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

Homework 2

1.

Let $X_1, \ldots, X_n$ be a random sample from a PDF

$$f(x|\theta) = \theta^{-2}, \quad 0 < \theta \leq x < \infty.$$

a.

Using the Factorization Theorem, we can find a sufficient statistic for θ. The joint PDF of $\boldsymbol{x}$ is,

$$\begin{aligned} f(\boldsymbol{x}|\theta) &= \prod_{i=1}^{n} \theta x_i^{-2} I_{[\theta,\infty)}(\boldsymbol{x}) \\ &= \theta^n \prod_{i=1}^{n} x_{i=1}^{-2} I_{[\theta,\infty)}(x_i)) \\ &= \frac{\theta^n}{\prod_{i=1}^{n} x_i^2} I_{[\theta,\infty)}(\min(X_i)) \\ &= g(T(\boldsymbol{x})|\theta)h(\boldsymbol{x}). \end{aligned}$$

Where $g(T(\boldsymbol{x}|\theta)) = I_{[\theta,\infty)}(\min(X_i))$ and $h(\boldsymbol{x}) = \dfrac{\theta^n}{\prod_{i=1}^{n} x_i^2}$. So $X_{(1)}$ is a sufficient statistic for θ.

b.

Using the method of moments, we equate the sample moments with the population moments to solve for $\hat{\theta}$.

$$\begin{aligned} m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i &= \int_{\theta}^{\infty} x\theta x^{-2}\, dx = \mu_1 \\ = \frac{1}{n}\sum_{i=1}^{n} x_i &= \theta \int_{\theta}^{\infty} \frac{1}{x}\, dx. \end{aligned}$$

This function diverges and so $\hat{\theta}_{\text{MOM}}$ does not exist.

c.

To find the maximum likelihood estimator for θ, we can find the global maximum for θ on the support of of the likelihood function.

$$\begin{aligned} L(\theta|\boldsymbol{x}) &= \prod_{i=1}^{n} \frac{\theta}{{}_{i}^{2}} I_{[\theta,\infty)}(x_i) \\ &= \frac{\theta^n}{\prod_{i=1}^{n} x_i^2} I_{[\theta,\infty)}(\min(x_i)). \end{aligned}$$

We can find the rate of change over θ,

$$\frac{\partial}{\partial\theta}\frac{\theta^n}{\prod_{i=1}^n x_i^2} = \frac{n\theta^{n-1}}{\prod_{i=1}^n x_i^2},$$

which is positive for all values of $\theta > 0$. We know that $0 < \theta \leq x_{(1)}$. Because $\frac{\theta^n}{\prod_{i=1}^n x_i^2}$ is monotonically increasing over θ on that interval, $\hat{\theta}_{MLE} = x_{(1)}$, because the likelihood is maximized at $x_{(1)}$.

2.

Let $Y_1, \ldots, Y_n$ be IID wit the PDF

$$f(y|\theta) = \theta^{-1}, \quad 0 \leq y \leq \theta, \quad \theta > 0.$$

a.

To find the method of moments estimator for θ, we we equate the sample moments with the population moments to solve for $\hat{\theta}$.

$$\begin{aligned} \mu_1 &= \frac{1}{\theta}\int_0^\theta y\,dy = \frac{1}{2}y^2\Big|_0^\theta = \frac{\theta^2}{2\theta} - 0 = \frac{\theta}{2}, \\ m_1 &= \bar{y}, \\ \implies \mu_1 &= \frac{\theta}{2} = \bar{y} = m_1 \\ \implies \hat{\theta}_{\text{MOM}} &= 2\bar{y} \end{aligned}$$

b.

