prior
b <- beta_binom(20, 30, a = 2.98, b = 1.28, title = "Informative Prior")
```



(e) A summary of the point and interval estimates

```
# point estimates
names <- c("Mean", "Median", "Mode", "Std. Dev.")
informative_estimates <- c(b$`posterior mean`, b$`posterior median`,
                           b$`posterior mode`, b$`posterior standard deviation`)
uninformative_estimates <- c(a$`posterior mean`, a$`posterior median`, a$`posterior mode`
                             , a$`posterior standard deviation`)
output <- data.frame(names, informative_estimates, uninformative_estimates)
output[is.na(output)] = 0
knitr::kable(output,
             col.names = c("Point Estimate", "Informative Prior", "Uninformative Prior"))
```

Point Estimate	Informative Prior	Uninformative Prior
Mean	0.6707531	0.6562500
Median	0.7365492	0.5000000
Mode	0.8761062	0.0000000
Std. Dev.	0.0791409	0.0826797

```r
# interval estimates
informative_interval <- data.frame(b$endpoint)
uninformative_interval <- data.frame(a$endpoint)
intervals <- rbind(informative_interval, uninformative_interval)
rownames(intervals) <- c("Informative", "Uninformative")
knitr::kable(intervals)
```

	Q05	Q95
Informative	0.507	0.815
Uninformative	0.486	0.808