Statistics 654 Comprehensive Written Exam
August, 2020

Instructions: All problems have equal weight; partial credit will be given for each part of a problem. In some 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. Please show your work, and briefly explain your reasoning: correct answers with no work or explanation will not receive full credit.

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density

$$f(x|\theta, \gamma) \;=\; \frac{\theta\, \gamma^{\theta}}{x^{\theta+1}}\, \mathbb{I}(x \geq \gamma)$$

having parameters $\theta > 0$ and $\gamma > 0$. Note that $f(x|\theta,\gamma)$ is supported on $[\gamma, \infty)$.

a. Find the MLE of γ when θ is fixed. Does the estimate depend on θ?

b. Find the MLE of θ when γ is set to the value you found above.

c. Describe the general form of the likelihood ratio test statistic for testing

$$H_0 : \theta = 1 \text{ and } \gamma \text{ is arbitrary} \quad \text{vs} \quad H_1 : \theta \neq 1 \text{ and } \gamma \text{ is arbitrary},$$

and write out the statistic using the results of a. and b. above. You need not simplify.

2. Let $X \sim \mathcal{N}_d(\mu, \Sigma)$ be a multinormal random vector. Define $Y = CX$ and $Z = DX$ where $C \in \mathbb{R}^{l\times d}$ and $D \in \mathbb{R}^{k\times d}$ are matrices.

a. What is the distribution of Y?

b. Find necessary and sufficient conditions on C and D under which Y and Z are independent. Justify your answer.

3. Let $\mathcal{P} = \{f(x|\theta) : \theta \in \Theta\}$ be a family of densities on a set $\mathcal{X}$, and suppose that we are interested in estimating θ from an observation $X \in \mathcal{X}$ with $X \sim f(x|\theta) \in \mathcal{P}$.

a. Define what is meant by an estimator and a loss function in this setting.

b. Define the risk function of an estimator.

c. Let π be a prior distribution on Θ. Define the Bayes risk of an estimator under π.

d. Let $\mathcal{D}$ be a family of estimators. Define what it means for an estimator to be admissible.

4. Let $V \subseteq \mathbb{R}^d$ be a finite set of vectors $v = (v_1, \ldots, v_d)^t$ with $L = \max_{v \in V} ||v||_2$, and let $\varepsilon_1, \ldots, \varepsilon_d$ be independent sign variables with $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = 1/2$. In answering the questions below, you may appeal to results from the lectures, but explain your reasoning.

a. Bound the moment generating functions of $\sum_{i=1}^d \varepsilon_i v_i$ in terms of the constant L.

b. Use the MGF bound to get an upper bound on $\mathbb{E}\left[\max_{v \in V} \sum_{i=1}^d \varepsilon_i v_i\right]$.

c. Use the MGF bound to get an upper bound on $\mathbb{P}(\max_{v \in V} \sum_{i=1}^d \varepsilon_i v_i > t)$ when $t > 0$.

5. Let $X_1, X_2, \ldots$ be i.i.d. positive random variables with finite expectation. For $n \geq 1$ define $S_n = X_1 + \cdots + X_n$.

a. Calculate $\mathbb{E}(S_n/S_m)$ when $m \geq n$.

b. Find a lower bound for $\mathbb{E}(S_n/S_m)$ when $n \geq m$.

6. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function that is Lipschitz with constant L, and let $U \in \mathbb{R}^d$ be uniformly distributed in the unit ball $B(1) = \{x : ||x|| = 1\}$. Fix $\delta > 0$ and define a new function $g(x) := \mathbb{E}f(x + \delta U)$. You may assume that g is well-defined for each x.

a. What is the expected value of U?

b. Is $g \leq f$, $g \geq f$, or neither? Justify your answer.

c. Is g convex, concave or neither?

d. Is g Lipschitz?

e. What can you say about $\sup_x |g(x) - f(x)|$?