To find the maximum likelihood estimator for θ, we can find the global maximum for θ on the support of the likelihood function. The joint PDF of Y is

$$\begin{aligned} f(\boldsymbol{y}|\theta) &= \prod_{i=1}^n \frac{1}{\theta} I_{[0,\theta]}(y_i) \\ &= \frac{1}{\theta^n} I_{[0,\theta]}(\max(y_{(1)})). \end{aligned}$$

To maximize the likelihood function, find the rate of change for the likelihood function to see where it is maximized. Normal calculus methods to find global maximums will not work to find the MLE as the support of the PDF of y is dependent on the parameter θ, the parameter we are trying to estimate.

$$\frac{\partial}{\partial\theta} L(\theta|\boldsymbol{y}) = \frac{-n}{\theta^{n-1}}.$$

Which is negative on $\theta > 0$. Therefore $\frac{-n}{\theta^{n-1}}$ is monotonically decreasing, and we know that $\theta \in [x_{(n)}, \infty)$. So $L(\theta|\boldsymbol{y})$ is maximized at $x_{(n)}$ because that is the point closest to $-\infty$ on a monotonically decreasing function.

3.

We have that $X_1, \dots, X_n$ have the common distribution

$$P(X_i \leq x|\alpha,\beta) = \begin{cases} 0 & \text{if } x < 0, \\ \left(\frac{x}{\beta}\right)^{\alpha} & 0 \leq x \leq \beta, \\ 1 & \text{if } x > \beta. \end{cases}$$

a.

Since the CDF is given, we can find the PDF,

$$\begin{aligned} f(x|\alpha,\beta) &= \frac{\partial}{\partial x}\left(\frac{x}{\beta}\right)^{\alpha} \\ &= \frac{\partial}{\partial x}\left(\frac{\alpha}{\beta^{\alpha}}\right)x^{\alpha-1}I_{[0,\beta]}(x) \end{aligned}$$

Using the PDF above, we can apply the Factorization Theorem, showing that

$$\begin{aligned} f(\boldsymbol{x}|\alpha,\beta) &= \prod_{i=1}^{n}\left(\frac{\alpha}{\beta^{\alpha}}\right)x^{\alpha-1}I_{[0,\beta]}(x_i) \\ &= \left(\frac{\alpha}{\beta^{\alpha}}\right)^{n} I_{[0,\beta]}(x_{(n)})\prod_{i=1}^{n}x^{\alpha-1} \\ &= g(T(\boldsymbol{x})|\theta)h(\boldsymbol{x}), \end{aligned}$$

where $g(T(\boldsymbol{x})|\theta) = \left(\frac{\alpha}{\beta^{\alpha}}\right)^{n} I_{[0,\beta]}(x_{(n)})\prod_{i=1}^{n}x^{\alpha}$, and $h(\boldsymbol{x}) = \prod_{i=1}^{n}\frac{1}{x_i}$. Here we have that $\prod_{i=1}^{n}I_{[0,\beta]}(x_i) = I_{[0,\beta]}(x_{(n)})$, as x is bounded between 0 and β, and β is unknown. So by the Factorization Theorem $T(\boldsymbol{X}) = \left(\prod_{i=1}^{n}x_i, X_{(n)}\right)$ is a sufficient statistic for (α,β).

b.

To find the MLE Of (α,β), we need to find the maximum value of the likelihood function over its support. Since β is a parameter found in the support, it not possible to use maximize the likelihood function with respect to β. We have that the likelihood function is

$$L(\alpha,\beta|\boldsymbol{x}) = \left(\frac{\alpha}{\beta^{\alpha}}\right)^{n} I_{[0,\beta]}(x_{(n)})\prod_{i=1}^{n}x^{\alpha-1}.$$

As β increases, the likelihood function decreases. We know that $x_{(n)} \leq \beta < \infty$, so then the likelihood is maximized with respect to β at $\hat{\beta}_{\text{MLE}} = x_{(n)}$, As for α, we can take the partial derivative with respect to α of the log-likelihood function (ℓ) and set it equal to zero, solving for α.

$$\begin{aligned}
0 &= \frac{\partial}{\partial \alpha}\ell(\alpha) \\
&= \frac{\partial}{\partial \alpha}\log\left(L(\alpha, \beta|\boldsymbol{x})\right) \\
&= \frac{\partial}{\partial \alpha}\left(n\log(\alpha) - n\alpha\log(\beta) + \log(I_{[0,\beta]}(x_{(n)})) + \alpha\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\log(x_i)\right) \\
&= \frac{n}{\alpha} - n\log(\beta) + \sum_{i=1}^{n}\log(x_i) \\
&= \frac{n}{\alpha} - n\log(x_{(n)}) + \sum_{i=1}^{n}\log(x_i) \quad (\text{Since } \hat{\beta}_{\text{MLE}} = x_{(n)}.) \\
\implies \frac{n}{\alpha} &= n\log(x_{(n)}) + \sum_{i=1}^{n}\log(x_i) \\
\iff \hat{\alpha}_{\text{MLE}} &= \left(\log(x_{(n)}) + \frac{1}{n}\sum_{i=1}^{n}\log(x_i)\right)^{-1}.
\end{aligned}$$

