STA 630 - Homework 2

Anthony Bernardi

February 17th, 2025

1 Problem 1 - Hoff 3.10

Change of variables problem. Let $\phi = g(\theta)$ be a monotone function with an inverse h, such that $\theta = h(\psi)$.

If $p_\theta(\theta)$ is the density of θ, then the density of ψ is given by the following.

$$p_\psi(\psi) = p_\theta(h(\psi)) \left| \frac{d}{d\psi} h(\psi) \right| \tag{1}$$

1.1 Part A

Let $\theta \sim beta(a, b)$, and let $\psi = log[\theta/(1-\theta)]$. Find the density of ψ and plot it for the case where $a = b = 1$.

We will start by finding the density and finding the form of $h(\psi)$ in this case.

$$\psi = g(\theta) = log\left[\frac{\theta}{1-\theta}\right] \tag{2}$$

$$exp(log(\theta/(1-\theta))) = \theta/(1-\theta) = exp(\psi) \tag{3}$$

$$\theta = \frac{exp(\psi)}{1 + exp(\psi)} \tag{4}$$

Now that we have the inverse, we can find its derivative and use this as the Jacobian in the change of variables formula.

$$h(\psi) = \frac{exp(\psi)}{1 + exp(\psi)} \tag{5}$$

$$\frac{d}{d\psi}h(\psi) = \frac{exp(\psi)}{(1+exp(\psi))^2} \tag{6}$$

This comes from a straightforward application of the quotient rule.

Finally, we can plug this into the change of variables formula to get the density of ψ, considering that θ is beta-distributed with parameters a and b.

We'll first delineate the density of θ with the accompanying parameters before applying the change of variables formula.

$$p_\theta(\theta) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)} \tag{7}$$

Where the denominator is, as typical, the beta function.

We can now write our density function with the change of variables formula, with respect to ψ.

$$p_\theta(h(\psi)) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{exp(\psi)}{1+exp(\psi)}\right)^{a-1}\left(1-\frac{exp(\psi)}{1+exp(\psi)}\right)^{b-1}\frac{exp(\psi)}{(1+exp(\psi))^2} \tag{8}$$

Simplifying the density function allows us to put it in a more interpretable form. After cancellations, we get the following.

$$p_\psi(\psi) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{exp(\psi)}{1+exp(\psi)}^{a+b} \tag{9}$$

This is the density function of ψ given that θ is beta-distributed with parameters a and b.

We can plot this density function using the following code, recognizing this as a beta density function.

Put another way, we recognize this as having a beta distribution with parameters a + 1 and b + 1.

Code and plot will go here...

1.2 Part B

Let $\theta \sim gamma(a, b)$. Let $\psi = log(\theta)$. Find the density of ψ.

We will do this problem in a very similar way, and start by finding the inverse function.

$$\psi = g(\theta) = log(\theta) \tag{10}$$

$$\theta = exp(\psi) \tag{11}$$

Here, $\theta = h(\psi)$, and we can find the derivative of this function.

$$\frac{d}{d\psi} h(\psi) = exp(\psi) \tag{12}$$

We can now plug this into the change of variables formula to get the density of ψ, using the gamma distribution as our model.

$$p_\theta(h(\psi)) = \frac{b^a}{\Gamma(a)} exp(\psi)^a \cdot exp(-b \cdot exp(\psi)) \tag{13}$$

After combining like terms, we get things in a more concise manner.

Code and plotting for the density function will go here...

2 Problem 2 - Hoff 3.12

2.1 Part a

Let $Y \sim binomial(n, \theta)$, obtain Jeffreys' prior for θ.

We will do this in the typical way, finding the square of the Fisher information.

We start by finding the likelihood, given by the following.

$$L(\theta) = \binom{n}{y} \theta^{\sum y} (1-\theta)^{n - \sum y} \tag{14}$$

We then find the log likelihood, given below.

$$l(\theta) = \sum y \cdot log(\theta) + (n - \sum y) \cdot log(1-\theta) \tag{15}$$

Taking the derivative of the log likelihood with respect to the parameter theta gives us the following.

$$\frac{d}{d\theta} l(\theta) = \frac{\sum y}{\theta} + nlog(1-\theta) - \sum ylog(\theta) \tag{16}$$

Taking the second derivative of the log likelihood with respect to the parameter theta gives us the following.

$$\frac{d^2}{d\theta^2}l(\theta) = -\frac{\sum y}{\theta^2} - \frac{n}{1-\theta} - \frac{n - \sum y}{1-\theta} \tag{17}$$

We can combine like terms and simplify to get the following.

$$\frac{d^2}{d\theta^2}l(\theta) = -\frac{\sum y}{\theta^2} + \frac{n - \sum y}{1-\theta}^2 \tag{18}$$

The Fisher information is given by the expectation of the square of the second derivative of the log likelihood, which we can write as the following.

$$I(\theta) = -E\left[\frac{d^2}{d\theta^2}l(\theta)\right] = -E\left[-\frac{\sum y}{\theta^2} + \frac{n - \sum y}{1-\theta}^2\right] \tag{19}$$

Which we can simplify to the following.

$$I(\theta) = \frac{n}{\theta(1-\theta)} \tag{20}$$

The Jeffreys' prior is given by the square root of the Fisher information, which we can write as the following.

$$\pi(\theta) = \sqrt{\frac{n}{\theta(1-\theta)}} \tag{21}$$

This is the Jeffreys' prior for the binomial distribution.

