ST559 Midterm

Nick Sun

May 14, 2020

Abstract

This midterm has 4 problems

1 Q1

Amphipods are routinely used in sediment toxicity tests. Here we consider the survival of an amphipod is case and control toxicity samples.

	Deaths	Survived	Total
case	$x = 30$	$n - x = 70$	100
control	$y = 10$	$n - y = 90$	100
Total	$t = 40$	$2n - t = 160$	200

Let p_1 and p_2 be the respective probabilities of death of an individual amphipod when exposed to contaimined and cleaned sediments samples. Assume that $x \sim B(n, p_1)$ and $y \sim Bin(n, p_2)$.

1.1 part a.

Write down a joint prior distribution for p_1 and p_2.

answer It is known that a good noninformative prior for a binomial distribution parameter p is Beta$(1/2, 1/2)$. If we assume that p_1 is independent of p_2, then we can establish a joint prior for p_1 and p_2 that is just the product of the marginal priors.

$$\pi(p_1, p_2) \propto p_1^{1/2}(1 - p_1)^{1/2} p_2^{1/2}(1 - p_2)^{1/2}$$

1.2 part b.

Write down the likelihood function and derive the corresponding posterior distribution of p_1 and p_2 given the data.

The joint likelihood function for x and y is just the product of the individual likelihoods since we can reasonably assume that x is independent from y.

$$p(x, y) = \binom{n}{x} p_1^x (1 - p_1)^{n-x} \binom{n}{y} p_2^y (1 - p_2)^{n-y}$$

The posterior then can be obtained by multiplying this likelihood with the joint prior we obtained from before.

$$\begin{aligned}\pi(p_1, p_2|x, y) &= p(x, y|p_1, p_2)\pi(p_1, p_2)\\ &\propto p_1^x(1-p_1)^{n-x}p_2^y(1-p_2)^{n-y}p_1^{-1/2}(1-p_1)^{-1/2}p_2^{-1/2}(1-P_2)^{-1/2}\\ &\propto p_1^{x-1/2}(1-p_1)^{n-x-1/2}p_2^{y-1/2}(1-p_2)^{n-y-1/2}\end{aligned}$$

The notice that the joint posterior distribution is equal to $Beta(x+1/2, n-x+1/2)Beta(y+1/2, n-y+1/2)$. This joint posterior can be factored apart into separate function of p_1 and p_2.

answer

1.3 part c.

Compute the 95% Bayesian intervals for the following quantities:

1. $p_1 - p_2$
2. p_1/p_2
3. $(1-p_1)/(1-p_2)$

answer We can answer this by drawing random samples from the posterior distributions, computing the appropriate quantities, then finding the 2.5% and 97.5% quantiles

```r
p_1_post <- rbeta(10000, shape1 = 30.5, shape2 = 70.5)
p_2_post <- rbeta(10000, shape1 = 10.5, shape2 = 90.5)

pdiff <- p_1_post - p_2_post
quantile(pdiff, c(.025, .975))

pdiv <- p_1_post/p_2_post
quantile(pdiv, c(.025, .975))

pratio <- ((1 - p_1_post)/(1 - p_2_post))
quantile(pratio, c(.025, .975))
```

Bayesian CI for $p_1 - p_2$:

2.5%	97.5%
.0911	.3058

Bayesian CI for p_1/p_2:

2.5%	97.5%
1.61	5.96

Bayesian CI for $(1-p_1)/(1-p_2)$:

2.5%	97.5%
.667	.893

1.4 part d.

We are interested in the odds ratio θ and the odds product ϕ:

$$\theta = \frac{p_1(1-p_2)}{(1-p_1)p_2}$$
$$\phi = \frac{p_1 p_2}{(1-p_1)(1-p_2)}$$