We can show that the likelihood is maximized at $\hat{\alpha}_{\text{MLE}}$ by showing that it is a global maximum. We can show this by showing $\frac{\partial^2}{\partial \alpha^2}\ell(\alpha) < 0$ when $\alpha = \hat{\alpha}_{\text{MLE}}$. We know that $n \in \mathbb{N}$. So then we have that

$$\begin{aligned}
\frac{\partial^2}{\partial \alpha^2}\ell(\alpha)\Big|_{\alpha=\hat{\alpha}_{\text{MLE}}} &= \frac{\partial}{\partial \alpha}\left(\frac{n}{\alpha} - n\log(\beta) + \sum_{i=1}^{n}\log(x_i)\right)\Big|_{\alpha=\hat{\alpha}_{\text{MLE}}} \\
&= -\frac{n}{\alpha^2}\Big|_{\alpha=\hat{\alpha}_{\text{MLE}}} < 0 \ \forall \ |\hat{\alpha}_{\text{MLE}}| \in (0, \infty).
\end{aligned}$$

Therefore,

$$(\hat{\alpha}_{\text{MLE}}, \hat{\beta}_{\text{MLE}}) = \left(\left(\log(x_{(n)}) + \frac{1}{n}\sum_{i=1}^{n}\log(x_i)\right)^{-1}, X_{(n)}\right).$$

4.

We have that $X_1, \ldots, X_n$ be a random sample from a population with PDF

$$P_\theta(X = x) = \theta^x(1-\theta), \quad \text{for } x = 0 \text{ or } 1 \quad 0 \le \theta \le 0.5.$$

a.

To find the method of moments estimator for θ, we can set the first sample moment equal to the first population moment.

$$\begin{aligned} m_1 = E[X] = \bar{x} = \theta &= \sum_{x=0}^{1} x\theta^x(1-\theta)^{1-x} = \mu_1, \\ \implies \hat{\theta}_{\text{MOM}} &= \bar{X}. \end{aligned}$$

Since $0 \le \theta \le 0.5$, we know then that $\hat{\theta}_{\text{MOM}} = \min(\bar{X}, 0.5)$.

b.

To find the MLE of θ, we can maximize the log-likelihood function by taking the derivative and setting it equal to zero to find an extremum. So we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial\theta}\ell(\theta) \\ &= \frac{\partial}{\partial\theta}\log\left(L(\theta|\boldsymbol{x})\right) \\ &= \frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^{n}\theta^{x_i}(1-\theta)^{1-x_i}\right) \\ &= \frac{\partial}{\partial\theta}\log\left(\theta^{\sum_{i=1}^{n}x_i}(1-\theta)^{n-\sum_{i=1}x_i}\right) \\ &= \frac{\partial}{\partial\theta}\left(\left(\sum_{i=1}^{n}x_i\right)\log(\theta) + \left(n - \sum_{i=1}^{n}x_i\right)\log(1-\theta)\right) \\ &= \frac{\sum_{i=1}^{n}x_i}{\theta} - \frac{n-\sum_{i=1}^{n}x_i}{1-\theta} \\ \iff \frac{1-\theta}{\theta} &= \frac{n-\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}x_i} \\ \iff \frac{1-\theta}{\theta} &= \frac{1-\frac{1}{n}\sum_{i=1}^{n}x_i}{\frac{1}{n}\sum_{i=1}^{n}x_i} \\ \iff \frac{1-\theta}{\theta} &= \frac{1-\bar{x}}{\bar{x}} \\ \implies \theta &= \bar{x}. \end{aligned}$$

We know that θ is bounded by 0 and 0.5, so we then would say that $\hat{\theta}_{\text{MLE}} = \min(\bar{X}, 0.5)$. To show $\theta = \bar{X}$ is a global maximum, we can show that $\frac{\partial^2}{\partial\theta^2}\ell(\theta)\Big|_{\theta=\hat{\theta}_{\text{MLE}}} < 0$ for all $\hat{\theta}_{\text{MLE}} \in [0, 0.5]$.

