Problem Set 6 **Statistical Methods In Engineering And Science**

Due Date: 10:00 PM, May 12, 2023 Prof. Alexander Giessing

Last Update: May 5, 2023 Spring Quarter, 2023

Study Group: ______________________ ______________________ ______________________

Please upload your solution in a single pdf file on Canvas. Include all calculations, R-code, and figures (if applicable). All data sets are available on Canvas https://canvas.uw.edu/courses/1584511.

Question 1. The "Roman Casino" in Seattle is for sale. You consider buying it, but you want to know how much money you are going to make. The current owner, the Maverick Gaming LLC, is only willing to share the following information with you: The roulette game "Red or Black" is played about 1,000 times a night, 365 days a year. Each time the game is played the owner of the casino wins \$1 with probability 20/38 and loses \$1 with probability 18/38.
Explain in detail why the weak law of large numbers can be used to determine the revenue (per year) of the casino, and determine how much it is. (For simplicity, assume that Seattleites only know to play "Red or Black").

Question 2. Let $X_1, \ldots, X_{730}$ be i.i.d. random variables, with pdf f given by

$$f(x) = \begin{cases} 4(1-x)^3 & \text{if } 0 \leq x \leq 1, \\ 0 & o/w. \end{cases}$$

Use the central limit theorem to approximate the probability $P(X_1 + \ldots + X_{730} < 153.3)$.

Question 3. Consider the following data set consisting of 12 observations

.92 .79 .90 .65 .86 .47 .73 .97 .94 .77 .79 .91

Suppose that the data are the realization of a random sample $X_1, \ldots, X_{12}$ from a distribution with pdf

$$f(x) = \begin{cases} (\theta+1)x^\theta & \text{if } 0 \leq x \leq 1, \\ 0 & o/w. \end{cases}$$

(a) What is the method of moment estimator for θ? Compute an estimate based on the data set.

(b) What is the maximum likelihood estimator for θ? Compute an estimate based on the data set.

Question 4. Let $X_1, \ldots, X_n$ be a random sample from $Ber(p)$. Show that the method of moment estimator and the maximum likelihood estimator, both, are equal to the sample proportion $\hat{p} =$

$n^{-1}\sum_{i=1}^{n} X_i$.

Question 5. Let $X_1, \ldots, X_n$ be a random sample from a $Bin(N, p)$.

(a) Suppose that N is known and only success probability p is unknown. Compute the method of moment estimator and the maximum likelihood estimator for p.

(b) Suppose that N and p are unknown. Compute the method of moment estimators for N and p.

Question 6. Let $X_1, \ldots, X_n$ be a random sample from $Gamma(\alpha, \beta)$ with pdf

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{if } x \in (0, \infty); \\ 0 & o/w, \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. Assume that the shape parameter $\alpha > 0$ is known and compute the maximum likelihood estimator of the rate parameter $\beta > 0$.

Question 7. Given a random sample $X_1, \ldots X_n$ from $Exp(\lambda)$ with unknown $\lambda > 0$ you want to estimate the expectation $1/\lambda$. From the lecture you know that the sample average $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ is an unbiased estimator for $1/\lambda$. Now, consider a more generally estimator of the form $T_c = c \times (\sum_{i=1}^{n} X_i)$ for $c \in \mathbb{R}$. You are interested in the MSE of this estimator and you would like to know whether there are choices for $c \in \mathbb{R}$ such that T_c has smaller MSE than $\bar{X}_n$.

(a) Compute the $\text{MSE}(T_c)$ for arbitrary $c \in \mathbb{R}$.

(b) For which $c \in \mathbb{R}$ does the estimator T_c perform best in the MSE sense? Compare the MSE of this "best" estimator with the MSE of the unbiased estimator $\bar{X}_n$ (which is just the estimator T_c with $c = 1/n$).

Question 8. Suppose that you have two unbiased estimators U and V with the same variance $\text{Var}(U) = \text{Var}(V)$. Based on the MSE criterion you cannot choose one estimator over the other. Consider now $W = (U + V)/2$.

(a) Show that W is unbiased.

(b) Show that

$$\frac{\text{Var}\big((U+V)/2\big)}{\text{Var}(U)} = \frac{1}{2} + \frac{1}{2}\rho(U, V),$$

where $\rho(U, V)$ is the correlation coefficient between U and V.

(c) Why do the results for part (a) and (b) imply that we should use W instead of U or V?