- You have 3 hrs to complete this exam
- The exam has two parts. Part I requires to solve all problems. Part II allows you to choose between two problems. Please solve just one problem in Part II. If you answer both, only the lowest grade out of the two will be taken into account.
- The last page of the exam has a list of pmf's and pdf's that you may (or may not) need to use throughout the exam.
- Hint: You do not need to use integrals to solve any of the problems on this exam.

Part I

1. (5) Let $X_1, \ldots, X_n$ be *iid* with pdf

$$f(x|\theta) = \frac{1}{\theta},\ 0 \leq x \leq \theta,\ \theta > 0.$$

Derive the MLE estimate of θ.

$$L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta) = \begin{cases} \left(\frac{1}{\theta}\right)^n & \text{if } x_{(1)} \geq 0 \text{ and } x_{(n)} \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

$\hat{\theta}_{MLE}$ is the value of θ that maximizes the above likelihood function. Taking the derivative of the likelihood function with respect to θ,

$$\frac{\partial L(\theta|\mathbf{x})}{\partial \theta} = -n\left(\frac{1}{\theta}\right)^{n-1} < 0$$

The likelihood function is decreasing in θ. Thus in order to maximize this function, we wish to choose the smallest θ possible subject to the constraint $\theta \geq X_{(n)}$. If $\theta > X_{(n)}$, then the $L(\theta|\mathbf{x}) = 0$. Therefore, $\hat{\theta}_{MLE} = X_{(n)}$.

2. (5) Let $X_1, X_2, \ldots$ be a sequence of random variables that converges in probability to a constant a. Assume that $P(X_i > 0) = 1$ for all i. Verify that the sequences defined by $Y_i = \sqrt{X_i}$ and $Y_i' = a/X_i$ converge in probability, and find the limits (what do they converge to?).

$$
\begin{aligned}
&X_n \to_p a \\
&\text{By Thm 5.5.4, } Y_n = \sqrt{X_n} \to_p \sqrt{a} \\
&\text{Also by Thm 5.5.4, } \frac{1}{X_n} \to_p \frac{1}{a} \\
&\text{By Slutsky Thm, } \frac{a}{X_n} \to_p \frac{a}{a} = 1
\end{aligned}
$$

3. (5) For any two random variables X and Y with finite variances, prove that X and $Y - \mathbb{E}(Y|X)$ are uncorrelated.

$$
\begin{aligned}
COV(X, Y - \mathbb{E}(Y|X)) &= \mathbb{E}[X(Y - \mathbb{E}(Y|X))] - \mathbb{E}[X]\mathbb{E}[Y - \mathbb{E}(Y|X)] \\
&= \mathbb{E}[XY - X\mathbb{E}(Y|X)] - \mathbb{E}[X](\mathbb{E}[Y] - \mathbb{E}[\mathbb{E}(Y|X)]) \\
&= \mathbb{E}[XY] - \mathbb{E}[X\mathbb{E}(Y|X)] - \mathbb{E}[X](\mathbb{E}[Y] - \mathbb{E}[Y]) \\
&= \mathbb{E}[XY] - \mathbb{E}[\mathbb{E}(XY|X)] \\
&= \mathbb{E}[XY] - \mathbb{E}[XY] \\
&= 0
\end{aligned}
$$

4. (5) Let $X_1, \ldots, X_n$ be iid Bernoulli(p). Show that the variance of $\bar{X}$ attains the Cramr-Rao Lower Bound, and hence $\bar{X}$ is the best unbiased estimator of p.

