Name: Carson Slater
Course: Theory of Statistics 2
Instructor: Dr. Harvill

Homework 3

1.

Consider $X_1, \dots, X_n \overset{iid}{\sim} \text{Gamma}(\alpha, \beta)$, were α is known.

a.

The methods of moments estimator for β can be found by setting the population moments equal to the sample moments.

$$\begin{aligned} m_1 = \bar{X} &= \int_0^\infty x \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx = \mu_1 \\ \iff \bar{X} &= \alpha\beta \int_0^\infty x \frac{1}{\Gamma(\alpha+1)\beta^\alpha} x^{\alpha} e^{-\frac{x}{\beta}}\, dx \\ \iff \tilde{\beta} &= \frac{\bar{X}}{\alpha}. \end{aligned}$$

b.

We have that

$$\begin{aligned} \text{MSE}(\hat{\beta}) = \text{MSE}(\tilde{\beta}) &= \text{Var}\left(\hat{\beta}\right) + \left(\text{Bias}(\hat{\beta})\right)^2 \\ &= \text{Var}\left(\frac{\bar{X}}{\alpha}\right) + \left(E\left[\frac{\bar{X}}{\alpha}\right] - \beta\right)^2 \\ &= \frac{\beta^2}{\alpha n} + \underbrace{\left(\frac{\alpha\beta}{\alpha} - \beta\right)^2}_{\text{Equals } 0} \\ &= \frac{\beta^2}{\alpha n}. \end{aligned}$$

c.

Both estimators have the same mean squared error, so neither is preferable over the other.

2.

We consider $Y_1, \cdots, Y_n \overset{iid}{\sim} \text{Unif}(0, \theta)$ We know the MLE of θ is $\hat{\theta} = Y_{(n)}$, and the method of moments estimator for θ is $\tilde{\theta} = 2\bar{Y}$. We first find the MSE of $\tilde{\theta}$,

$$\begin{aligned}
\text{MSE}(\tilde{\theta}) E &\left[\left(2\bar{Y} - \theta\right)^2\right] \\
&= \frac{4}{n^2} E\left[\left(\sum_{i=1}^{n} y_i\right)^2\right] - \frac{4\theta}{n} E\left[\sum_{i=1}^{n} y_i\right] + \theta^2 \\
&= \frac{4}{n^2}\left[\frac{n(n-1)\theta^2}{4} + \frac{n\theta^2}{3}\right] - \frac{4\theta}{n}\left[\frac{n\theta}{2}\right] + \theta^2 \\
&= \frac{4}{n^2}\left[-\frac{n\theta^2}{4} + \frac{n\theta^2}{3}\right] \\
&= \frac{\theta^2}{12n}.
\end{aligned}$$

Then, the MSE of $\hat{\theta}$ is

$$\begin{aligned}
\text{MSE}(\tilde{\theta}) &= E[Y_{(n)}^2] - 2\theta E[Y_{(n)}] + \theta^2 \\
&= \frac{n}{n+2}\theta^2 - 2\theta\frac{n}{n+1}\theta + \theta^2 \\
&= \frac{\theta^2(n^2 + n - 2n^2 - 4n + n^2 + 3n + 2)}{(n+1)(n+2)} \\
&= \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}$$

Here, as $n \to \infty$, the $\text{MSE}(\hat{\theta})$ decreases an order of magnitude faster than the $\text{MSE}(\tilde{\theta})$. So the MLE is preferable to the method of moments estimator.

3.

We are given $X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$, where $0 \leq \theta \leq 0.5$. We know that the MLE of θ is $\hat{\theta} = \min(\bar{X}, 0.5)$, and the method of moments estimator is $\tilde{\theta} = \bar{X}$.

a.

Let $Y = \sum_{i=1}^n X_i$ as $\sum_{i=1}^n X_i$ is a sufficient statistic for θ. Let $\lfloor \cdot \rfloor$ denote the floor function.

$$\begin{aligned}
\text{MSE}(\hat{\theta}) &= E\left[(\hat{\theta} - \theta)^2\right] \\
&= \underbrace{\sum_{y=0}^{\lfloor n/2 \rfloor} \left(\frac{y}{n} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}}_{\text{When } Y/n < 0.5} + \underbrace{\sum_{y=\lfloor n/2 \rfloor + 1}^{n} \left(\frac{1}{2} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}}_{\text{When } Y/n \geq 0.5} \text{ (Since } Y \sim \text{Bin}(n, \theta))
\end{aligned}$$

$$\begin{aligned}
\text{MSE}(\tilde{\theta}) &= E\left[(\tilde{\theta} - \theta)^2\right] \\
&= \sum_{y=0}^{n} \left(\frac{y}{n} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}
\end{aligned}$$

b.

