PH.D. QUALIFYING EXAMINATION – THEORETICAL STATISTICS
Time: 9:00 AM – 1:00 PM, Friday, August 22, 2025
Instructions

- The exam consists of two parts: STA 584 and STA 684. You are required to answer all questions. The raw score for each part will be converted to its percentage.
- Write on one side of the paper only. Begin each subpart on a new sheet with the problem number clearly labeled. You must show all your work and justifications completely and correctly to receive full credit. Partial credit may be given for partially correct solutions.
- For each problem or subproblem, submit only the answer you want to be graded. Any crossed-out work will be ignored. Failure to follow this instruction for a problem will result in a zero score for the problem.
- If you apply a theorem, you must state the theorem, identify its assumptions and conclusions, and justify why it is applicable. New notations must be defined before use. You need to specify the support of the distribution in your answer.
- When finished, please collate all pages based on problem labels and then number the pages accordingly. Hand in the exam paper.

By signing below, I hereby acknowledge that I have read and fully understand the instructions.

____

Signature

____

Printed Name

Part A - STA 584

---

1. (8 pts, 2 pts/each) Find the probability in the questions below.

1.a) An automobile safety engineer claims that 1 in 10 automobile accidents is due to driver fatigue. Find the probability that at least 2 of 5 randomly selected automobile accidents are due to driver fatigue.

1.b) An expert sharpshooter misses a target 5 percent of the time. Find the probability that she will miss the target for the second time on the fifth shot.

1.c) Among the 8 applicants for a job, 5 have college degrees. If three of the applicants are randomly chosen for an interview, find the probability that all three have college degrees.

1.d) Binary digits are transmitted over a communication system. If a **1** is sent, it will be received as a **1** with probability 0.95 and as a **0** with probability 0.05. If a **0** is sent, it will be received as a **0** with probability 0.99 and as a 1 with probability 0.01. A series of 0's and 1's is sent in random order, with 0's and 1's each being equally likely. If a digit is received as a **1,** find the probability it was sent as a **1**.

2. (9 pts, 3 pts/each) Let X and Y be independent random variables with a continuous uniform distribution on the interval (0, 1).

2.a) Find the cumulative distribution function (CDF) and probability density of the random variable $W = -2\,ln(X)$. What are the name and the mean value of the probability distribution of W?

2.b) Consider the two random variables $U = X + Y$ and $V = Y$. Find the joint probability density of U and V. Specify the ranges of the random variables U and V clearly.

2.c) Find the marginal density of U.

3. (15 pts, 2 pts/each) The joint probability density of X and Y is given by

$$f(x,y) = \begin{cases} 2 & \text{for } x>0, y>0, x+y<1 \\ 0 & \text{otherwise.} \end{cases}$$

3.a) Find $P(X \le 0.5)$.

3.b) Find $P(X > 2Y)$.

3.c) Find the variance of X.

3.d) Find the covariance of X and Y.

3.e) Find the variance of $W = X - Y - 0.5$.

3.f) Find the conditional probability function of Y given $X = x$, where $0 < x < 1$.

3.g) Are X and Y independent? Explain. (1 point)

3.h) Find the mean of Y given $X = x$.

4. (7 pts, STA 584 & STA 684) Let X_1, X_2, and X_3 be a random sample of size *n* = 3 from the uniform distribution defined
over the unit interval (0,1). By definition, the range, *R*, of a sample is the difference between the largest $X_{(3)}$ and smallest $X_{(1)}$ order statistics. That is, $R = X_{(3)} - X_{(1)}$.

4.a) Find $P(X_{(1)} \le x)$, where $0 < x < 1$. (2 pts)

4.b) Find the joint PDF of $X_{(1)}$ and $X_{(3)}$. (1 pt)

4.c) Find $P(R \le r)$, where $-\infty < r < \infty$. (3 pts)

4.d) Find the PDF for the range. (1 pt)

Part B - STA 6 84

---

1. (*20 points*) Let Y_n denotes a random variable that is $b(n,p),\ 0 < p < 1$. Prove that
$$\frac{Y_n - np}{\sqrt{n\left(\frac{Y_n}{n}\right)\left(1-\frac{Y_n}{n}\right)}} \xrightarrow{D} N(0,1).$$

2. (*20 points*) The inverse Gaussian distribution, $IG(\theta,\mu)$, is given by
$$f(x;\theta,\mu) = \sqrt{\frac{\theta}{2\pi x^3}}\, exp\left\{-\frac{\theta}{2x}\left(\frac{x-\mu}{\mu}\right)^2\right\}; 0 < x < \infty, 0 < \theta, 0 < \mu.$$
(2.a) Show that the statistics $\bar{X} = (1/n)\sum_{i=1}^{n} x_i$ and $S^* = \sum_{i=1}^{n}(1/x_i - 1/\bar{X})$ are complete sufficient statistics.
(2.b) Show that the distribution of $\bar{X}$ and S^* are, respectively, inverse Gaussian and Chi-square distributions.

3. (*15 points*) Consider the following trivariate pdf of $X_1, X_2,$ and X_3
$$f(x_1,x_2,x_3) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_4)}\left(\prod_{i=1}^{3}\frac{x_i^{\alpha_i-1}}{\Gamma(\alpha_i)}\right)\left(1-\sum_{i=1}^{3}x_i\right)^{\alpha_4-1},$$
where $\alpha_0 = \sum_{j=1}^{4}\alpha_j$ such that $\alpha_j > 0$, for $j = 1,2,3,4$ and $x_i > 0$ for $i = 1,2,3$, such that $\sum_{i=1}^{3} x_i < 1$, while the function f is equal to zero elsewhere.

Find the conditional pdf $f(x_1|x_2,x_3)$ and state its name.

4. (*15 points*)
(4.a) State the Neyman–Pearson Theorem.
(4.b) Consider a random sample $X_1, X_2, \dots, X_n$ from the pdf
$$f(x;\mu) = \sqrt{2/\pi}\,\mu x^{-2} e^{-\frac{1}{2}(x-\mu/x)^2}, 0 < x < \infty, \mu > 0.$$
Let $\overline{x_r^2} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i^2}$. Prove that the most powerful critical region of size α to test $H_0{:}\,\mu = 1$ vs $H_a{:}\,\mu = 2$ for some constant k is $\overline{x_r^2} \le k$. Explain how to find the unknown constant k.

5. (30 *points*) Suppose a random sample of size n is taken from a distribution having the pdf
$$f(x) = \frac{\alpha}{\sqrt{2\pi x}} e^{-\frac{1}{2}\alpha^2 \ln^2(x/\theta)}; x > 0, \alpha > 0, \theta > 0.$$
(5.a) Find the maximum likelihood estimators, $\hat{\alpha}$ and $\hat{\theta}$ for α and θ respectively.
(5.b) Derive the formula for the expected (Fisher) information matrix, $I(\alpha,\theta)$.
(5.c) Find the maximum likelihood estimator, $\hat{E}[X]$, for the mean of the distribution.
(5.d) Using the Delta method, show that the variance of the estimator obtained in Part (5.c) is given by
$$Var\big(\hat{E}[X]\big) = \frac{\theta^2}{n\alpha^2}\left(1+\frac{1}{2\alpha^2}\right) exp\left(\frac{1}{\alpha^2}\right).$$