PH.D. QUALIFYING EXAMINATION – THEORETICAL STATISTICS
Time: 1:00 PM – 5:00 PM, August 28, 2020

General Instructions

- This exam consists of two parts. Part A (STA584) has 5 problems and a possible total of 50 points. Part B (STA684) consists of 8 problems and has a possible total of 100 points. Your raw score in Part A will be converted to a percentage.
- Write your name on the first page only. Begin each question on a new sheet with its number noted. Write on one side only. Once finished, please arrange all pages in the order of question numbers and then number the pages accordingly.
- When presenting the probability distribution of a random variable, the range of the random variable must be specified.
- You must show all your work and justifications correctly and clearly to receive full credits, and partial credits may be given for partially correct solutions.

Part A

Question #1 [**12 points**]

(a) In an election, there are five candidates for mayor and four candidates for city treasurer.
  (i) In how many ways can a voter mark her ballot for both of these offices?
  (ii) In how many ways can a person vote if she exercises her option of not voting for a candidate for any or both of these offices?

(b) Five identical cars pull into a parking lot that has eight empty spaces.
  (i) In how many ways can the cars occupy five of those spots?
  (ii) How many arrangements are such that the leftmost and the rightmost spaces remain unoccupied?

(c) (i) How many different five-digit numbers can be formed from the eight digits 2 through 9 with no digit being used more than once?
  (ii) Find the probability that the number in (i) will be more than 72000.

(d) As of July 1, suppose a major league baseball team had won 53% of their games in 2019. Eighteen percent of their games had been played against left-handed starting pitchers. The team won 36% of the games played against left-handed starting pitchers.
  (i) Set up a two-way table showing the results expected for every 100 games. Use the table columns for the pitchers and the table rows for the games.
  (ii) What percentage of their games did they win against right-handed starting pitchers?
  (iii) What percentage of the lost games were played against left-handed starting pitchers?

Question #2 [**10 points**]

(a) Chebyshev's Theorem: If μ and σ are the mean and the standard deviation of a random variable X, then for any positive constant k, $P\left(|X-\mu|<k\sigma\right)\geq 1-1/k^2$. Prove the theorem for a continuous random variable X.

(b) Use the result in (a) to find a lower bound on P(– 4 < X < 26) where the random variable X has a mean $\mu = 11$ and variance $\sigma^2 = 9$.

(c) Let X be a discrete random variable with the negative binomial distribution

$$P(X=x)=\binom{n+x-1}{x}\theta^x(1-\theta)^n,\quad x=0,1,2,\ldots,$$

(i) What does the random variable X represent? Explain the meanings of n and θ.

(ii) Take the first derivative with respect to θ the expressions on both sides of the equation $\sum_{x=0}^{\infty}\binom{n+x-1}{x}\theta^x(1-\theta)^n = 1,$ find the mean of the negative binomial distribution.

Question #3 [10 points]

The lifetime (in days), X, of an electronic equipment is a random variable with probability density function

$$f(x) = \begin{cases} \frac{1}{4}e^{-x/4}, & x > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the probability that the equipment will operate for at least 6.4 days without failure. Round your answer to two places of decimal.

(b) Suppose five such equipment are considered. Find the probability that none of the equipment will operate for at least 6.4 days without failure

(c) Find P$[-4\ln(1-p) < X < -4\ln(1-kp)]$ where $0 < kp < 1$. Can the values of k lie between 0 and 1? Explain.

(d) Find the moment generating function $M_X(t)$, of the random variable X. For what values of t is the moment generating function $M_X(t)$ defined?

Question #4 [13 points]

Suppose X and Y are continuous random variables with joint probability density function

$$f(x, y) = \begin{cases} kxy, & 0 < x < 1,\ 0 < y < 1,\ x + y > 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the value of k that makes the function a probability density function.

(b) Find P$(X + Y < \frac{3}{2})$

(c) Find P$(X \geq \frac{3}{4}, Y \geq \frac{1}{2})$

(d) Find the marginal probability density function (PDF) of X. Include the support of the PDF.

(e) Find the conditional PDF of Y given that $X = x$. Include the support of the PDF.

Question #5 [5 points]

The joint probability density function for the length of life of two different types of components operating in a system is given by

$$f(x, y) = \begin{cases} \frac{x}{8}e^{-(x+y)/2}, & x > 0,\ y > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

The relative efficiency of the two types of components is measured by $W = Y/X$. Find the probability density function for W. Include the support of the probability density function.

Theoretical Statistics Qualifying Exam: Part B STA 684

General Instruction: Answer all questions. Print your answer, your name, number the pages and number the problems on provided exam paper. Write on one side only.

1. Let $X_1, \ldots, X_n$ be iid random variables each having a normal distribution with mean μ and variance σ^2.

   Prove that $t = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ has a Student t-distribution with $n - 1$ degrees of freedom, where $\bar{X}$ and S are the sample mean and the sample standard deviation. (12)

2. Let the random variable X_n have a distribution that is $b(n, p)$. Prove that $(X_n/n)(1 - X_n/n)$ converges in probability to $p(1 - p)$. (12)

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\theta, \sigma^2)$ distribution, where σ^2 is fixed but $-\infty < \theta < \infty$. Show that the mle of θ is $\bar{X}$. (12)

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu_0, \sigma^2 = \theta)$ distribution, where $0 < \theta < \infty$ and μ_0 is known. Show that the likelihood ratio test of $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$ can be based upon the statistic

$$W = \frac{\sum_{i=1}^{n} (X_i - \mu_0)^2}{\theta_0}$$

   Determine the null distribution of W and give, explicitly, the rejection rule for a level α test. (12)

5. Let $X_1, \ldots, X_{n1}$ be a random sample from the distribution of $X \sim N(\mu_1, \sigma_1^2)$ and let $Y_1, \ldots, Y_{n2}$ be a random sample from the distribution of $Y \sim N(\mu_2, \sigma_2^2)$. Prove that (12)

$$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \Big/ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \xrightarrow{P} 1$$

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution having parameter θ, $0 < \theta < \infty$. Prove that the sum of the observations of the random sample of size n is a sufficient statistic for θ. (10)

7. Suppose $X_1, \ldots, X_n$ are iid with pdf (15)

$$f(x, \theta) = \begin{cases} \theta e^{-\theta x} & 0 < x < \infty \\ 0 & otherwise \end{cases}$$

   a. Find an MVUE of θ.
   b. Prove your result in part a. is an MVUE of θ.

8. (15)

   a. State Neyman–Pearson Theorem.
   b. Prove Neyman–Pearson Theorem.