So when $\frac{Y}{n} < 0.5$ the estimators are the same, which means their mean squared errors are the same when $\frac{Y}{n} < 0.5$, which implies that for $Y \in [0, \frac{n}{2}]$, the MSE's of the MLE and method of moment's estimator are the same. To find the better estimator, we can take the difference of them to see if it is positive or negative.

$$\begin{aligned}
\text{MSE}(\tilde{\theta}) - \text{MSE}(\hat{\theta}) &= \left(\sum_{y=0}^{n} \left(\frac{y}{n} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y} \right) \\
&\quad - \left(\sum_{y=0}^{\lfloor n/2 \rfloor} \left(\frac{y}{n} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y} \right) + \left(\sum_{y=\lfloor n/2 \rfloor + 1}^{n} \left(\frac{1}{2} - \theta\right)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y} \right) \\
&= \sum_{y=\lfloor n/2 \rfloor + 1}^{n} \left(\left(\frac{y}{n} - \theta\right)^2 - \left(\frac{1}{2} - \theta\right)^2 \right) \binom{n}{y} \theta^y (1-\theta)^{n-y}
\end{aligned}$$

Every term in this summation is positive, as $\left(\left(\frac{y}{n} - \theta\right)^2 - \left(\frac{1}{2} - \theta\right)^2 \right) > 0$ for all $Y \in (\frac{n}{2}, n]$. So then we have that

$$\begin{aligned}
\text{MSE}(\tilde{\theta}) - \text{MSE}(\hat{\theta}) &> 0 \\
\text{MSE}(\tilde{\theta}) &> \text{MSE}(\hat{\theta}).
\end{aligned}$$

Hence, the MLE is preferred to the method of moments estimator.

4.

Suppose the random variables $Y_1, \dots, Y_n$ satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, 2, \dots, n,$$

where $x_1, \dots, x_n$ are constants and ϵ_i, $i = 1, 2 \dots, n$ are IID $\mathrm{N}(0, \sigma^2)$ such that $\sigma > 0$ and is unknown.

a.

The MLE of β, $\hat{\beta}$ is

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}.$$

We can show this is unbiased for β,

$$\begin{aligned} E[\hat{\beta}] &= E\left[\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}\right] = \frac{\sum_{i=1}^n x_i E[Y_i]}{\sum_{i=1}^n x_i^2} \\ &= \frac{\sum_{i=1}^n x_i^2 \beta}{\sum_{i=1}^n x_i^2} \quad (E[Y_i] = \beta x_i) \\ &= \beta. \end{aligned}$$

To find the $\mathrm{MSE}(\hat{\beta})$,

$$\begin{aligned} \mathrm{MSE}(\hat{\beta}) = \mathrm{Var}(\hat{\beta}) + (\mathrm{Bias}(\hat{\beta}))^2 &= \mathrm{Var}\left(\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}\right) \\ &= \sum_{i=1}^n \mathrm{Var}(c_i Y_i) \quad \left(\text{Let } \ c_i = \frac{x_i}{x_i^2}, \text{ and } Y_i\text{'s ind.}\right) \\ &= \sum_{i=1}^n c_i^2 \sigma^2 = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}. \end{aligned}$$

b.

Considering the estimator

$$\tilde{\beta} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i},$$

we can show this is unbiased with respect to β.

$$\begin{aligned} E[\tilde{\beta}] &= E\left[\frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}\right] \\ &= \frac{\sum_{i=1}^n x_i \beta}{\sum_{i=1}^n x_i} \quad (E[Y_i] = \beta x_i) \\ &= \beta. \end{aligned}$$

The MSE of $\tilde{\beta}$ is

$$\begin{aligned}
\text{MSE}(\tilde{\beta}) = \text{Var}(\tilde{\beta}) + (\text{Bias}(\tilde{\beta}))^2 = \text{Var}\left(\frac{\sum_{i=1}^n Y_i}{n\bar{x}}\right) \\
&= \frac{\sum_{i=1}^n \text{Var}(Y_i)}{(n\bar{x})^2} \quad (Y_i\text{'s ind.}) \\
&= \frac{n\sigma^2}{n^2\bar{x}^2} = \frac{\sigma^2}{n\bar{x}^2}.
\end{aligned}$$

c.