First, lets find the Cramer-Rao Lower Bound:

$$
\begin{aligned}
Var(\hat{p}) &\geq \frac{1}{-n\mathbb{E}\left[\frac{\partial^2}{\partial p^2}\log P(X=x|p)\right]} \\
&\geq \frac{1}{-n\mathbb{E}\left[\frac{\partial^2}{\partial p^2}\left(x\log(p)+(1-x)\log(1-p)\right)\right]} \\
&\geq \frac{1}{-n\mathbb{E}\left[\frac{\partial}{\partial p}\left(\frac{x}{p}-\frac{(1-x)}{(1-p)}\right)\right]} \\
&\geq \frac{1}{-n\mathbb{E}\left(\frac{-x}{p^2}-\frac{(1-x)}{(1-p)^2}\right)} \\
&\geq \frac{1}{-n\left(\frac{-\mathbb{E}(x)}{p^2}-\frac{(1-\mathbb{E}(x))}{(1-p)^2}\right)} \\
&\geq \frac{1}{-n\left(\frac{-p}{p^2}-\frac{(1-p)}{(1-p)^2}\right)} \\
&\geq \frac{1}{-n\left(\frac{-1}{p}-\frac{1}{(1-p)}\right)} \\
&\geq \frac{1}{n\left(\frac{1}{p(1-p)}\right)} \\
&\geq \frac{p(1-p)}{n}
\end{aligned}
$$

Next, lets find the variance of $\hat{p} = \bar{X}$

$$
\begin{aligned}
Var(\bar{X}) = Var(\frac{1}{n}\sum_i X_i) &= \frac{1}{n^2}\sum_i Var(X_i) \\
&= \frac{1}{n^2} n Var(X_i) \\
&= \frac{p(1-p)}{n}
\end{aligned}
$$

Thus, the estimator $\hat{p} = \bar{X}$ achieves the Cramer-Rao Lower Bound.

5. (5) Define $X_1, X_2, \ldots, X_n$ as a random sample of exponentially distributed variables with parameter λ, $f_{X_i}(x) = \lambda\exp(-\lambda x)$, $F_{X_i}(x) = 1-\exp(-\lambda x)$. Define the statistic $X_{\{1\}}$ as $\min\{X_1, \ldots, X_n\}$. Derive the cdf of $X_{\{1\}}$.

$$
\begin{aligned}
F_{X_{(1)}}(x) &= Pr(X_{(1)} < x) \\
&= 1 - Pr(X_1 > x) * Pr(X_2 > x) * \ldots * Pr(X_n > x) \\
&= 1 - (1 - Pr(X_1 < x)) * (1 - Pr(X_2 < x)) * \ldots * (1 - Pr(X_n < x)) \\
&= 1 - (1 - (1 - \exp(-\lambda x)))^n \\
&= 1 - (\exp(-\lambda x))^n \\
&= 1 - (\exp(-n\lambda x))
\end{aligned}
$$

Note, the cdf of $X_{\{1\}}$, is the cdf of an exponential distribution with parameter $n\lambda$.

Part II

6. (25) Consider a random sample, $X_1, X_2, \ldots, X_n$, where X_i is distributed exponential with parameter λ ($f_{X_i}(x) = \lambda \exp(-\lambda x)$).

   (a) (3) What is the method of moments estimator of μ, where $\mu = \mathbb{E}(X_i)$? Denote the method of moments estimator you obtained μ_{MM}. Is μ_{MM} consistent? Explain why.

$$\mu_{MM} = \frac{1}{n}\sum_{i}^{n} X_i$$

Our method of moments is consistent, since by the Law of Large Numbers,

$$\mu_{MM} = \frac{1}{n}\sum_{i}^{n} X_i \rightarrow_p E(X_i) = \mu$$

   (b) (3) Show that μ_{MM} is unbiased.

$$
\begin{aligned}
E(\mu_{MM}) &= E(\frac{1}{n}\sum_i^n X_i) \\
&= \frac{1}{n}\sum_i^n E(X_i) \\
&= \frac{1}{n} n\mu \\
&= \mu
\end{aligned}
$$

(c) (3) What is the method of moments estimator of λ?

$$
\begin{aligned}
E(X_i) &= \frac{1}{\lambda} \\
\Rightarrow \lambda_{MM} &= \frac{1}{\frac{1}{n}\sum_i^n X_i}
\end{aligned}
$$

