National Sun Yat-sen University
Department of Applied Mathematics

Ph.D. Qualifying Exam

Mathematical Statistics

INSTRUCTIONS:

- You have **240 minutes** to complete this exam which consists of **10 questions**.
- Please show all your steps to receive partial credit.
- Please organize your work in a reasonably neat and coherent way.
- Please label the answer clearly.
- Notations: i.i.d., independent and identically distributed; cdf, cumulative distribution function; pdf, probability density function; mle, maximum likelihood estimator.

Question	1	2	3	4	5	6	7	8	9	10	Total
Points	10	10	10	10	10	10	10	10	10	10	100
Score											

Question 1. Let X and Y be i.i.d. $N(0,1)$ random variables, and define $Z = \min(X, Y)$. Find the pdf of Z^2.

Question 2. Let $X_1, \dots, X_n$ be i.i.d. uniform random variables on the interval $[0, \theta]$ with $\theta > 0$ and let $X_{(1)} < \cdots < X_{(n)}$ denote the order statistics. Define the range and midrange as $R = X_{(n)} - X_{(1)}$ and $V = (X_{(1)} + X_{(n)})/2$, respectively. Find the joint pdf of (R, V).

Question 3. Let X and Y be random variables with finite means and variances. Find a function g^* such that

$$g^* = \arg\min_{g \in \mathcal{G}} E[\{Y - g(X)\}^2],$$

where $\mathcal{G}$ is a class of all square-integrable functions of X, i.e., $E|g(X)^2| < \infty$ for all $g \in \mathcal{G}$.

Question 4. Let F be any cdf with finite mean θ and variance $\tau^2 > 0$, and $X_1, \ldots, X_n$ be i.i.d., where $X_i \sim F$ with probability $\delta \in (0,1)$ and $X_i \sim N(\mu, \sigma^2)$ with probability $1-\delta$. Prove that

$$\operatorname{Var}(\bar{X}) = (1-\delta)\frac{\sigma^2}{n} + \delta\frac{\tau^2}{n} + \delta(1-\delta)\frac{(\theta-\mu)^2}{n}, \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Question 5. Let $X_1, \ldots, X_n$ be i.i.d. Poisson(λ), and let λ have a $\Gamma(\alpha, \beta)$ distribution with $\alpha, \beta > 0$, the conjugate family for the Poisson.

(a) Find the posterior distribution of λ.

(b) Calculate the posterior mean and variance.

Question 6. Let $X_1, \ldots, X_n$ be i.i.d. uniform random variables on the interval $[0, \theta]$ with unknown $\theta > 0$.

(a) Find the mle for θ, denoted as $\hat{\theta}_{\mathrm{mle}}$.

(b) Show that $\hat{\theta}_{\mathrm{mle}}$ is a consistent estimator for θ.

Question 7. Let $X_1, \ldots, X_n$ be i.i.d. random samples from the pdf

$$f(x) = 1, \quad \theta - \frac{1}{2} < x < \theta + \frac{1}{2},$$

where θ is an unknown parameter.

(a) Find a point estimator for θ.

(b) Find a $100(1-\alpha)\%$ (or approximate $100(1-\alpha)\%$) confidence interval for θ, where $0 < \alpha < 1$.

Question 8. Let $X_1, \ldots, X_n$ be i.i.d. random samples from the pdf

$$f(x) = \theta x^{\theta-1}, \quad 0 \leq x \leq 1,$$

where $\theta > 0$ is an unknown parameter.

(a) Find a point estimator for θ.

(b) Find a level α (or approximate level α) test for testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ for some given $\theta_0 > 0$, where $0 < \alpha < 1$.

Question 9. Let X be a random variable taking on the values 0 and 1 with probabilities p and $1-p$, respectively, where $p \in [1/3, 2/3]$ is an unknown parameter. We consider the estimation of p based on a single observation X.

(a) Find the mle for p, denoted as $\hat{p}_{\text{mle}}$.

(b) Let $\hat{p}_{\text{naive}} = 1/2$ be the naive estimator that always estimates p as $1/2$. Determine which estimator is better by comparing their mean squared errors.

Question 10. State and prove the Central Limit Theorem for a sequence of i.i.d. random variables. Please clearly specify the assumptions and conditions required for the theorem, and explain how they are utilized in your proof.