Considering the estimator

$$\beta^* = \frac{1}{n}\sum_{i=1}^n \left(\frac{Y_i}{x_i}\right),$$

we can show this is unbiased with respect to β.

$$\begin{aligned}
E[\beta^*] = E\left[\frac{1}{n}\sum_{i=1}^n \left(\frac{Y_i}{x_i}\right)\right] &= \frac{\sum_{i=1}^n E[Y_i]}{n^2\bar{x}} \\
&= \frac{\sum_{i=1}^n (x_i\beta)}{n^2\bar{x}} = \frac{n\bar{x}(n\beta)}{n^2\bar{x}} \\
&= \beta.
\end{aligned}$$

The MSE of β^* is

$$\begin{aligned}
\text{MSE}(\beta^*) = \text{Var}(\beta^*) + (\text{Bias}(\beta^*))^2 &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^n \left(\frac{Y_i}{x_i}\right)\right) \\
&= \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^n \left(\frac{Y_i}{x_i}\right)\right) = \frac{1}{n^2}\sum_{i=1}^n \text{Var}\left(\frac{Y_i}{x_i}\right) \quad (\text{Since } Y_i\text{'s ind.}) \\
&= \frac{1}{n^2}\sum_{i=1}^n \frac{\text{Var}(Y_i)}{x_i^2} = \frac{1}{n^2}\sum_{i=1}^n \frac{\sigma^2}{x_i^2}.
\end{aligned}$$

d.

We have that by Jensen's Inequality, $\text{MSE}(\beta^*) > \text{MSE}(\hat{\beta})$, we do not consider β^*. Comparing $\tilde{\beta}$ and $\hat{\beta}$, we observe their denominators and determine which one is greater.

$$\begin{aligned}
0 &\leq \sum_{i=1}^n (x_i - \bar{x})^2 \\
\implies \sum_{i=1}^n x_i^2 &\geq n\bar{x}^2 \\
\implies \text{MSE}(\hat{\beta}) &\leq \text{MSE}(\tilde{\beta}).
\end{aligned}$$

So the MLE of β would be the best unbiased estimator based on these comparisons.

5.

We have that $X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Unif}(-\theta, \theta)$, where $\theta > 0$. To find the best unbiased estimator for θ. Consider $T(\mathbf{X}) = |X_{(n)}|$. We know the joint PDF of $X|\theta$ can be written as

$$\begin{aligned} f(\boldsymbol{x}|\theta) &= \prod_{i=1}^{n} \left(\frac{1}{2\theta}\right) I[|x_i| < \theta] \\ &= \left(\frac{1}{2\theta}\right)^n I[x_{(n)} < \theta] \\ &= g(T(\boldsymbol{x})|\theta)h(\boldsymbol{x}) \end{aligned}$$

where $g(T(\boldsymbol{x})|\theta) = \frac{1}{2\theta} I[x_{(n)} < \theta]$ and $h(\boldsymbol{x}) = 1$. By the factorization theorem, $T(\mathbf{X}) = |X|_{(n)}$ is a sufficient statistic. Now, suppose there exists a function $g : \langle 0, \infty \rangle \mapsto \mathbb{R}$ such that $E[g(Y)] = 0$ for all values of θ, where $Y = |X_{(n)}|$. Then we have that

$$\begin{aligned} 0 &= E_\theta[g(Y)] \\ &= \frac{d}{d\theta} \int_0^\theta g(y) \frac{ny^{n-1}}{\theta^n} dy \quad \text{(From Properties of Order of Statistics)} \\ &= \left(\frac{1}{\theta^n}\right) \frac{d}{d\theta} \int_0^\theta g(y) n y^{n-1} dy + \left(\frac{d}{d\theta}\frac{1}{\theta^n}\right) \underbrace{\int_0^\theta g(y) n t^{n-1} dy}_{E[g(y)]=0} \quad \text{(Product Rule)} \\ &= \left(\frac{1}{\theta^n}\right) g(y) n y^{n-1} = \frac{n g(y)}{\theta}. \end{aligned}$$