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\ell(\theta)\Big|_{\theta=\hat{\theta}_{\text{MLE}}} &= \frac{\partial}{\partial\theta}\left(\frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1-\theta}\right)\Big|_{\theta=\hat{\theta}_{\text{MLE}}} \\ &= -\frac{\sum_{i=1}^n x_i}{\theta^2} + \frac{n - \sum_{i=1}^n x_i}{(1-\theta)^2}\Big|_{\theta=\hat{\theta}_{\text{MLE}}} < 0 \ \forall \ |\hat{\theta}_{\text{MLE}}| \in [0, 0.5].\end{aligned}$$

Because $\left|\frac{\sum_{i=1}^n x_i}{\theta^2}\right| > \left|\frac{n - \sum_{i=1}^n x_i}{(1-\theta)^2}\right| \ \forall \ \theta \in [0, 0.5]$, we know that $\frac{\partial^2}{\partial\theta^2}\ell(\theta)\Big|_{\theta=\hat{\theta}_{\text{MLE}}} < 0$ over the viable values of θ. So $\hat{\theta}_{\text{MLE}} = \min(\bar{X}, 0.5)$.

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

5.

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.

We know that $Y_i = \beta x_i + \epsilon_i$, and that $_i$ is a constant, so then we have that $Y_i \sim \mathrm{N}(\beta x_i, \sigma^2)$ for all $i = 1, \dots, n$. So the joint distribution of each Y_i is

$$\begin{aligned} f(\boldsymbol{y}|\beta, \sigma^2) &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \beta x_i)^2\right\} \quad \text{(Since } Y_i\text{'s not IID)}. \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i^2 - 2\beta x_i y_i - \beta^2 x_i^2)^2\right\} \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{\beta^2}{2\sigma^2}\sum_{i=1}^n x_i^2\right\} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 + \frac{\beta}{\sigma^2}\sum_{i=1}^n x_i y_i\right\} \\ &= g(T(\boldsymbol{y})|\beta, \sigma^2) h(\boldsymbol{y}), \end{aligned}$$

where $g(T(\boldsymbol{y})|\beta, \sigma^2) = f(\boldsymbol{y}|\beta, \sigma^2)$ and $h(\boldsymbol{y}) = 1$. Then, by the Factorization Theorem, that a sufficient statistic for (β, σ^2) is

$$T(\boldsymbol{Y}) = \left(\sum_{i=1}^n x_i Y_i, \sum_{i=1}^n Y_i^2\right).$$

b.

We now will find the MLE of β We have it that the likelihood of β is

$$L(\beta|\boldsymbol{y}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{\beta^2}{2\sigma^2}\sum_{i=1}^n x_i^2\right\} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 + \frac{\beta}{\sigma^2}\sum_{i=1}^n x_i y_i\right\},$$

which yields the log-likelihood

$$\ell(\beta|\boldsymbol{y}, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{\beta^2}{2\sigma^2}\sum_{i=1}^n x_i^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 + \frac{\beta}{\sigma^2}\sum_{i=1}^n x_i y_i.$$

Maximizing $\ell(\beta|\boldsymbol{y}, \sigma^2)$, we have

$$\begin{aligned} 0 = \frac{\partial \ell(\beta)}{\partial \beta} &= -\frac{\beta}{\sigma^2}\sum_{i=1}^n x_i + \frac{1}{\sigma^2}\sum_{i=1}^n x_i y_i. \\ \implies \hat{\beta}_{\text{MLE}} &= \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}. \end{aligned}$$

To check if the solution is unique (i.e. a global maximum), we can show the second partial derivative of $\ell(\beta|\boldsymbol{y}, \sigma^2)$.

$$\frac{\partial \ell(\beta)}{\partial \beta}\Big|_{\beta=\hat{\beta}_{\text{MLE}}} = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i\Big|_{\beta=\hat{\beta}_{\text{MLE}}} < 0 \ \forall \ \beta \in \mathbb{R}.$$

Therefore,

$$\hat{\beta}_{\text{MLE}} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

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

6.

a.

