---
title: "HWK1_STA6570"
author: "Julian Zapata-Hall"
date: "2024-03-17"
output: html_document

```{r, warning=FALSE, message=FALSE, comment=FALSE}
library(tinytex)
```
Questions

Question 1
1. Let θ be the true proportion of men in your community over the age of 40 with hypertension. Consider the following “thought experiment”:
(a) Though you may have little or no expertise in this area, give an initial point estimate of θ.
Before looking at any data I would have said $g_0=0.3$. After a quick look at sources (https://www.cdc.gov/nchs/products/databriefs/db364.htm#:~:text=Key%20findings,-Data%20from%20the&text=In%20survey%20period%202017%E2%80%932018,%25%20(60%20and%20over) in 2017-18 59.4% of men ages 40-59 and 75.2% and of men ages 60+ had hypertension in the USA. Additionally, from https://www.statista.com/statistics/241488/population-of-the-us-by-sex-and-age/) as of July 1, 2022 41.46 million USA residents were men ages 40-59 and 36.33 million USA residents were men ages 60+. Avery unscientific but likely more accurate point estimate would be $$g_1=\frac{41.46(0.594)+36.33(0.752)}{41.46+36.33}=\frac{51.9474}{77.79}=0.6677902$$.

(b) Now suppose a survey to estimate θ is established in your community, and of the first 5 randomly selected men, 4 are hypertensive. How does this information affect your initial estimate of θ? It does not modify it. The sample is small and a sample proportion of 0.8 for a sample size of 5 seems reasonable given our point estimate. In the bayesian philosophy it would be superior to include this data in the point estimate in some way.

Question 2
BDA3 Exercise 2.8. Hint: For Part b, see pages 7 and 41.
8. Suppose that $f(x|\theta,\sigma)$ is normal with mean θ and standard deviation σ.
(a) Suppose that σ is known, we will show that the Jeffreys prior (2.12) for θ is given by $$p(\theta) = 1 , \theta \in \mathbb{R}$$

By definition of Jeffreys prior, we have $p(\theta) \propto [I(\theta)]^{1/2}$, where $I(\theta)$ is the expected Fisher information in the model, namely, $-E_{x|\theta}(\frac{d^2}{d\theta^2} log (f(x|\theta,\sigma)))$. We proceed algebraically:

$$log(f(x|\theta,\sigma)) = log((\frac{1}{\sigma \sqrt{2\pi}})e^{\frac{-(x-\theta)^2}{2\sigma^2}}) = (\frac{-(x-\theta)^2}{2\sigma^2})+ log(\frac{1}{\sigma \sqrt{2\pi}})$$

Thus,
$$-E_{x|\theta}(\frac{d^2}{d\theta^2}(log (f(x|\theta,\sigma)))) = -E_{x|\theta}(\frac{d}{d\theta}(\frac{x-\theta}{\sigma^2})) = -E_{x|\theta}(-\frac{1}{\sigma^2}) = \frac{1}{\sigma^2}, \theta \in \mathbb{R}$$

Therefore,
$$p(\theta) \propto [I(\theta)]^{1/2} = [\frac{1}{\sigma^2}]^{1/2} = \frac{1}{\sigma}$$

Since σ is known, however:

$$p(\theta) = 1 , \theta \in \mathbb{R}$$

as desired.

(b) Next, suppose that θ is known, and show that the Jeffreys prior for σ is given by
$$p(\sigma) = \frac{1}{\sigma}, \sigma > 0$$.

By definition of Jeffreys prior, we have $p(\sigma) \propto [I(\sigma)]^{1/2}$, where $I(\sigma)$ is the expected Fisher information in the model, namely, $-E_{x|\sigma}(\frac{d^2}{d\sigma^2} log (f(x|\sigma)))$. We proceed algebraically:

$$log(f(x|\theta,\sigma)) = log((\frac{1}{\sigma \sqrt{2\pi}})e^{\frac{-(x-\theta)^2}{2\sigma^2}}) = (\frac{-(x-\theta)^2}{2\sigma^2})+ log(\frac{1}{\sigma \sqrt{2\pi}})$$

Thus,
$$-E_{x|\sigma}(\frac{d^2}{d\sigma^2}(log (f(x|\theta,\sigma)))) = -E_{x|\theta}(\frac{d}{d\sigma}(\frac{(x-\theta)^2}{\sigma^3}-\frac{1}{\sigma})) = E_{x|\sigma}(\frac{3(x-\theta)^2}{\sigma^4}-\frac{1}{\sigma^2}) = \frac{2\sigma^2}{\sigma^4} = \frac{2}{\sigma^2}, \sigma > 0$$

Therefore,
$$p(\sigma) \propto [I(\sigma)]^{1/2} = [\frac{2}{\sigma^2}]^{1/2} = \frac{\sqrt{2}}{\sigma}$$

Since σ is unknown, however:

$$p(\sigma) = \frac{1}{\sigma} , \sigma > 0$$

as desired.