Since $n/\theta > 0$, $g(Y) = 0$. So for every value of $\theta > 0$, $g(Y) = 0$. Hence $P(g(Y) = 0) = 1 \;\forall\; \theta > 0$. So then $|X|_{(n)}$ is a complete, sufficient statistic for θ. So by Lehmann–Scheffé theorem, if some function of $T(\boldsymbol{X})$ is unbiased then it is UMVUE. Also, there exists a theorem that states

> Let T be a complete sufficient statistic for a parameter, and let $\phi(T)$ be any estimator based only on T. Then $\phi(T)$ is the unique UMVUE of its expected value.

We have that $|X|_{(n)}$ is a complete, sufficient statistic for θ. Then

$$E[|X|_{(n)}] = \int_0^\theta x \frac{n x^{n-1}}{\theta^n} dx = \frac{n}{n+1}\theta$$

From here we can say that the UMVUE of θ is

$$\theta^* = \frac{n+1}{n}|X|_{(n)} = \frac{n+1}{n} T(\boldsymbol{X})$$

as $|X|_{(n)}$ is biased on its own.

6.

a.

We are given $X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$. Let $y = \sum_{i=1}^{n} x_i$. The log-likelihood of p is

$$\begin{aligned} \ell(p|\boldsymbol{x}) &= \log(L(p|\boldsymbol{x}) \\ &= \log\left(\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\right) \\ &= y\log(p) + (n-y)\log(1-p). \end{aligned}$$

To minimize $\ell(p|\boldsymbol{x})$,

$$\begin{aligned} 0 &\overset{\text{set}}{=} \frac{\partial}{\partial p}\left(y\log(p) + (n-y)\log(1-p)\right) \\ \iff \frac{y}{p} &= \frac{n-y}{1-p} \\ \iff y - py &= np - py \\ \iff y &= np \\ \implies \hat{p}_{\text{MLE}} &= \bar{x}. \end{aligned}$$

We can show this is the global maximum by showing $\left.\frac{\partial^2}{\partial p^2}\ell(p)\right|_{p=\hat{p}_{\text{MLE}}} < 0 \ \forall \ p$.

$$\begin{aligned} \left.\frac{\partial^2}{\partial p^2}\ell(p)\right|_{p=\hat{p}_{\text{MLE}}} &= \frac{\partial}{\partial p}\left(\frac{\sum_{k=1}^{n} x_k}{p} - \frac{\sum_{k=1}^{n}(1-x_k)}{(1-p)}\right) \\ &= -\left(\frac{\sum_{k=1}^{n} x_k}{p^2} + \frac{\sum_{k=1}^{n}(1-x_k)}{(1-p)^2}\right). \end{aligned}$$

So $\left.\frac{\partial^2}{\partial p^2}\ell(p)\right|_{p=\hat{p}_{\text{MLE}}} < 0 \ \forall \ p$.

We know that $f(\boldsymbol{x}|\theta)$ satisfies the conditions of the Cramér-Rao Lower Bound Theorem. So then we can write $\frac{\partial}{\partial p}\ell(p|\boldsymbol{x})$ as

$$\begin{aligned} \frac{\partial}{\partial p}\ell(p|\boldsymbol{x}) &= \frac{n\bar{x}}{p} - \frac{n(1-\bar{x})}{1-p} \\ &= \frac{n(\bar{x}-p)}{p(1-p)} \\ &= a(p)(w(\boldsymbol{x}) - \tau(p)) \end{aligned}$$

where $a(p) = \dfrac{n}{p(1-p)}$, $w(\boldsymbol{x}) = \bar{x}$ and $\tau(p) = p$. We know that $w(\boldsymbol{x}) = \bar{x}$ is unbiased w.r.t. $\tau(p)$, and so then $\hat{p}_{\text{MLE}} = \bar{x}$ attains the Cramér-Rao lower bound.

b.