2.2 Part b

Reparameterize the binomial sampling model with $\psi = log\left(\frac{\theta}{1-\theta}\right)$. Obtain Jeffreys' prior for ψ.

As earlier, this starts with finding the inverse function, which we can write as the following.

$$\theta = \frac{exp(\psi)}{1 + exp(\psi)} \tag{22}$$

We then use the given density function to find the likelihood function, and eventually the log likelihood.

The density function is given by the following.

$$p(y|\psi) = \binom{n}{y} \left(exp(\psi y) \cdot (1 + exp(\psi))^{-n} \right) \tag{23}$$

We use this to find the likelihood function.

$$L(\psi) = \binom{n}{y} \left(exp(n\psi \sum y) \cdot (1 + exp(\psi))^{-n} \right) \tag{24}$$

This continues to the log likelihood.

$$l(\psi) = \sum y \cdot n\psi - n \cdot log(1 + exp(\psi)) \tag{25}$$

We can now take the derivative of the log likelihood with respect to the parameter ψ.

$$\frac{d}{d\psi} l(\psi) = n \sum y - n \cdot \frac{exp(\psi)}{1 + exp(\psi)} \tag{26}$$

Taking the second derivative of the log likelihood with respect to the parameter ψ gives us the following.

$$\frac{d^2}{d\psi^2} l(\psi) = -n \cdot \frac{exp(\psi)}{(1 + exp(\psi))^2} \tag{27}$$

We now use this to get our Fisher information.

$$I(\psi) = -E\left[\frac{d^2}{d\psi^2} l(\psi)\right] = -E\left[-n \cdot \frac{exp(\psi)}{(1 + exp(\psi))^2}\right] \tag{28}$$

$$I(\psi) = n \cdot E\left[\frac{exp(\psi)}{(1 + exp(\psi))^2}\right] \tag{29}$$

We then square root this to get the Jeffreys' prior.

$$\pi(\psi) = \sqrt{n \cdot \left[\frac{exp(\psi)}{(1 + exp(\psi))^2}\right]} \tag{30}$$

This is the Jeffreys' prior for the binomial distribution reparameterized with ψ.

2.3 Part c

Take the prior distribution and apply the change of variables formula to get the induced prior density for ψ. The density ought to be the same as the one derived in part b.

We can use the change of variables formula to get the induced prior density for ψ.

As before, we will start the change of variables formula by finding the function $h(\psi)$ and its derivative.

$$h(\psi) = \frac{exp(\psi)}{1 + exp(\psi)} \tag{31}$$

$$\frac{d}{d\psi} h(\psi) = \frac{exp(\psi)}{(1 + exp(\psi))^2} \tag{32}$$

We can now plug this into the change of variables formula to get the induced prior density for ψ, and match this up with our earlier example.

$$\pi(\psi) = \sqrt{n \cdot \left[\frac{exp(\psi)}{(1 + exp(\psi))^2} \right]} \tag{33}$$

This is the induced prior density for ψ, and it matches up with the Jeffreys' prior derived in part b.

3 Problem 3 - Hoff 5.1

Given the school files, use the following using the normal model.

Use the conjugate prior distribution, with the given parameters for mu, sigma, k, and v.

3.1 Part a

Find the posterior means and 95% confidence intervals for the mean and standard deviation parameters.