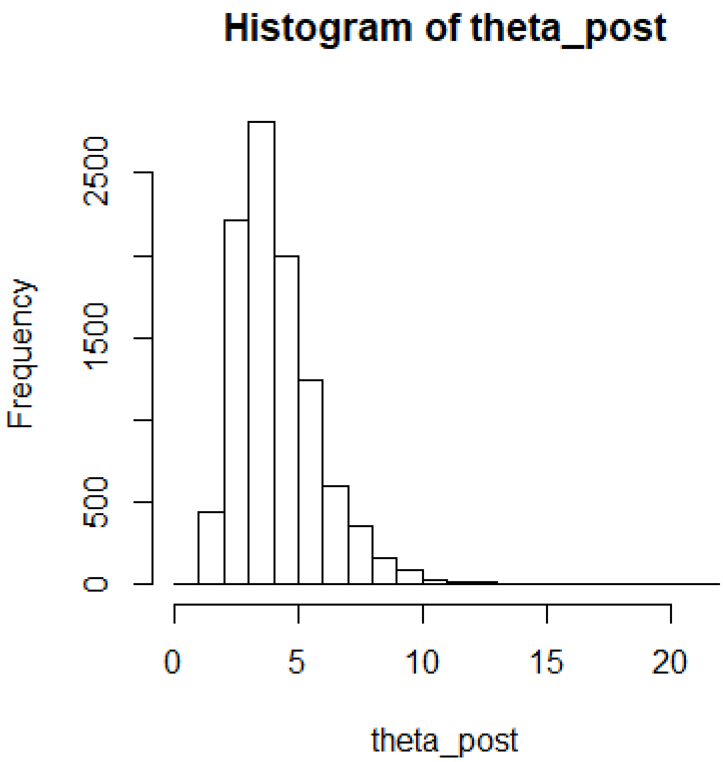
Summarize the posterior distribution of θ and ϕ given the data.

answer Assuming that we have drawn random observations from posterior distributions for p_1 and p_2, we can use the following code to compute the distributions of θ and ϕ.

```
theta_post <- (p_1_post*(1 - p_2_post))/((1 - p_1_post)*p_2_post)
hist(theta_post)
summary(theta_post)

phi_post <- (p_1_post*p_2_post)/((1 - p_1_post)*(1 - p_2_post))
hist(phi_post)
summary(phi_post)
```

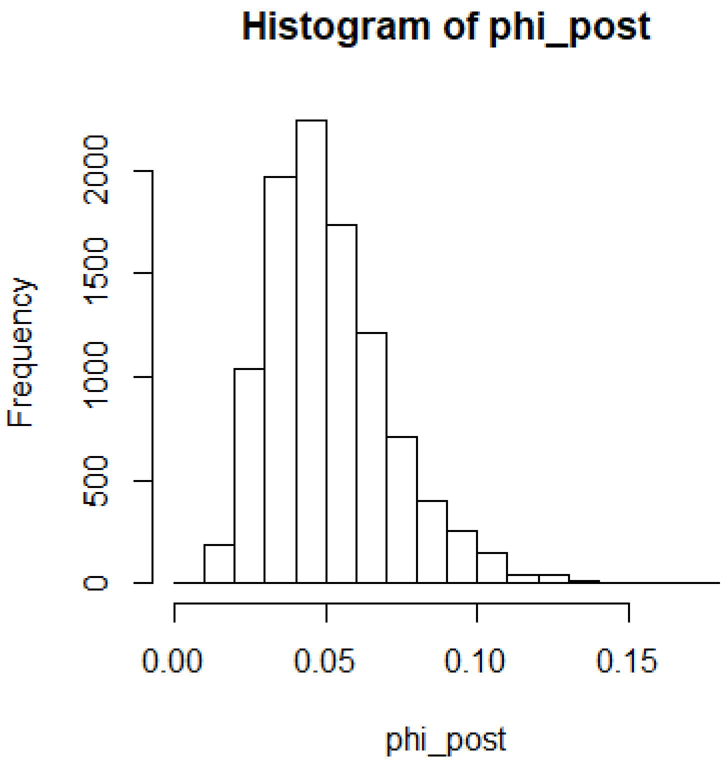
The distribution of θ looks like the following:



We can see that it is right skewed curve with the following numerical summary:

Min	Q1	Median	Mean	Q3	Max
.9616	2.9388	3.816	4.178	5.02	21.34

The distribution of ϕ looks like similar in shape to θ.



We have the following numerical summary.

Min	Q1	Median	Mean	Q3	Max
.009	.036	.048	.051	.062	.173

1.5 part e.

Compute the probability $P(\theta > 1|data)$

answer Running the following code, we get .9998 as our probability.

```
sum(theta_post > 1)/length(theta_post)
```

1.6 part f.

