Probability and Statistics

PhD Qualification Exam

Mock Examination No. 5

Total: 100 points

- There are eight compulsory problems. The point values sum to 100.
- Give all necessary details. Unsupported final answers receive little credit.
- You may quote standard theorems only after verifying their assumptions.

Problem 1 **10 points**

Let $X \sim N_n(0, I_n)$ and let P be a deterministic symmetric idempotent matrix of rank r, where $0 < r < n$.

(a) Prove that PX and $(I - P)X$ are independent Gaussian vectors. [3]

(b) Identify the joint distribution of $\|PX\|^2$ and $\|(I - P)X\|^2$. [3]

(c) Identify the distributions of

$$\frac{\|PX\|^2/r}{\|(I - P)X\|^2/(n - r)} \quad \text{and} \quad \frac{\|PX\|^2}{\|X\|^2}.$$

[4]

Problem 2 **15 points**

For each n, let $a_{n1}, \dots, a_{nm_n}$ be deterministic real numbers such that

$$\sum_{k=1}^{m_n} a_{nk}^2 = 1, \qquad \max_{1 \le k \le m_n} |a_{nk}| \longrightarrow 0.$$

Let $X_1, X_2, \dots$ be i.i.d. with $\mathbb{E}X_1 = 0$, $\mathrm{Var}(X_1) = 1$, and $\mathbb{E}|X_1|^{2+\delta} < \infty$ for some $\delta > 0$.

(a) Prove that $\sum_{k=1}^{m_n} a_{nk} X_k \xrightarrow{d} N(0, 1)$ by verifying a named triangular-array CLT condition. [9]

(b) Show by counterexample that the conclusion can fail if the maximum-weight condition is removed, even when all other assumptions hold. [4]

(c) State a weaker standard condition that can replace the $(2 + \delta)$ moment assumption. [2]

Problem 3 **10 points**

Let $H \in L^2$ and let $(\mathcal{F}_n)$ be an increasing sequence of sigma-fields. Set

$$M_n = \mathbb{E}(H \mid \mathcal{F}_n), \qquad \mathcal{F}_\infty = \sigma\left(\bigcup_{n\geq 1} \mathcal{F}_n\right).$$

(a) Prove that $(M_n, \mathcal{F}_n)$ is an L^2-bounded martingale. [3]

(b) Prove that its increments are orthogonal in L^2 and derive

$$\sum_{k=1}^{n-1} \mathbb{E}(M_{k+1} - M_k)^2 = \mathbb{E}M_n^2 - \mathbb{E}M_1^2.$$

[3]

(c) Prove that $M_n \to \mathbb{E}(H \mid \mathcal{F}_\infty)$ in L^2. [4]

Problem 4 **15 points**

Let $(B_t)_{0\leq t\leq 1}$ be standard Brownian motion and define the ordinary path integral

$$A = \int_0^1 B_s \, ds.$$

(a) Compute $\operatorname{Var}(A)$ and $\operatorname{Cov}(B_t, A)$ for $0 \leq t \leq 1$. [4]

(b) Find the conditional distribution of B_t given $A = a$. [4]

(c) Construct a centered Gaussian process (W_t) that is independent of A and such that $B_t = W_t + \mathbb{E}(B_t \mid A)$. Give its covariance function. [4]

(d) Compute $\mathbb{P}(B_1 > 0 \mid A = a)$. [3]

Problem 5 **15 points**

Let $X_1, \ldots, X_n$ be i.i.d. with density

$$f_\theta(x) = 2(x-\theta)\mathbf{1}_{\{\theta<x<\theta+1\}}, \qquad \theta \in \mathbb{R}, \quad n \geq 2.$$

(a) Verify normalization and write the sample likelihood with its full parameter-dependent support. [3]

(b) Use the likelihood-ratio criterion to prove that the full vector of order statistics $T = (X_{(1)}, \ldots, X_{(n)})$ is minimal sufficient. [8]

(c) Prove that T is not complete by constructing a nonzero unbiased-zero function of T. [4]

Problem 6 **10 points**

Let $X_1, \ldots, X_n$ be i.i.d. from a continuous distribution F having a unique median m, so $F(m) = 1/2$. If an unbounded endpoint is needed, use the conventions $X_{(0)} = -\infty$ and $X_{(n+1)} = +\infty$.

(a) Use order statistics and a binomial probability to construct a finite-sample, distribution-free confidence interval for m with coverage at least $1 - \alpha$. State exactly how its index is chosen. [6]

(b) If F has a density f continuous and positive at m, state and justify the asymptotic distribution of the sample median. [4]

Problem 7 **15 points**

Let $(N_1, N_2, N_3) \sim \text{Multinomial}(n; p_1, p_2, p_3)$, where all null probabilities p_{10}, p_{20}, p_{30} are positive and sum to one. Test

$$H_0 : (p_1, p_2, p_3) = (p_{10}, p_{20}, p_{30}) \quad \text{against the unrestricted alternative.}$$

(a) Derive the unrestricted MLE and the likelihood-ratio statistic

$$G^2 = 2 \sum_{i=1}^{3} N_i \log \frac{N_i}{np_{i0}},$$

including the convention required when $N_i = 0$. [5]

(b) Derive the asymptotic null distribution of G^2, stating the dimension and regularity conditions behind Wilks' theorem. [4]

(c) Let

$$X^2 = \sum_{i=1}^{3} \frac{(N_i - np_{i0})^2}{np_{i0}}.$$

Prove that $G^2 - X^2 \xrightarrow{p} 0$ under H_0, and give the resulting asymptotic level-α rule. [6]

Problem 8 **10 points**

Suppose a scalar parameter is known to satisfy $\theta \in [a, b]$, where $a < b$, and squared-error loss is used. Let δ be any estimator taking real values, and define its projection

$$\delta^* = \min\{b, \max\{a, \delta\}\}.$$

(a) Prove that $R(\theta, \delta^*) \leq R(\theta, \delta)$ for every $\theta \in [a, b]$. [5]

(b) Give a condition for strict domination and explain the consequence for admissible estimators in this bounded-parameter problem. [5]