The following R code was used to find the posterior means and 95% confidence intervals for the mean and standard deviation parameters.

```
# reading in the data
setwd('/home/adbucks/Downloads')
school1 <- read.table("school1-1.dat.txt")
head(school1)
school2 <- read.table("school2.dat.txt")
school3 <- read.table("school3.dat.txt")

# read in the parameters for the posterior mean estimation
mu_0 <- 5
var_0 <- 4
k_0 <- 1
gamma_0 <- 2

# observed data
y1 <- school1$V1
n <- length(y1)
y1bar <- mean(y1)
s21 <- var(y1)

# posterior inference
kn <- k_0 + n
gamma_n <- gamma_0 + n
mu_n <- (k_0 * mu_0 + n*y1bar) / kn
s2n <- (gamma_0*var_0 + (n-1)
* s21 + k_0*n*(y1bar - mu_0)^2/(kn))/(gamma_n)

paste0("The Posterior mean is: " , round(mu_n,2))

# repeating this for each school
y2 <- school2$V1
n2 <- length(y2)
y2bar <- mean(y2)
s22 <- var(y2)

kn2 <- k_0 + n2
gamma_n2 <- gamma_0 + n2
mu_n2 <- (k_0 * mu_0 + n2*y2bar) / kn2
s2n2 <- (gamma_0*var_0 + (n2-1) *
```

```
s22 + k_0*n2*(y2bar - mu_0)^2/(kn2))/(gamma_n2)

paste0("The Posterior Mean for School 2 is: ",
       round(mu_n2 , 2))


#### school 3
y3 <- school3$V1
n3 <- length(y3)
y3bar <- mean(y3)
s23 <- var(y3)

kn3 <- k_0 + n3
gamma_n3 <- gamma_0 + n3
mu_n3 <- (k_0 * mu_0 + n3*y3bar) / kn3
s2n3 <- (gamma_0*var_0 + (n3-1) *
s23 + k_0*n3*(y3bar - mu_0)^2/(kn3))/(gamma_n3)

paste0("The Posterior Mean for School 3 is: ",
       round(mu_n3 , 2))

# confidence intervals for each now
margin1 <- qt(0.975 ,
              df = n - 1)*sqrt(s21)/sqrt(n)
l1 <- mu_n - margin1
u1 <- mu_n + margin1

paste0("The 95% Confidence Interval for the mean is: ",
       "(" , round(l1,2), ", ", round(u1,2), ")")

# now for the standard dev
library(MKinfer)
cisd1 <- sdCI(y1)

# school 2
margin2 <- qt(0.975 ,
              df = n2 - 1)*sqrt(s22)/sqrt(n2)
l2 <- mu_n2 - margin2
```

```
u2 <- mu_n2 + margin2

paste("The 95% Confidence Interval for the mean of
      school 2 is: ", "(", round(l2,2) , ", ",
      round(u2,2), ")")

(cisd2 <- sdCI(y2))

# school 3
margin3 <- qt(0.975,
              df = n3 - 1)*sqrt(s23)/sqrt(n3)
l3 <- mu_n3 - margin3
u3 <- mu_n3 + margin3

paste("The 95% Confidence Interval for the mean of
      school 3 is: ", "(", round(l3,2), ", ",
      round(u3,2), ")")

(cisd3 <- sdCI(y3))
```

This was done for each of the three schools, with the following results.

```
[1] "The Posterior mean is: 9.28"
[1] "The Posterior Mean for School 2 is: 6.95"
[1] "The Posterior Mean for School 3 is: 7.81"
[1] "The 95% Confidence Interval for the mean is: (7.61,10.97)"
[1] "The 95% Confidence Interval for the mean of school 2 is: (5.01 , 8.89)"
[1] "The 95% Confidence Interval for the mean of school 3 is: (6.04 , 9.58)"
```

We use the following output for the confidence intervals for the standard deviation.

```
> cisd1
[1] 3.034 5.405
> cisd2
[1] 3.469 6.348
> cisd3
[1] 2.876 5.524
```

3.2 Part b

4 Problem 4

Suppose $x_1, x_2, \ldots, x_n$ is a random sample from an exponential distribution with mean $\frac{1}{\theta}$.

4.1 Part a

Derive Jeffreys' prior for θ.

We will start by finding the likelihood function for the exponential distribution with the given mean.

$$L(\theta) = \theta^n \cdot exp(-\theta \sum x_i) \tag{34}$$

We use this to then find the log likelihood.

$$l(\theta) = n \cdot log(\theta) - \theta \sum x_i \tag{35}$$

We use this to take the derivative.

$$\frac{d}{d\theta} l(\theta) = \frac{n}{\theta} - \sum x_i \tag{36}$$

We then take the second derivative.

$$\frac{d^2}{d\theta^2} l(\theta) = -\frac{n}{\theta^2} \tag{37}$$

This then leads us to the Fisher information criteria.

$$I(\theta) = -E\left[-\frac{n}{\theta^2}\right] = \frac{n}{\theta^2} \tag{38}$$

Taking the square root gives us the Jeffreys' prior.

$$\pi(\theta) = \sqrt{\frac{n}{\theta^2}} = \frac{\sqrt{n}}{\theta} \tag{39}$$

This is the Jeffreys' prior for the exponential distribution from our prompt.

4.2 Part b

Derive the posterior distribution of θ using Jeffreys' prior.

We can start with taking our likelihood function, and implementing the Jeffreys' prior.

$$p(\theta|x) \propto \theta^n \cdot exp(-\theta \sum x_i) \cdot \frac{\sqrt{n}}{\theta} \tag{40}$$

We can simplify this to the following.

$$p(\theta|x) \propto \theta^{n-1} \cdot exp(-\theta \sum x_i) \cdot \sqrt{n} \tag{41}$$

This is the posterior distribution of θ using Jeffreys' prior, which we can say is proportional to the above.

We will want to note that this follows the gamma distribution, with parameters $n, \sum x_i$.

4.3 Part c

Derive the predictive distribution of a future observation z.

$$p(z|x) = \int p(z|\theta) \cdot p(\theta|x) d\theta \tag{42}$$

We note that this is the general form for the exponential, and we would proceed given the context of the problem in this way.