What conclusions can you draw about the test results?

answer We have significant evidence that the odds ratio is greater than 1, thus odds of death is much greater in the case group than the control group.

2 Q2

The following data is the amount of aluminum in 19 samples of pottery at two kiln sites.

Summary statistics of the two samples

- $n_1 = 14$
- $\bar{x}_1 = 12.275$
- $s_1 = 1.31$
- $n_2 = 5$
- $\bar{x}_2 = 18.18$
- $s_2 = 1.78$

Assume that the amount of aluminum at site 1 has a normal distribution $N(\mu_1, \sigma_1^2)$ and site 2 has another normal distribution $N(\mu_2, \sigma_2^2)$. Answer the following questions.

2.1 part a.

Construct a 95% Bayesian intervals for:

1. The difference in means $\mu_1 - \mu_2$
2. The ratio of the two variances σ_1^2/σ_2^2

In order to answer this question, we have to use the following facts on the posterior distributions for the Normal parameters which were derived on page 65 of *Bayesian Data Analysis*. If we define our data vector as X, the marginal posterior distributions for μ and σ^2 are:

- $\mu|\sigma^2, X \sim N(\bar{X}, \sigma^2/n)$
- $\sigma^2|X \sim \text{Scaled Inv-}\chi^2(n-1, s^2)$

If we make the reasonable assumptions that the samples are independent from one another, we can compute the posterior distributions for each μ and σ^2, sample from those posterior distributions, and then get the Bayesian intervals using those random samples.

There is no built-in sampling function in `R` for the inverse χ^2 distribution, but if we use the fact that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$, we can work around this.

If we random sample X from the χ^2_{n-1} distribution, we can get an inverse χ^2 distributed random observation σ^2 using $\sigma^2 = \frac{(n-1)s^2}{X}$. We then use this σ^2 value in our conditional Bayesian posterior for μ.

The `R` code that I used for this problem is provided here:

```
llanderyn <- c(14.4, 13.8, 14.6, 11.5, 13.8, 10.9, 10.1, 11.6,
    11.1, 13.4, 12.4, 13.1,12.7, 12.1)
island_thomas <- c(18.3, 15.8, 18.0, 18.0, 20.8)

n1 <- length(llanderyn)
xbar_1 <- mean(llanderyn)
s_1 <- sd(llanderyn)

sigma2_1_post <- ((n1-1)*s_1^2)/(rchisq(10000, df = n1 -1))
```

```
mu_1_post <- sapply(sigma2_1_post, FUN = function(x) rnorm(n = 1,
    mean = xbar_1, sd = sqrt(x/n1)))

n2 <- length(island_thomas)
xbar_2 <- mean(island_thomas)
s_2 <- sd(island_thomas)

sigma2_2_post <- ((n2-1)*s_2^2)/(rchisq(10000, df = n2 -1))
mu_2_post <- sapply(sigma2_2_post, FUN = function(x) rnorm(n = 1,
    mean = xbar_2, sd = sqrt(x/n2)))

mudiff_post <- mu_1_post - mu_2_post
quantile(mudiff_post, c(.025, .975))
```

The 95% Bayesian credible interval we get for $\mu_1 - \mu_2$ is:

2.5%	97.5%
-8.24	-3.58

There are a few ways to compute the Bayesian posterior interval for $\frac{\sigma_1^2}{\sigma_2^2}$. The first way is to recognize that since the conditional posterior for σ_1^2 and σ_2^2 is scaled inverse-χ^2, we can reorganize the ratio of the variables into an F-distribution.

$$\begin{aligned}
\sigma_1^2|data &\sim \text{Scaled Inv-}\chi^2(n_1-1, s_1^2) \text{ so} \\
\frac{\sigma_1^2}{(n_1-1)s_1^2} &\sim \text{Inv-}\chi^2(n_1-1) \text{ and} \\
\frac{(n_1-1)s_1^2}{\sigma_1^2} &\sim \chi^2(n-1)
\end{aligned}$$

This also holds for σ_2^2. Therefore, the ratio of $\frac{\sigma_1^2}{\sigma_2^2}$ can be turned into a scaled F-distribution through the following steps.