Let $Y = X_1X_2X_3X_4$ be an estimator for $\tau(p) = p^4$. Using properties of expectation for independent random variables, we have that

$$E[Y] = E[X_1X_2X_3X_4] = E[X_1]E[X_2]E[X_3]E[X_4] = p^4.$$

So Y is unbiased for p^4. We also know that $T(\boldsymbol{x}) = \sum_{i=1}^n x_i$ is a complete sufficient statistic for p. Now we define $\phi(t) = E[Y|T(\boldsymbol{x})]$. Then Since Y is unbiased, then $E_p[\phi(t)] = \tau(p) = p^4$. To find the best unbiased estimator of p^4, we need to find $\phi(t)$.

$$
\begin{aligned}
\phi(t) &= E[Y|T(\boldsymbol{x})] \\
&= \sum_{y=0}^{1} y_i P(Y|T=t) \\
&= \sum_{y=0}^{1} y_i \frac{P(Y=y, T=t)}{P(T=t)} \\
&= \frac{P(Y=1, T=t)}{P(T=t)} \\
&= \frac{P(X_1X_2X_3X_4 = 1, \sum_{i=1}^n x_i = t)}{P(\sum_{i=1}^n x_i = t)} \\
&= \frac{P(X_1X_2X_3X_4 = 1)P(\sum_{i=5}^n x_i = 0)}{P(\sum_{i=1}^n x_i = t} \\
&= \frac{p^4\binom{n-4}{t-4}p^{t-4}(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}} \\
&= \frac{\binom{n-4}{t-4}}{\binom{n}{t}}, \quad t \geq 4.
\end{aligned}
$$

For $t < 4$, $X_1X_2X_3X_4 = 0$, so then the estimator is $E[X_1X_2X_3X_4|T=t] = 0$.

7.

We have a $X_1, \dots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$, where $\theta \sim \mathcal{N}(\mu, \tau^2)$.

a.

To find the joint PDF of $\bar{X}$ and θ, we first establish that $\bar{X}|\theta \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, which is known from sampling distributions of the sample mean from a normal population. Using this, we have that

$$\begin{aligned} f(\bar{x}|\theta) &= f(\bar{x}|\theta)f(\theta) \\ &= \frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}} \exp\left\{-\frac{1}{2}\left(\frac{(\bar{x}-\theta)^2}{\sigma^2/n} + \frac{(\theta-\mu)^2}{\tau^2}\right)\right\}. \end{aligned}$$

b.

To find $m(\bar{x}|\sigma^2, \mu, \tau^2)$, we can integrate the result in (a) over the parameter space, Θ.

$$m(\bar{x}|\sigma^2, \mu, \tau^2) = \int_\Theta \frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}} \exp\left\{-\frac{1}{2}\left(\frac{(\bar{x}-\theta)^2}{\sigma^2/n} + \frac{(\theta-\mu)^2}{\tau^2}\right)\right\} d\theta$$

It can be shown that

$$\begin{aligned} &\int_\Theta \frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}} \exp\left\{-\frac{1}{2}\left(\frac{(\bar{x}-\theta)^2}{\sigma^2/n} + \frac{(\theta-\mu)^2}{\tau^2}\right)\right\} d\theta \\ &= \underbrace{\frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}}}_{\text{Constant}} \exp\left\{-\frac{(\bar{x}-\mu)^2}{2(\sigma^2/n+\tau^2)}\right\} \int_\Theta \underbrace{\exp\left\{-\frac{(\sigma^2/n+\tau^2)}{2(\tau^2(\sigma^2/n))}\left(\theta - \frac{\tau^2\bar{x}+(\sigma^2/n)\mu}{(\sigma^2/n)+\tau^2}\right)^2\right\}}_{\text{Normal Kernel (Integrates to 1 with Constant } C)} \\ &\propto \exp\left\{-\frac{(\bar{x}-\mu)^2}{2(\sigma^2/n+\tau^2)}\right\}. \end{aligned}$$

So $m(\bar{x}|\sigma^2, \mu, \tau^2) \sim \mathcal{N}(\mu, (\sigma^2/n + \tau^2))$.

c.

