# Written Comprehensive Examination

# 2nd Year Theory Exam – 220A, 220B

# Department of Statistics, UC Irvine Wednesday, June 17, 2020, 9:00 am to 1:00 pm

- This is a closed book and notes examination. You are to answer exactly 5 of the following 6 questions. Use your time wisely. Clearly justify each step. The 5 questions you choose to answer will be worth equal credit.
- Your solutions to each problem should be written on separate sheets of paper. Label each sheet with your student identification number, the problem number, and the page number of that solution written in the upper right hand corner. For example, the labeling on a page may be:

ID# 912346378<br>Problem 2, page 3

- You have 4 hours to complete your solution. Please be prepared to turn in your exam at 1:00pm.

1. For a nonnegative random variable $X$, define its Laplace transform as $\psi_X(t) = Ee^{-tX}$.

   (a) Show that $\psi_X(t)$ exists and is continuous in $t \in [0, \infty)$ for any random variable $X \geq 0$.

   (b) Suppose $X \geq 0$ and let $F(x) = P(X \leq x)$. Prove the formula $\psi_X(t) = \int_0^\infty te^{-tx}F(x)dx$ for $t > 0$.

   (c) Let $X_n, n = 1, 2, \ldots$ be a sequence of nonnegative random variables such that $X_n \xrightarrow{D} X$. Show that $\psi_{X_n}(t) \to \psi_X(t)$ for every $t \geq 0$.

   (d) It is known that if two nonnegative random variables have the same Laplace transform, then they have the same distribution. Use this to establish the following result. Let $U_k, k = 1, 2, \ldots, n$, be independent Uniform$(0, 1)$ random variables. Then $\sum_{k=1}^n U_k$ has Lebesgue density

$$f(x) = \frac{1}{(n-1)!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}(x-k)^{n-1}I(x \geq k),$$

   where $I(\cdot)$ is the indicator function.

2. Let $X_n \sim \text{Expo}(\lambda_n)$ independently where the exponential distribution has Lebesgue density $f_n(x) = \lambda_n e^{-\lambda_n x}$ for $x \geq 0$.

   (a) Give necessary and sufficient conditions on $\lambda_n$ such that $X_n \to 0$ in probability. Justify.

   (b) Give necessary and sufficient conditions on $\lambda_n$ such that $X_n \to 0$ almost surely. Justify.

   (c) Suppose $\lambda_n = n$. Identify constants $b_n > 0$ such that $\sum_{j=1}^n X_j/b_n \to 1$ in probability. Justify.

   (d) Suppose $\lambda_n = 1/n$. Identify constants $a_n$ and $b_n > 0$ such that $(\sum_{j=1}^n X_j - a_n)/b_n \xrightarrow{D} \text{N}(0, 1)$. Justify.

3. **Uniform integrability**

   Consider a sequence of non-negative random variables $X_n$ defined on a probability space $(\Omega, \mathcal{A}, P)$. We say $\{X_n\}$ is uniformly integrable (u.i.) if $\sup_n E\{X_n I\{X_n > C\}\} \to 0$ as $C \to \infty$.

   (a) Prove a sufficient condition for u.i.: if $\sup_n EX_n^{1+\delta} < \infty$ for some $\delta > 0$, then $\{X_n\}$ is u.i.

   (b) Prove an *equivalent definition* of u.i.: (a) $\sup_n EX_n < \infty$; and (b) for any $\epsilon > 0$, there exists $\delta > 0$ such that for every measurable set $A$ with $P(A) \leq \delta$, then $E(X_n I(A)) \leq \epsilon$ for every $n$.

   (c) Now prove this result: if $X_n \xrightarrow{p} X$ as $n \to \infty$ and $\{X_n\}$ is u.i., then $X_n \xrightarrow{L_1} X$.

   (d) Is the converse of part (c) true? If yes, prove it; if not, give a counter-example and explain.

   (e) Consider an example where $X_n$ satisfies $P(X_n = 0) = 1 - n^{-k}$ and $P(X_n = n) = n^{-k}$. What is the requirement on $k$ to have $X_n \xrightarrow{p} 0$? What is the requirement on $k$ to have $X_n \xrightarrow{L_1} 0$? What happens in between?

4. **Decision theory**

Let $\theta$ be a parameter that only takes two values, $\theta_1$ and $\theta_2$. Consider a class of decision rules $D = \{d(\alpha) : \alpha \in [0, 2\pi]\}$. For $d(\alpha)$, its risk function $R(\cdot, \cdot)$ satisfies $R(d(\alpha), \theta_1) = 4 + \cos\alpha, R(d(\alpha), \theta_2) = 2 + \sin\alpha$.

(a) Draw a picture for the risk set spanned by $D$. Identify the class of admissible rules and explain why.

(b) Find the minimax rule and its risk function.

(c) Consider a prior on $\theta$, i.e., $\pi(\theta_1) = p, \pi(\theta_2) = 1 - p$ for some $p \in [0, 1]$. Find the Bayes rule and its Bayes risk.

(d) Now consider a randomized rule based on the class $D$, i.e., $d = \sum_{i=1}^{K} w_i d(\alpha_i)$, where $w_1, \ldots, w_K > 0, \sum_{i=1}^{K} w_i = 1, 0 \le \alpha_1 < \cdots < \alpha_K \le 2\pi$ for a finite integer $K > 1$. Is it possible for $d$ to be admissible? If yes, give an example and explain. If not, prove it.