$$\begin{aligned}
\frac{s_2^2/\sigma_2^2}{s_1^2/\sigma_1^2} &\sim F(n_2-1, n_1-1) \\
\frac{\sigma_1^2}{\sigma_2^2}\left(\frac{s_2^2}{s_1^2}\right) &\sim F(n_2-1, n_1-1) \\
\frac{\sigma_1^2}{\sigma_2^2} &\sim \frac{s_1^2}{s_2^2}F(n_2-1, n_1-1)
\end{aligned}$$

So now we can either simulate a lot of scaled inverse-χ^2 random variables for σ_1^2 and σ_2^2 which we have already done, simulate a lot of random observations from an F distribution and scale them, or just find the quantiles of this scaled F distribution. Any of the following produce approximately the same intervals:

```
s1_s2_ratio <- sigma2_1_post/sigma2_2_post
quantile(s1_s2_ratio, c(.025, .975))
# or...
s1_s2_ratio <- (s_1^2/s_2^2)*rf(length(sigma2_1_post), n2-1, n1-1)
```

```r
quantile(s1_s2_ratio, c(.025, .975))
# or...
(s_1^2/s_2^2)*(qf(c(.025, .975), n2-1, n1-1))
```

The 95% Bayesian credible interval we get then for σ_1^2/σ_2^2 is:

2.5%	97.5%
.07	2.4

2.2 part b.

What prior knowledge are you assuming in constructing these intervals?

answer First, we need to know first that the samples are independent from each other. We also assume that we have a noninformative prior on the random vector (μ, σ^2). *Bayesian Data Analysis* suggests using a prior proportional to $\frac{1}{\sigma^2}$.

We could have used a conjugate prior, such as the Normal-Scaled-inverse-χ^2 distribution which was mentioned in the textbook, but this would have required us to specify guesses for μ_1, μ_2, σ_1^2, and σ_2^2. Since we did not have this information available to us, it was probably in our best interest to let the data "speak for itself".

2.3 part c.

How does these results compare with classical t and F confidence intervals?

answer Using the base `R` functions `t.test` and `var.test`, we find that the t-test with the unequal variance assumption gives us a 95% frequentist confidence interval for $\mu_1 - \mu_2$ of

2.5%	97.5%
-7.8	-3.4

and a 95% frequentist inverval for σ_1^2/σ_2^2:

2.5%	97.5%
.069	2.42

Note that `var.test()` uses this formula to calculate the confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$:

$$\left[\frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2,n1-1,n2-1}}, \frac{s_1^2}{s_2^2}F_{\alpha/2,n2-1,n1-1}\right]$$

These intervals are *close* to our Bayesian intervals, but not exactly the same.

3 Q3

Samples are taken from twenty batches of geological materials which are mined for their commericial values. The amount of impurities found are found to have the following statistics:

1. $n = 20$
2. $\bar{x} = 51.56$
3. $s = 3.23$

Let's regard these as independent samples from a normal distribution with μ and σ^2. Find a 95% posterior credible interval for μ under each of the conditions.

3.1 part a.

The values of σ^2 is known to be 10 and our prior distribution for μ is normal with mean 60 and standard deviation 20.

answer In this scenario, the variance is known! This allows us to use the known posterior distribution for μ with known variance (which we previously saw in homework 2, question 8, part a.)

$$
\begin{aligned}
f(\bar{x}|\mu)f(\mu) &\propto exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right) exp\left(-\frac{n(\bar{x}-\mu)^2}{2\sigma^2}\right) \\
&\propto N\left(\frac{\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}, \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}\right)
\end{aligned}
$$

By inputting $\bar{x} = 51.445$, $n = 20$, $\sigma^2 = 10$, $\sigma_0^2 = 20^2$, $\mu_0 = 60$, we get the following distribution:

$$
\begin{aligned}
\mu|\text{data}, \sigma^2 &\sim N\left(\frac{\frac{20\bar{x}}{10}+\frac{60}{20^2}}{\frac{20}{10}+\frac{1}{20^2}}, \frac{1}{\frac{20}{10}+\frac{1}{20^2}}\right) \\
&\sim N(51.55, .499)
\end{aligned}
$$

We can compute a 95% credible interval using this posterior distribution: $51.55 \pm 1.96\sqrt{.499}$. This gives us:

2.5%	97.5%
50.07	52.84

3.2 part b.