To find the posterior distribution of $\theta|\bar{x}, \mu, \tau^2$, we first take the ratio of the joint distribution of $\bar{x}$ and θ and the marginal of $\bar{x}|\sigma^2, \mu, \tau^2$.

$$\begin{aligned} \pi(\theta|\bar{x}, \mu, \tau^2) &= \frac{f_{\bar{X},\theta}(\bar{X}, \theta)}{m(\bar{x}|\sigma^2, \mu, \tau^2)} \\ &= \frac{\frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}} \exp\left\{-\frac{1}{2}\left(\frac{(\bar{x}-\theta)^2}{\sigma^2/n} + \frac{(\theta-\mu)^2}{\tau^2}\right)\right\}}{\underbrace{\frac{1}{\sqrt{2\pi(\sigma^2/n+\tau^2)}} \exp\left\{-\frac{(\bar{x}-\mu)^2}{2(\sigma^2/n+\tau^2)}\right\}}_{\text{Function of } \bar{X}\text{, so constant w.r.t } \theta.}} \\ &\propto \frac{1}{\sqrt{2\pi(\sigma^2/n)\tau^2}} \exp\left\{-\frac{1}{2}\left(\frac{(\bar{x}-\theta)^2}{\sigma^2/n} + \frac{(\theta-\mu)^2}{\tau^2}\right)\right\} \\ &\propto \exp\left\{-\frac{[\tau^2+\sigma^2/n]\theta^2 - 2[\bar{x}\tau^2 + \mu\sigma^2/n]\theta}{2\tau^2(\sigma^2/n)}\right\} \end{aligned}$$

By completing the square, we have that

$$\pi(\theta|\bar{x},\mu,\tau^2) \propto \exp\left\{-\frac{\left(\theta - \dfrac{\bar{x}\tau^2+\mu\sigma^2/n}{\tau^2+\sigma^2/n}\right)^2}{\dfrac{2\tau^2(\sigma^2/n)}{\tau^2+\sigma^2/n}}\right\}.$$

which is the kernel of a normal distribution with mean and variance

$$\begin{aligned} E[\theta|\bar{x},\mu,\tau^2] &= \left(\frac{\tau^2}{\tau^2+\sigma^2/n}\right)\bar{x} + \left(\frac{\sigma^2/n}{\tau^2+\sigma^2/n}\right)\mu, \\ \operatorname{Var}\left[\theta|\bar{x},\mu,\tau^2\right] &= \frac{(\sigma^2/n)\tau^2}{\tau^2+\sigma^2/n}. \end{aligned}$$

d.

The risk function, $R(\theta,\delta^\pi)$, where δ^π is the squared error loss for the Bayes estimator $E[\theta|\bar{x}]$, is written as

$$\begin{aligned} R(\theta,\delta^\pi) &= \operatorname{MSE}(E[\theta|\bar{X}]) \\ &= \operatorname{Var}(E[\theta|\bar{X}]) + \left(\operatorname{Bias}(E[\theta|\bar{X}])\right)^2 \\ &= \operatorname{Var}\left(\left(\frac{\tau^2}{\tau^2+\sigma^2/n}\right)\bar{X} + \left(\frac{\sigma^2/n}{\tau^2+\sigma^2/n}\right)\mu\right) \\ &\quad + \left(E\left[\left(\frac{\tau^2}{\tau^2+\sigma^2/n}\right)\bar{X} + \left(\frac{\sigma^2/n}{\tau^2+\sigma^2/n}\right)\mu\right] - \theta\right)^2 \\ &= \left(\frac{\tau^2}{\tau^2+\sigma^2/n}\right)^2 \operatorname{Var}(\bar{X}) + \left(\frac{\tau^2 E[\bar{X}] + \mu(\sigma^2/n) - \theta\tau^2 - \theta(\sigma^2/n)}{\tau^2+\sigma^2/n}\right) \\ &= \frac{\sigma^2}{n}\left(\frac{\tau^2}{\tau^2+\sigma^2/n}\right)^2 + \left(\frac{(\sigma^4/n^2)(\mu-\theta)^2}{(\tau^2+\sigma^2/n)^2}\right). \end{aligned}$$

e.