(c) Finally, assume that θ and σ are both unknown, and show that the bivariate Jeffreys prior (2.15) is given by
$$p(\theta,\sigma) = \frac{1}{\sigma^2}, \theta \in \mathbb{R}, \sigma > 0,$$

a slightly different form than that obtained by simply multiplying the two individual Jeffreys priors obtained above.

Let $I(\theta,\sigma)_{ij}$ be as defined as in (2.15) with (θ,σ) instead of the parameter vector notation in the textbook which I could not replicate. It follows that $I(\theta,\sigma)_{11} = \frac{1}{\sigma^2}$ and $I(\theta,\sigma)_{22} = \frac{2}{\sigma^2}$.
$$I(\theta,\sigma)_{12} = \frac{1}{\sigma^2} = -E_{x|\theta}(\frac{d}{d\theta}(\frac{(x-\theta)^2}{\sigma^3}-\frac{1}{\sigma})) = E_{x|\theta}(\frac{2(x-\theta)}{\sigma^3})=0,\theta \in \mathbb{R}, \sigma > 0$$

Thus, by definition of a multi-parameter case:

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

where $|I(\theta,\sigma)|$ is the determinant of the matrix we just found the elements of. We proceed algebraically:
$$|I(\theta,\sigma)|^{1/2} = ((\frac{2}{\sigma^2})(\frac{1}{\sigma^2})-0)^{1/2} = \frac{\sqrt{2}}{\sigma^2}$$
Thus, $$p(\theta,\sigma) = \frac{1}{\sigma^2}, \theta \in \mathbb{R},\theta>0$$
as desired.

Question 3
BDA3 Exercise 2.9. In parts (a) and (b), you should construct the density plots using R. In addition, find the
exact (i.e. not using a normal approximation) posterior probability that the true percentage of Californians
who support the death penalty is less than 64%.

9. Setting parameters for a beta prior distribution: suppose your prior distribution for θ,
the proportion of Californians who support the death penalty, is beta with mean 0.6 and standard deviation 0.3.

(a) Determine the parameters α and β of your prior distribution. Sketch the prior density
function.

Given the mean of beta prior is 0.6, we know that $0.6 = \frac{\alpha}{\alpha+\beta}$.
Additionally, since the standard deviation is 0.3, $0.3 = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}$. Using the first equation we obtain
$\beta=\frac{2\alpha}{3}$. Using the second equation,
$$0.3 = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}=
\sqrt{\frac{\alpha(\frac{2\alpha}{3})}{(\alpha+(\frac{2\alpha}{3}))^2(\alpha+(\frac{2\alpha}{3})+1)}}= \sqrt{\frac{18\alpha^2}{125\alpha^3+75\alpha^2}}=
\sqrt{\frac{18}{125\alpha+75}}$$

Thus, $\alpha = (\frac{18}{0.3^2}-75)/125 = 1$ and therefore, $\beta=\frac{2\alpha}{3} = \frac{2}{3]$

Approximate Prior Density Plot

```{r}
plot(density(rbeta(1000000,1,2/3)))
```

(b) A random sample of 1000 Californians is taken, and 65% support the death penalty.
What are your posterior mean and variance for θ? Draw the posterior density function.

Note that $Y|\theta\sim Bin(1000,\theta)$. We want to find the posterior for θ given $Y=650$.
Note that by a common conjugacy result $\theta|Y \sim Beta(y +\alpha, n-y+\beta)$ in this case:
$\theta|Y=560 \sim Beta(651, 350+2/3)$

The posterior Mean is therefore $\frac{651}{651+\frac{1052}{3}}\approx 0.65$.
And the posterior variance is $\frac{651(\frac{1052}{3})}{(651+\frac{1052}{3})^2 (652+\frac{1052}{3})}=\frac{684852}{(\frac{3005}{3})^2 (3008)}\approx 0.000227$

Approximate Posterior Density Plot

```{r}
plot(density(rbeta(1000000,651,350+2/3)))
```

Additionally, will find the
exact (i.e. not using a normal approximation) posterior probability that the true percentage of Californians
who support the death penalty is less than 64%. A simple integral is calculated. The posterior pdf is integrated over the region $\theta \in (0,64)$. or the R code below is used.

```{r}
pbeta(.64, 651, 350+2/3,lower.tail = TRUE)
```

Thus, there is a posterior probability of approximately 0.2541 that the true percentage of Californians
who support the death penalty is less than 64%.