The value of σ^2 is unknown. Our prior distribution for σ^2 is an inverse gamma distribution with mean 0.1 and standard deivation .05. Our conditional prior distribution for μ given σ^2 is normal with mean 60 and variance $40\sigma^2$.

answer There are a couple ways of solving this problem depending on which parameterization we use. The first step is solving a simple system of equations to find the parameters α and β for the inverse Gamma prior distribution using the mean and variance that we have been given. Thankfully, the inverse Gamma distribution has closed formulas for mean and variance.

$$\begin{aligned} E[\sigma^2] &= \frac{\beta}{\alpha-1} = .1 \\ V[\sigma^2] &= \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = .05^2 \\ &\rightarrow \pi(\sigma^2) = \text{Inv-Gamma}(6, 1/2) \end{aligned}$$

I decided to find the posterior distribution form for a Normal-Inv-χ^2 prior since the inverse Gamma prior we selected is equivalent to a scaled inverse χ^2 distribution through the following relation:

$$X \sim \text{Inv-Gamma}(\alpha/2, 1/2) = \text{Scaled Inv-}\chi^2(\alpha, 1/\alpha)$$

Since $\beta = 1/2$, we can instead represent the prior as Inv-Gamma$(6, 1/2)$ = Scaled Inv-$\chi^2(12, 1/12)$. Our joint prior distribution then is:

$$\begin{aligned} \pi(\mu, \sigma^2) &\propto N(\mu|\mu_0, 40\sigma^2) \times \text{Inv-Gamma}(\sigma^2|6, 1/2) \\ &\propto N(\mu|\mu_0, 40\sigma^2) \times \text{Inv-}\chi^2(\sigma^2|12, 1/12) \\ &\propto \sigma^{-1}(\sigma^2)^{-7} exp\left(-\frac{1}{2\sigma^2}\left(1 + \frac{1}{40}(\mu-\mu_0)^2\right)\right) \end{aligned}$$

The posterior distribution for this scenario is:

$$\begin{aligned} p(\mu, \sigma^2|data) &= N(\mu|\mu, 40\sigma^2) \times \text{Inv-}\chi^2(\sigma^2|6, \frac{1}{6})p(data|\mu, \sigma^2) \\ &\propto \sigma^{-1}(\sigma^2)^{-4} exp\left(-\frac{1 + \frac{(\mu-60)^2}{40}}{2\sigma^2}\right) \times (\sigma^2)^{-n/2} exp\left(-\frac{1}{2\sigma^2}\left(ns^2 + n(\bar{x}-\mu)^2\right)\right) \end{aligned}$$

This is a pretty complex posterior, but it is able to be shown[1]that this is equivalent to a Normal-Inv-$\chi^2(\mu_n, \kappa_n, \eta_n, \sigma_n^2)$ where:

- $\kappa_n = \kappa_0 + n = \frac{1}{40} + 20$
- $\eta_n = \eta_0 + n = 6 + \frac{1}{40}$
- $\mu_n = \frac{\kappa_0\mu_0 + n\bar{x}}{\kappa_n}$
- $\sigma_n^2 = \frac{1}{\eta_n}(\eta_0\sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0}{\kappa_0+n}(\mu_0-\bar{x})^2$
- $\eta_0 = 6$
- $\sigma_0^2 = \frac{1}{60}$

[1]See https://www.cs.ubc.ca/ murphyk/Papers/bayesGauss.pdf

- $\kappa_0 = \frac{1}{40}$

In short, the Normal-Inv-χ^2 prior is a conjugate prior! Now we have to integrate this posterior distribution over σ^2 to get the marginal posterior distribution for μ. This is also pretty complex, but thankfully someone has already done this for me and showed that the marginal posterior distribution for μ is a scaled and shifted t-distribution (see this wikipedia article on Bayesian inference for an unknown mean).

$$\begin{aligned} p(\mu|data) &\propto \left(1 + \frac{\kappa_n}{\eta_n \sigma_n^2}(\mu - \mu_n)^2\right)^{-(\eta_n+1)/2} \\ &\propto t_{\eta_n}(\mu|\mu_n, \sigma_n^2/\kappa_n) \end{aligned}$$