(d) (3) Obtain the Maximum Likelihood Estimator (MLE) of parameter λ, λ_{MLE}.

$$
\begin{aligned}
&L(\lambda|\mathbf{x}) = f(\mathbf{x}|\lambda) = \lambda^n \exp(-\lambda \sum_i^n X_i) \\
&\log L(\lambda|\mathbf{x}) = n\log(\lambda) - \lambda \sum_i^n X_i \\
&\frac{\partial}{\partial\lambda}\log L(\lambda|\mathbf{x}) = \frac{n}{\lambda_{MLE}} - \sum_i^n X_i = 0 \\
&\Rightarrow \lambda_{MLE} = \frac{n}{\sum_i^n X_i} = \frac{1}{\bar{X}}
\end{aligned}
$$

(e) (3) Is the MLE of λ unbiased? Explain. (You do not need to integrate to answer this question)

λ_{MLE} is not unbiased. To prove this, we can use Jensen's Inequality. Since the function $\frac{1}{x}$ is strictly convex, we have

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\bar{x}}\right) > \frac{1}{E(\bar{x})} \\
&\Rightarrow \mathbb{E}\left(\frac{1}{\bar{x}}\right) = \mathbb{E}(\lambda_{MLE}) > \frac{1}{E(\bar{x})} = \lambda \\
&\Rightarrow \mathbb{E}(\lambda_{MLE}) > \lambda
\end{aligned}$$

Thus, the MLE estimate is biased.

(f) (4) Show that the MLE for λ is consistent.

From the Law of Large Numbers,

$$\bar{X} = \frac{1}{n}\sum_{i}^{n} X_i \to_p E(X_i) = \frac{1}{\lambda}$$

Then, from Thm. 5.5.4,

$$\lambda_{MLE} = \frac{1}{\bar{X}} \to_p \frac{1}{\frac{1}{\lambda}} = \lambda$$

(g) (3) Derive the asymptotic distribution for $\sqrt{n}(\lambda_{MLE} - \lambda)$?

From the Central Limit Theorem,

$$\sqrt{n}\left(\bar{X} - \frac{1}{\lambda}\right) \to_d \text{n}\left(0, \frac{1}{\lambda^2}\right)$$

Then, using the Delta Method,

$$\sqrt{n}(\lambda_{MLE} - \lambda) \to_d \text{n}\left(0, \frac{1}{\lambda^2}\lambda^4\right)$$

(h) (3) Construct a Likelihood Ratio Test for $H_0 : \lambda = 1$, $H_1 : \lambda \neq 1$.

A likelihood ratio test statistic, denoted $\gamma(\mathbf{x})$ is constructed as follows (Note: usually we use $\lambda(\mathbf{x})$ to denote the likelihood ratio test statistic, but will not here since λ is already defined as a parameter of the distribution of X_i).

$$\gamma(\mathbf{x}) = \frac{L(\lambda = 1|\mathbf{x})}{L(\lambda_{MLE}|\mathbf{x})}$$

The distribution of $\gamma(\mathbf{x})$ is unknown, but we can transform this statistic to one which we know the asymptotic distribution,

$$-2\log\gamma(\mathbf{x}) \to_d \chi^2_1$$

7. (25) A researcher is interested in estimating the proportion of the population whose incomes are below the poverty line, a prespecified level of income. Denote income Y and c the poverty line. The parameter of interest is $\theta = \Pr(Y \le c)$. The researcher considers estimating θ using a random sample of size n, $Y_1, Y_2, ..., Y_n$. Define a random sample, $C_1, C_2, ..., C_n$, where C_i is defined as

$$C_i = \begin{cases} 1 \text{ if } Y_i \le c \\ 0 \text{ otherwise} \end{cases}$$

   (a) (3) The method of moments estimator for $\theta = \Pr(Y \le c)$ can be defined in terms of this newly defined random sample as

$$\theta_{MM} = \frac{1}{n}\sum_{i=1}^{n} C_i$$

   Show that θ_{MM} is an unbiased estimator of θ.

