Statistics 655 Comprehensive Written Exam
August 2023

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 $A \subseteq \mathbb{R}^n$ be a closed set, and let $x^* \in A$ be fixed. Note that A may be unbounded. Suppose that we observe $y_i = x_i^* + \varepsilon_i$ for $i = 1, \ldots, n$, where $\varepsilon_1, \ldots, \varepsilon_n$ are independent noise variables with mean zero and variance one. Let $y = (y_1, \ldots, y_n)^t$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^t$. Consider the estimate

$$\hat{x} = \arg\min_{x \in A} ||y - x||$$

(a) Argue as carefully as you can that $\hat{x}$ is well defined (that is, the minimum in the definition exists) with probability one.

(b) Show that $\mathbb{E}||y - \hat{x}|| \leq \sqrt{n}$.

(c) Show that $||\hat{x} - x^*||^2 \leq 2\langle \varepsilon, \hat{x} - x^* \rangle$

2. Let $(X_1, Y_1)^t, (X_2, Y_2)^t, \ldots \in \mathbb{R}^2$ be iid such that $\mathbb{E}X_i = 1$, $\mathbb{E}Y_i = 2$, $\text{Var}(X_i) = \text{Var}(Y_i) = 1$, and $\text{Cov}(X_i, Y_i) = \rho$. What can you say about the limiting behavior of

$$\left(\sum_{i=1}^{n} X_i \right) \left(\sum_{i=1}^{n} Y_i \right)$$

after appropriate scaling and centering?

3. Let $X_1, X_2, \ldots, X \in \mathbb{R}^d$ be random vectors defined on the same probability space.

(a) Define the relations $X_n = O_p(1)$ and $X_n = o_p(1)$.

(b) Argue carefully that if X_n converges to X in probability then $X_n = O_p(1)$.

4. Let $A \subseteq \mathbb{R}^n$ be a convex set containing the zero vector, and let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^t$ where $\varepsilon_1, \ldots, \varepsilon_n$ are independent noise variables with mean zero and variance one. For $s > 0$ define the function

$$H(s) = \mathbb{E}\left[\sup_{x \in A,\, ||x|| \leq s} \langle x, \varepsilon \rangle \right]$$

(a) Show that $H(s) \geq 0$. Find an upper bound on $H(s)$ that involves s and the dimension n.

(b) Show that $H(s)$ is *non-decreasing.*

(c) Show that $H(s)/s$ is *non-increasing.*

(d) Suppose now that the noise variables $\varepsilon_1, \ldots, \varepsilon_n$ are iid $\mathcal{N}(0,1)$. For $t > 0$ find a bound on

$$\mathbb{P}\left(\sup_{x \in A,\, ||x|| \leq s} \langle x, \varepsilon \rangle > t \right)$$

(e) Use the probability bound in (d) to get a bound on $H(s)$ in the normal noise case. How does this bound compare to that in part (a)?