We can simulate an observation x from this location-scale t-distribution using $x = \mu_n + (\sigma_n^2/\kappa_n)T$ where T is an observation from the t_{η_n} distribution. Here is the relevant `R` code for running this simulation.

```
q3data <- c(44.3, 50.2, 51.7, 49.4, 50.6, 55, 53.5, 48.6, 48.8,
    53.3, 59.4, 51.4, 52.0, 51.9, 51.6, 48.3, 49.3, 54.1, 52.4, 53.1)
xbar <- mean(q3data)

mu0 <- 60
n <- length(q3data)
k0 <- 1/40
kn <- k0 + n
eta0 <- 12
etan <- eta0 + n
mu0 <- 60
mun <- (k0*mu0 + n*xbar)/kn
sigma2n <- (1/etan)*(eta0*(1/60) +
           (n - 1)*sd(q3data) +
           (n*k0)/(k0 + n)*(xbar - 60)^2)

t_etan <- rt(n = 10000, df = etan)
mu_post <- mun + t_etan*sqrt(sigma2n/kn)
quantile(mu_post, c(.025, .975))
```

The 95% posterior credible interval is found to be $(50.921, 52.014)$, which is close to what we got in part a.

4 Q4

Let $X_1, \ldots, X_n$ be an iid sample from a normal distribution with mean $\mu = \theta\sigma$, variances σ^2. Here, θ is known as the noncentrality parameter.

4.1 part a.

Show that the likelihood function can be written as

$$L(\theta,\sigma) \propto \sigma^{-n} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta\sigma)^2\right)$$

answer Let's do some algebra!

$$\begin{aligned}
\prod_{i=1}^{n} f(x_i|\theta,\sigma^2) &= \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x_i-\theta\sigma)^2}{2\sigma^2}\right)\\
&= (2\pi)^{-n/2}\sigma^{-n} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta\sigma)^2\right)\\
&\propto \sigma^{-n} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta\sigma)^2\right)
\end{aligned}$$

4.2 part b.

Derive the Fisher Information Matrix.

answer Define the log likelihood function as:

$$L(\theta,\sigma) = -n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta\sigma)^2$$

The Fisher Information Matrix then is:

$$I(\theta,\sigma) = -E\begin{pmatrix} \frac{\partial}{\partial\theta^2}L(\theta,\sigma) & \frac{\partial}{\partial\theta\sigma^2}L(\theta,\sigma) \\ \frac{\partial}{\partial\sigma^2\theta}L(\theta,\sigma) & \frac{\partial}{\partial\sigma^4}L(\theta,\sigma)\end{pmatrix}$$

This next part involves some calculus. Instead, of finding the information for the entire sample of n observations, I will find the information for one observation since we know from the properties of Fisher information that the information of an iid sample is just a multiple of the information of a single observation.

$$\begin{aligned}
\frac{\partial}{\partial\theta^2}L(\theta,\sigma) &= \frac{\partial}{\partial\theta}\frac{x-\theta\sigma}{\sigma} = -1\\
\frac{\partial}{\partial\theta\partial\sigma^2}L(\theta,\sigma) &= \frac{\partial}{\partial\sigma^2}\frac{x-\theta\sigma}{\sigma}\\
&= \frac{\partial}{\partial\sigma^2}\frac{x}{(\sigma^2)^{1/2}} = -\frac{x}{2\sigma^3}
\end{aligned}$$

We have to take the partial derivatives with respect to σ^2 as opposed to σ. If we define $\tau=\sigma^2$, we can rewrite the likelihood in terms of τ and find the partial derivatives using that reparameterized likelihood.

$$\begin{aligned}
\frac{\partial}{\partial\tau\partial\theta}L(\theta,\tau) &= \frac{\partial}{\partial\theta} - \frac{1}{2\tau} - \frac{2\theta x\tau^{1/2} - 2x^2}{4\tau^2} \\
&= \frac{\partial}{\partial\theta} - \frac{2x\theta\tau^{1/2}}{4\tau^2} \\
&= -\frac{x}{2\tau^{3/2}} = -\frac{x}{2\sigma^3} \\
\frac{\partial}{\partial\tau^2}L(\theta,\tau) &= \frac{\partial}{\partial\tau}\frac{1}{2\tau} + \frac{x^2}{2\tau^2} - \frac{x\theta}{2\tau^{3/2}} \\
&= -\frac{1}{2\tau^2} - \frac{x^2}{\tau^3} + \frac{3x\theta}{4\tau^{5/2}} \text{ (now substitute } \sigma^2 \text{ back in)} \\
&= -\frac{1}{2\sigma^4} - \frac{x^2}{\sigma^6} + \frac{3x\theta}{4\sigma^5} = \frac{3x\theta\sigma - 2\sigma^2 - 4x^2}{4\sigma^6}
\end{aligned}$$