We have that $S^2|\sigma^2 \sim \chi^2_{n-1}$ with sample size n. This is equivalent to saying that $S^2|\sigma^2 \sim \text{Gamma}\left(\frac{n-1}{2}, \frac{2\sigma^2}{n-1}\right)$. We suppose the prior distribution $\pi(\sigma^2)$ is an inverse gamma with shape parameter α and scale parameter β,

$$\pi(\sigma^2) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\frac{1}{(\sigma^2)^{\alpha+1}}\exp\left\{-\frac{1}{\beta\sigma^2}\right\}, \quad 0 < \sigma^2 < \infty.$$

We know that the posterior distribution for $\sigma^2|S^2$ is

$$\pi(\sigma^2|S^2) = \frac{f(S^2|\sigma^2)\pi(\sigma^2)}{m(S^2)}.$$

Knowing $f(S^2|\sigma^2)$ and $\pi(\sigma^2)$, we can say that

$$\begin{aligned}
\pi(\sigma^2|S^2) &\propto \frac{1}{(\sigma^2)^{(n-1)/2}}\exp\left\{-\frac{(n-1)S^2}{2\sigma^2}\right\}\frac{1}{(\sigma^2)^{\alpha+1}}\exp\left\{-\frac{1}{\beta\sigma^2}\right\} \\
&= \frac{1}{(\sigma^2)^{(n-1)/2+\alpha+1}}\exp\left\{-\left(\frac{(n-1)S^2}{2\sigma^2} + \frac{1}{\beta\sigma^2}\right)\right\} \\
&= \text{IG}\left(\alpha + \frac{(n-1)}{2}, \left\{\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right\}^{-1}\right).
\end{aligned}$$

This implies that the kernel of the posterior is proportional to an $\text{IG}\left(\alpha + \frac{(n-1)}{2}, \left\{\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right\}^{-1}\right)$. So $\sigma^2|S^2 \sim \text{IG}\left(\alpha + \frac{(n-1)}{2}, \left\{\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right\}^{-1}\right)$.

b.

We know the Bayes estimator for σ^2 is

$$\sigma^2_{\text{B}} = E[\sigma^2|S^2].$$

Since $\sigma^2|S^2 \sim \text{IG}\left(\alpha + \frac{(n+1)}{2}, \left\{\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right\}^{-1}\right)$, we need to find the expectation for the inverse gamma distribution.

$$\begin{aligned}
E[X] &= \int_0^\infty x \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{1}{(x)^{\alpha+1}} \exp\left\{-\frac{1}{\beta x}\right\} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \underbrace{\int_0^\infty \frac{1}{(x)^\alpha} \exp\left\{-\frac{1}{\beta x}\right\} dx}_{\text{Kernel of IG}(\alpha-1,\beta)} \\
&= \frac{\Gamma(\alpha-1)\beta^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \underbrace{\int_0^\infty \frac{1}{\Gamma(\alpha-1)\beta^{\alpha-1}} \frac{1}{(x)^\alpha} \exp\left\{-\frac{1}{\beta x}\right\} dx}_{\text{IG}(\alpha-1,\beta)\text{ Density}} \\
&= \frac{\Gamma(\alpha-1)\beta^{\alpha-1}}{(\alpha-1)\Gamma(\alpha-1)\beta^\alpha} \\
&= \frac{1}{\beta(\alpha-1)}.
\end{aligned}$$

Knowing the expectation for $\text{IG}(\alpha, \beta)$, we can find the Bayes estimator for σ^2.

$$\begin{aligned}
E[\sigma^2|S^2] &= \frac{1}{\left(\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right)^{-1}\left(\alpha + \frac{(n+1)}{2} - 1\right)} \\
&= \frac{\left(\frac{(n-1)S^2}{2} + \frac{1}{\beta}\right)}{\alpha + \frac{(n+3)}{2}}.
\end{aligned}$$