$$\begin{aligned} E(\theta_{MM}) &= E(\frac{1}{n}\sum_{i}^{n} C_i) \\ &= \frac{1}{n}\sum_{i}^{n} E(C_i) \\ &= \frac{1}{n} nPr(Y < c) \\ &= \theta \end{aligned}$$

   (b) (4) Derive the variance of θ_{MM}.

$$
\begin{aligned}
Var(\theta_{MM}) &= Var(\frac{1}{n}\sum_{i}^{n} C_i) \\
&= \frac{1}{n^2}\sum_{i}^{n} Var(C_i) \\
&= \frac{1}{n^2} n\theta(1-\theta) \\
&= \frac{\theta(1-\theta)}{n}
\end{aligned}
$$

(c) (4) For the remaining parts of this question, assume that you now know that income is normally distributed with known variance but unknown mean, *e.g.* $Y_i \sim \mathrm{n}\left(\mu, \sigma^2\right)$. What is the MLE of μ? Denote the obtained estimate as μ_{MLE}.

$$
\begin{aligned}
&f(Y_i|\mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(Y_i-\mu)^2/(2\sigma^2)} \\
&L(\mu|\mathbf{Y}) = f(\mathbf{Y}|\mu) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\sum_i^n (Y_i-\mu)^2/(2\sigma^2)} \\
&\log L(\mu|\mathbf{Y}) = n\log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{(2\sigma^2)}\sum_{i}^{n}(Y_i-\mu)^2 \\
&\frac{\partial L(\mu|\mathbf{Y})}{\partial\mu} = \frac{1}{(2\sigma^2)} 2\sum_{i}^{n}(Y_i-\mu_{MLE}) = 0 \\
&\Rightarrow \mu_{MLE} = \frac{1}{n}\sum_{i}^{n} Y_i
\end{aligned}
$$

(d) (3) Show whether μ_{MLE} is unbiased or not.

$$
E(\mu_{MLE}) = \frac{1}{n}\sum_{i}^{n} E(Y_i) = \mu
$$

Thus, μ_{MLE} is an umbiased estimator of μ.

(e) (4) Denote $F(z)$ the cdf of the a standard normal evaluated at z and $f(z)$ the pdf of the standard normal evaluated at z. Notice that with these asumptions $\theta = \Pr(Y < c) = F\left(\frac{c-\mu}{\sigma}\right)$, or the cdf of a standard normally distributed variable evaluated at $\frac{c-\mu}{\sigma}$. Describe how would you use μ_{MLE} to obtain the MLE of θ, θ_{MLE}.

From the invariance property of MLEs,

$$\theta_{MLE} = F\left(\frac{c-\mu_{MLE}}{\sigma}\right)$$

We could evaluate the standard normal cdf at μ_{MLE} to obtain θ_{MLE}.

(f) (3) Is θ_{MLE} consistent? Explain (no need to show).

Yes. All MLEs are consistent. This is always a property of a MLE.

(g) (4) Find the asymptotic distribution of θ_{MLE}.

First, not that the asymptotic distribution of μ_{MLE} (which is just the sample average of Y_i) is

$$\sqrt{n}(\mu_{MLE}-\mu) \to_d \mathrm{n}(0,\sigma^2)$$

Since we can write θ as a function of μ, $\theta = \Pr(Y < c) = F\left(\frac{c-\mu}{\sigma}\right)$, we can use the Delta Method to derive the asymptotic distribution of θ_{MLE}:

$$\begin{aligned}
\sqrt{n}(\theta_{MLE}-\theta) &\to_d \mathrm{n}(0,\sigma^2 F'\left(\frac{c-\mu}{\sigma}\right)^2) \\
&\to_d \mathrm{n}(0,\sigma^2\left(f\left(\frac{c-\mu}{\sigma}\right)\frac{-1}{\sigma}\right)^2) \\
&\to_d \mathrm{n}(0,f\left(\frac{c-\mu}{\sigma}\right)^2)
\end{aligned}$$