Using the fact that $E[x] = \theta\sigma$ and $E[x^2] = \sigma^2 + \theta^2\sigma^2$, we can finally calculate the Fisher information matrix.

$$\begin{aligned}
-E\left[\frac{3x\theta\sigma - 2\sigma^2 - 4x^2}{4\sigma^6}\right] &= \frac{4\sigma^2 + 4\theta^2\sigma^2 + 2\sigma^2 - 3\theta^2\sigma^2}{4\sigma^6} \\
&= \frac{\sigma^2(6+\theta^2)}{4\sigma^6} \\
-E\left[-\frac{x}{2\sigma^3}\right] &= \frac{\theta\sigma}{2\sigma^3} = \frac{\theta}{2\sigma^2}
\end{aligned}$$

And finally...

$$I(\theta,\sigma^2) = -E\begin{pmatrix} 1 & \frac{\theta}{2\sigma^2} \\ \frac{\theta}{2\sigma^2} & \frac{6+\theta^2}{4\sigma^4} \end{pmatrix}$$

I also found the Fisher information matrix for θ and σ to double check my work. This is shown below.

$$\begin{aligned}
\frac{\partial}{\partial\theta}L(\theta,\sigma) &= \frac{x}{\sigma} - \theta \\
\frac{\partial}{\partial\theta^2}L(\theta,\sigma) &= -1 \\
\frac{\partial}{\partial\theta\sigma}L(\theta,\sigma) &= -\frac{x}{\sigma^2} \\
\frac{\partial}{\partial\sigma}L(\theta,\sigma) &= -\frac{1}{\sigma} + \frac{x^2}{\sigma^3} - \frac{x\theta}{\sigma^2} \\
\frac{\partial}{\partial\sigma\theta}L(\theta,\sigma) &= -\frac{x}{\sigma^2} \\
\frac{\partial}{\partial\sigma^2}L(\theta,\sigma) &= \frac{-3x^2 + 2x\theta\sigma + \sigma^2}{\sigma^4}
\end{aligned}$$

$$I(\theta,\sigma) = -E\begin{pmatrix} 1 & \frac{\theta}{\sigma} \\ \frac{\theta}{\sigma} & \frac{2+\theta^2}{\sigma^2} \end{pmatrix}$$

4.3 part c.

Derive the Jeffrey's prior for θ and σ

The determinant of the Fisher information matrix we found is:

$$\frac{6+\theta^2}{4\sigma^4} - \frac{\theta^2}{4\sigma^4} = \frac{3}{2}\left(\frac{1}{\sigma^4}\right)$$

Then, since Jeffrey's prior is defined as the square root of the determinant of the information matrix, we have:

$$\det(I(\theta,\sigma^2))^{1/2} \propto \frac{1}{\sigma^2}$$

However, if we are interested instead in Jeffrey's prior for θ and σ, we are not yet done. While Jeffrey's prior is invariant, we still have to multiply this by $\frac{\partial\sigma^2}{\partial\sigma}$ (*Bayesian Data Analysis, pg. 53*).

This gives us

$$\begin{aligned}\det(I(\theta,\sigma))^{1/2} &\propto \det(I(\theta,\sigma))^{1/2} \times |\frac{\partial\sigma^2}{\partial\sigma}| \\ &\propto \frac{1}{\sigma^2}2\sigma \\ &\propto \frac{1}{\sigma}\end{aligned}$$

This matches up with our Jeffrey's prior for θ and σ if we had calculated it directly from the information matrix for σ:

$$\det(I(\theta,\sigma))^{1/2} = \left(\frac{2+\theta^2}{\sigma^2} - \frac{\theta^2}{\sigma^2}\right)^{1/2} \propto \left(\frac{2}{\sigma^2}\right)^{1/2} \propto \frac{1}{\sigma}$$