Statistics 655 Comprehensive Written Exam
August 2018

All problems have equal weight; partial credit will be given for each part of a problem. In budgeting your time, note that some parts will take longer than others. In many cases, different parts of a problem can be done independently, so if you are unsure how to handle one part of a problem, don't hesitate to try the others. No answer should require a great deal of computation or a complicated proof.

1. Let $\mathbb{R}^{n\times p}$ denote the set of $n \times p$ real matrices, and let T be a bounded subset of $\mathbb{R}^p$. Define a function $f : \mathbb{R}^{n\times p} \to \mathbb{R}$ by

$$f(A) = \inf_{v\in T} ||Av||^2$$

where $||\cdot||$ is the usual Euclidean norm on $\mathbb{R}^n$.

(a) Argue carefully that $\sqrt{f(\cdot)}$ is Lipschitz, taking care to specify the Lipschitz constant, and the appropriate norm.

(b) Let let $U = f(W)$ where W is an $n \times p$ array of i.i.d. standard normal random variables. Find a simple upper bound on $\mathbb{P}(U \geq \mathbb{E}U + 2\sqrt{t\,\mathbb{E}U} + t)$ for $t > 0$. Hint: Relate the numerical quantity on the right hand side inside the probability to an expression involving $\mathbb{E}\sqrt{U}$ and $\sqrt{t}$.

2. Let $(U_1, V_1), (U_2, V_2), \ldots \in \mathbb{R}^2$ be i.i.d. such that $\mathbb{E}U_i = 1$, $\mathbb{E}V_i = 2$, $\mathrm{Var}(U_i) = \mathrm{Var}(V_i) = 1$, and $\mathrm{Cov}(U_i, V_i) = \rho$. What can you say about the limiting behavior of

$$\left(\sum_{i=1}^{n} U_i\right) \log\left(\frac{\sum_{j=1}^{n} U_j}{\sum_{k=1}^{n} V_k}\right)?$$

3. Let $\mathcal{P} = \{\mathcal{N}(\mu, \sigma^2) : \mu \in (-\infty, \infty),\ \sigma^2 > 0\}$ be the family of univariate normal distributions, and let $X_1, X_2, \ldots$ be i.i.d. with $X_i \sim \mathcal{N}(\mu_0, \sigma_0^2)$.

(a) Find the Fisher information of $\mathcal{P}$.

(b) Find a lower bound on the variance of any unbiased estimate of μ_0/σ_0 based on $X_1, \ldots, X_n$.

(c) What is the maximum likelihood estimate of $\theta_0 = (\mu_0, \sigma_0^2)$ based on $X_1, \ldots, X_n$? (You may simply give the answer – no calculations are necessary.) What can be said about the limiting behavior of this estimate as n tends to infinity? Briefly justify your answer.

4. Let $\{U_n\}_{n\geq 1}$ and $\{V_n\}_{n\geq 1}$ be sequences of random variables in $\mathbb{R}$ and let $\{X_n\}_{n\geq 1}$ be a sequence of random vectors in $\mathbb{R}^d$, with all random quantities being defined on the same probability space. Suppose that $U_n \to U$ in probability, $V_n = 2 + o_P(1)$, and $X_n \Rightarrow X$ in law with $\mathbb{P}(X = \mathbf{0}) = 0$. In each of the following cases, (i) indicate whether the quantity converges as n tends to infinity, (ii) identify the limit if one exists, and (iii) identify the type of convergence. Justify your answers.

(a) U_n/V_n

(b) $X_n \log(V_n)$

(c) $\mathbb{E}[V_n/(1 + |V_n|)]$

(d) $X_n \, \mathbb{E}[V_n]$

(e) $U_n \, X_n$

(f) $||V_n \, X_n||^{-1}$

5. Let $X \sim \mathcal{N}_d(0, \Sigma_X)$ and $Y \sim \mathcal{N}_d(0, \Sigma_Y)$ be d-dimensional multinormal random vectors such that $\Sigma_Y - \Sigma_X$ is non-negative definite.

(a) Carefully define a random vector $Z \in \mathbb{R}^d$ with the property that $X + Z$ and $X - Z$ have the same distribution as Y.

(b) Let $C \subseteq \mathbb{R}^d$ be a convex set. Show that $\mathbb{P}(X \in C^c) \leq 2\mathbb{P}(Y \in C^c)$.