The risk functions for $\pi_1(\theta)$ and $\pi_2(\theta)$ are

$$\pi_1(\theta) = \frac{\theta^2+10000}{10201} \qquad \pi_2(\theta) = \frac{\theta^2+1}{4}.$$

In Figure 1, we observe that inducing a prior distribution on θ with greater variance seems to make the Bayes estimator generally less risky for all true values of θ, but ultimately $\pi_1(\theta)$ (has smaller variance) is better for a specific values (specifically when $\theta \in (-1.709, 1.709)$) of the true parameter θ.

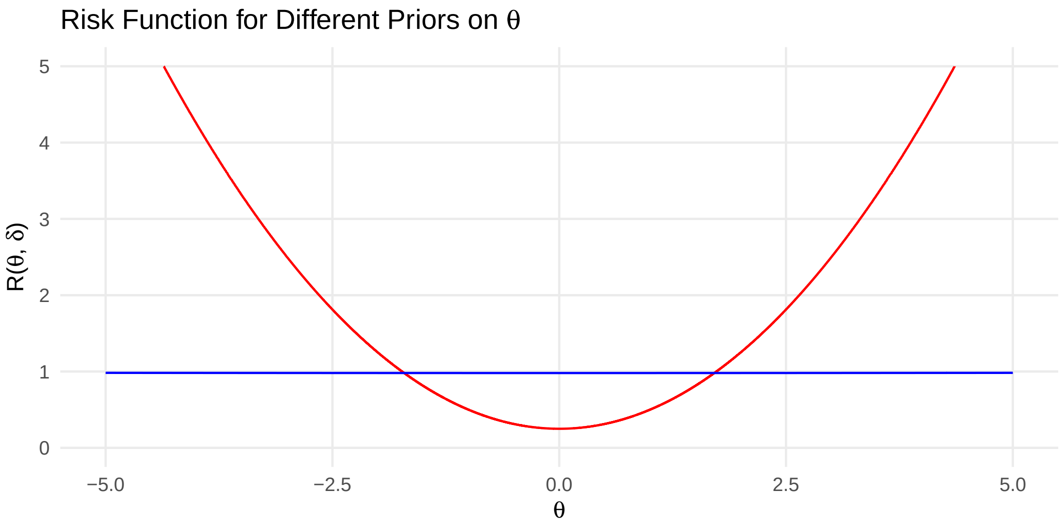

Figure 1: Risk functions for priors $\pi_1(\theta) \sim \mathcal{N}(0, 1)$ (red) and $\pi_2(\theta) \sim \mathcal{N}(0, 10)$ (blue).

R Code for Figure 1

```
# Carson Slater ------------------------------------------------------------
# Theory 2 Homework 3 ------------------------------------------------------

library("tidyverse"); library("latex2exp")
theme_set(theme_minimal())
theme_update(panel.grid.minor = element_blank())

# Parameter Definitions ----------------------------------------------------

n <- 1; sigma <- sqrt(1)
mu <- 0; theta <- seq(-10, 10, length.out = 100000)

# For pi_1 -----------------------------------------------------------------

tau1 <- 1
risk1 <- (tau1^2 / (tau1^2 + ((sigma^2) / n)))^2*(sigma^2/n) +
  ((sigma^4/n^2)*(mu - theta)^2) / (tau1^2 + (sigma^2/n))^2

plot1 <- cbind(tau1, risk1) |> as.data.frame()

# For pi_2 -----------------------------------------------------------------

tau2 <- 10
risk2 <- (tau2^2 / (tau2^2 + ((sigma^2) / n)))^2*((sigma^2) / n) +
  (((sigma^4) / (n^2))*(mu - theta)^2) / (tau2^2 + ((sigma^2) / n))^2
```

```r
plot2 <- cbind(tau2, risk1) |> as.data.frame()

# Plot --------------------------------------------------------------------

pdf(file = "risk_fns.pdf", width = 7, height = 3.5)
ggplot() +
  geom_line(data = plot1, aes(x = theta, y = risk1), col = "red") +
  geom_line(data = plot1, aes(x = theta, y = risk2), col = "blue") +
  scale_x_continuous(limits = c(-5, 5)) +
  scale_y_continuous(limits = c(0, 5)) +
  labs(title = latex2exp::TeX(r"(Risk Function for Different Priors on \theta)"),
       x = latex2exp::TeX(r"(\theta)"),
       y = latex2exp::TeX(r"(R(\theta, \delta))"))
dev.off()
```