(e) Now redo part (d) by letting $K = \infty$. In other words, consider a randomized rule, $d^* = \sum_{i=1}^{\infty} w_i d(\alpha_i)$, where $w_1, w_2, \ldots > 0, \sum_{i=1}^{\infty} w_i = 1, 0 \le \alpha_1 < \alpha_2 < \cdots \le 2\pi$. Is it possible for $d^*$ to be admissible? If yes, give an example and explain. If not, prove it.

5. Let $X$ be a random variable with density function $f(x; \theta), x \in R, \theta \in \Theta \subset R$. For any two values $\eta, \psi \in \Theta$, the Kullback-Leibler divergence (KLD) is defined as follows:

$$D(\eta, \psi) = E_\psi \log\left\{\frac{f(x;\psi)}{f(x;\eta)}\right\} = \int_R \log\left\{\frac{f(x;\psi)}{f(x;\eta)}\right\} f(x;\psi)dx,$$

where $E_\psi$ denotes the expectation taken under the density of $f(x; \psi)$. It is known that $D(\eta, \psi) \ge 0$ and $D(\eta, \psi) = 0$ if and only if $f(x, \eta) = f(x, \psi), a.s.$

Let $X_1, \ldots, X_n$ be an independent and identically distributed sample drawn from $f(x; \theta)$. To test the hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, we apply the likelihood ratio test approach. The log-likelihood ratio statistic is given by

$$LR(\boldsymbol{X}) = \log L(\boldsymbol{X}, \theta_0, \theta_1) = \log \prod_{i=1}^{n} \frac{f(X_i;\theta_1)}{f(X_i;\theta_0)} = \sum_{i=1}^{n} \ell(X_i),$$

where $\ell(X_i) = \log\{f(X_i; \theta_1)/f(X_i; \theta_0)\}$. Let $\mu_0 = E_{\theta_0}\ell(X)$, $\mu_1 = E_{\theta_1}\ell(X)$, and $\sigma_0^2 = Var_{\theta_0}(\ell(X)) < \infty$, $\sigma_1^2 = Var_{\theta_1}(\ell(X)) < \infty$.

(a) Let $u(x; \theta)$ be the score function of a single observation, i.e. $u(x; \theta) = \partial \log f(x; \theta)/\partial\theta$. Consider a neighborhood of parameter $\theta_0$, $N_0 \subset \Theta$. Show that

$$E_{\theta_1} u(X; \theta_0) = I(\theta_0)(\theta_1 - \theta_0)(1 + o(1)), \ \theta_1 \in N_0,$$

where $I(\theta_0) < \infty$ is the Fisher Information of a single observation under $H_0$, and $o(1)$ is in the sense of $|\theta_1 - \theta_0| \to 0$.

(b) Let $D(\theta_0, \theta_1)$ be the KLD of a single observation. Assume that the Fisher information $I(\theta)$ is continuous in $\theta \in N_0$. Using the result in Part (a) show that

$$D(\theta_0, \theta_1) = \frac{1}{2} I(\theta_0)(\theta_1 - \theta_0)^2\{1 + o(1)\}, \ \theta_1 \in N_0.$$

(c) Using the asymptotic normality of $LR(\boldsymbol{X})$, construct a size $\alpha$ Neyman-Pearson (N-P) test for $H_0 : \theta = \theta_0$ versus (a local alternative) $H_1 : \theta = \theta_0 + \delta/n^\lambda$, where both $\delta$ and $\lambda > 0$ are given constants.

(d) Give a range of $\lambda$ over which the N-P test given in Part (c) is consistent, namely its power increases to 1 when the sample size $n \to \infty$.

6. Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function $F$ and continuous density function $f > 0$. Let $\mathbb{F}_n$ be the empirical distribution function of the $X_i$'s, in other words, $\mathbb{F}_n(x) = (1/n)\sum_{i=1}^n I(X_i \le x)$, where $I(\cdot)$ is an indicator function. For a sequence of positive numbers $b_n$, define a uniform kernel density estimator of $f$ by

$$\hat{f}_n(x) = \frac{\mathbb{F}_n(x + b_n) - \mathbb{F}_n(x - b_n)}{2b_n}.$$

(a) Show that $E\left(\hat{f}_n(x)\right) \to f(x)$ if $b_n \to 0$.

(b) Show that $Var\left(\hat{f}_n(x)\right) \to 0$ if $b_n \to 0$ and $nb_n \to \infty$.

(c) Show that if $b_n \to 0$ and $nb_n \to \infty$,

$$\sqrt{2nb_n}\left\{\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right\} \to_d N(0, f(x)),$$

where $\to_d$ means convergence in distribution.

(d) Now consider $x, y \in R$, the set of real numbers, with $x \neq y$. Under the same assumptions in (c), show that

$$\sqrt{2nb_n}\begin{pmatrix} \hat{f}_n(x) - E\hat{f}_n(x) \\ \hat{f}_n(y) - E\hat{f}_n(y) \end{pmatrix} \to_d \text{“something”}$$

and find “something”.