Department of Applied Mathematics
PROBABILITY AND STATISTICS PRELIMINARY EXAMINATION
August 2020

Instructions:

Do two of three problems in each section (Prob and Stat).	Prob
Place an **X** on the lines next to the problem numbers	1. ____
that you are **NOT** submitting for grading.	2. ____
	3. ____
Do not write your name anywhere on this exam.	Stat
You will be identified only by your student number.	4. ____
Write this number **on each page** submitted for grading.	5. ____
Show all relevant work!	6. ____
	Total ____

Student Number ____________________________

Probability Section

Problem 1.

(a) Let X be a non-negative continuous random variable with cdf F. Show that

$$\mathsf{E}\left[\frac{1}{1+X}\right] = \int_0^\infty \frac{F(x)}{(1+x)^2}\,dx$$

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from the exponential distribution with rate 1. Define

$$Y_n = X_1 + \frac{1}{2}X_2 + \frac{1}{3}X_3 + \cdots + \frac{1}{n}X_n.$$

Find expressions for the moment generating functions for $X_{(n)}$ and Y_n. (Your expressions may contain sums or products but may not contain integrals.)

(c) Compare the two moment generating functions for $n = 1, 2$, and 3. Assuming that the pattern you see continues for all $n \geq 1$, what can you say about the distributions of $X_{(n)}$ and Y_n as they relate to each other?

Problem 2.

Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of iid random variables from the Poisson distribution with parameter λ. Let $Y_n = X_n X_{2n}$ for $n \geq 1$ and consider the nth partial sum

$$S_n = Y_1 + Y_2 + \cdots + Y_n.$$

(a) Find $\mathsf{E}[S_n]$.

(b) Find a constant C, which may depend on λ but which may not depend on n, such that

$$Var[S_n] \leq Cn \qquad \forall n \geq 1.$$

(c) Find a sequence of real numbers $\{a_n\}$ such that

$$\frac{S_n}{a_n} \xrightarrow{P} 1.$$

(Your a_n may depend on λ.)

Problem 3.

In a disease outbreak, there are three different states of an individual: the first state is "susceptible" (denote by s), the second is "infected" (denoted by i), and the third is "recovered" (denoted by r). The state of an individual at time $t \geq 0$, $X(t)$, is modeled as a continuous-time Markov chain with the infinitesimal generator (or rate matrix)

$$Q = \begin{bmatrix} -\eta & \eta & 0 \\ 0 & -\lambda & \lambda \\ 0 & \gamma & -\gamma \end{bmatrix}, \tag{1}$$

for some $\eta, \lambda, \gamma > 0$ with $\gamma < \eta$. We assume that $X(0) = s$.

(a) Consider the first infection time $\tau_i := \inf\{t > 0 : X(t) = i\}$. Given $t > 0$, find the probability $P(\tau_i > t)$.

(b) Given $t > 0$, find the probability that $X(t) = i$ and the state r has not yet been visited, i.e.

$$P\left(X(t) = i \text{ and } X(u) \in \{s, i\} \ \forall u \in [0, t)\right).$$

(c) Given $t > 0$, what is the probability that the individual get infected three times during the period $[0, t]$?

(d) Suppose that there are $N > 0$ individuals in a population. Each individual is susceptible at time 0, and subject to the spread of the disease as in (1) *independently of other individuals.* As $t \to \infty$, what are the limiting fractions of population that are susceptible, infected, and recovered?

Statistics Section

Problem 4.

Consider $X_1, X_2, \ldots, X_n$ where X_i is exponentially distributed with mean λ/θ_i. Let $Y_1, Y_2, \ldots, Y_n$ be exponential random variables with $\mathsf{E}[Y_i] = \lambda\theta_i$. Assume that the X's and Y's are all mutually independent.

In this problem, the parameters $\lambda, \theta_1, \theta_2, \ldots, \theta_n$ are all positive and unknown.

(a) Find the maximum likelihood estimator (MLE) of λ.

For parts (b) and (c), assume that $\theta_1, \theta_2, \ldots, \theta_n$ are known.

(b) Find the MLE for λ and the UMVUE (uniformly minimum variance unbiased estimator) for λ.

(c) Compute the relative efficiency of your estimators from part (b). What can you say as $n \to \infty$?

Problem 5.

Suppose that X and Y are iid $N(0,1)$ random variables. It is well known that X^2 and Y^2 each have a $\chi^2(1)$ distribution.

(a) Let $W = \min(X, Y)$. Show that $W^2 \sim \chi^2(1)$.

(b) Now suppose that X and Y are iid $N(\mu, \sigma^2)$ random variables with μ known and σ^2 unknown. Use part (a) to derive a $100(1-\alpha)\%$ confidence interval for σ^2 based on the statistic $W = \min(X, Y)$.

Problem 6.

Suppose that we have a random sample, $X_1, X_2, \ldots, X_n$ from the distribution with pdf

$$f(x;\theta) = \frac{1}{6\theta^3} x^2 e^{-x/\theta} I_{(0,\infty)}(x)$$

(a) Find the best (most powerful) test of size α of $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, assuming that $\theta_1 > \theta_0$. Give your answer in terms of a chi-squared critical value.

(b) Is your test uniformly most powerful (UMP) for the alternative hypothesis $H_1 : \theta > \theta_0$? Explain.

(c) Is your test uniformly most powerful (UMP) for the alternative hypothesis $H_1 : \theta \neq \theta_0$? Explain.

(d) Derive an approximate large-sample generalized likelihood ratio test (GLRT) of size α for the hypotheses in parts (b) and (c) <u>if</u> your test was not a UMP test. (Note: Depending on how you answered (b) and (c), you may have nothing to do here, you may have one test to do, or you may have 2 tests to do.)
