Department of Statistics
University of Wisconsin, Madison
PhD Qualifying Exam Option A
August 28, 2018
12:30-4:30pm, Room 133 SMI

- There are a total of FOUR (4) problems in this exam. Please do all FOUR (4) problems.
- Each problem must be done in a separate exam book.
- Please turn in FOUR (4) exam books.
- Please write your code name and **NOT** your real name on each exam book.

1. Let X, Y, and Z be random variables defined in a common probability space.

Definition. X and Y are independent conditional on Z, denoted as $X \perp Y \mid Z$, if and only if

$$P(A \mid Y, Z) = P(A \mid Z) \text{ a.s. for any } A \in \sigma(X)$$

where $\sigma(U)$ denotes the σ-field generated by the random variable U.

(a) Show that $X \perp Y \mid Z$ if and only if

$$P(A \cap B \mid Z) = P(A \mid Z)P(B \mid Z) \text{ a.s. for any } A \in \sigma(X) \text{ and } B \in \sigma(Y)$$

(b) Show that $X \perp Y \mid Z$ implies that $g(X) \perp h(Y) \mid Z$ for any Borel functions g and h.

(c) Define $\phi_{X,Y|Z}(t,s) = E[e^{i(tX+sY)} \mid Z]$, $\phi_{X|Z}(t) = E[e^{itX} \mid Z]$, and $\phi_{Y|Z}(s) = E[e^{isY} \mid Z]$, a.s., where t and s are real numbers and $i = \sqrt{-1}$. Show that $X \perp Y \mid Z$ if and only if

$$\phi_{X,Y|Z}(t,s) = \phi_{X|Z}(t)\phi_{Y|Z}(s) \quad \text{for all } t \text{ and } s \text{ a.s.}$$

(d) Provide an example in which $X \perp Y \mid Z$ but X and Y are not independent. Provide an example in which X and Y are independent but $X \perp Y \mid Z$ does not hold.

2. Let $\bar{X}_n$ and s_n^2 denote the mean and variance of a random sample of n observations from a distribution F with mean θ and finite variance σ^2. The level-α t-test of

$$H_0 : \theta = 0 \text{ vs } H_1 : \theta > 0$$

rejects H_0 if $\sqrt{n} s_n^{-1} \bar{X}_n > t_{n-1,1-\alpha}$, where $t_{\nu,\alpha}$ denotes the α-quantile of the t distribution with ν degrees of freedom.

Let $\alpha_n(F) = P_F(\sqrt{n} s_n^{-1} \bar{X}_n > t_{n-1,1-\alpha} \,|\, \theta = 0)$ denote the probability of rejecting H_0 when $\theta = 0$ (Type I error). The level of the t-test is often said to be *robust* against non-normality because $\alpha_n(F) \to \alpha$ as $n \to \infty$ for any F with finite σ^2. Given a class of distributions $\mathcal{F}$, a test is said to be *uniformly robust* over $\mathcal{F}$ if $\sup_{\mathcal{F}} \alpha_n(F) \to \alpha$ as $n \to \infty$.

(a) Prove that the t-test is not uniformly robust over the class $\mathcal{F}$ of distributions with finite variance by showing that for every α, $\sup_{\mathcal{F}} \alpha_n(F) = 1$ for sufficiently large n. [Hint: consider $F = (1-\epsilon)\, U(\mu_1 - \epsilon, \mu_1 + \epsilon) + \epsilon\, U(\mu_2 - \epsilon, \mu_2 + \epsilon)$ for suitable values of μ_1, μ_2 and $0 < \epsilon < 1$, where $U(a, b)$ denotes the uniform distribution on the interval (a, b).]

(b) Prove that the t-test is not uniformly robust even if $\mathcal{F}$ is restricted to the class of absolutely continuous distributions with uniformly bounded support.

3. Consider a linear regression model $y_i = x_i^T\beta + \epsilon_i$, where $(x_i, y_i) \in \mathbb{R}^p \times \mathbb{R}$, $x_i \sim \mathcal{N}(0, \Sigma)$ and $\epsilon_i \sim \mathcal{N}(0, 1)$. Furthermore, assume that $\{(x_i, y_i)\}_{i=1}^n$ are independent and identically distributed samples where x_i and ϵ_i are independent for all i. Here, $\Sigma \in \mathbb{R}^{p\times p}$ is a positive definite matrix and $\beta \in \mathbb{R}^p$ are unknown parameters of interest.

   (a) Find equations satisfied by the maximum likelihood estimators MLE $\widehat{\beta}_{MLE}$ for β.

   (b) Find equations satisfied by the maximum likelihood estimators MLE $\widehat{\Sigma}_{MLE}$ for Σ (*Hint:* For a positive definite matrix $A \in \mathbb{R}^{p\times p}$, $\frac{d \log \det A}{dA} = A^{-1}$).

   (c) Determine conditions under which the MLEs are unique. Are $\widehat{\beta}_{MLE}$ and $\widehat{\Sigma}_{MLE}$ independent? Justify your answer.

   (d) Find ν such that $n^\nu(\widehat{\beta}_{MLE} - \beta)$ converges to a non-trivial limit and find the limiting distribution. State any conditions you need for your result.

   (e) Find ν such that $n^\nu(\widehat{\Sigma}^{jk}_{MLE} - \Sigma^{jk})$ converges to a non-trivial limit and find the limiting distribution for every pair (j, k). State any conditions you need for your result. Note thatA^{jk} denotes the $(j, k)^{th}$ element for the matrix A.

4. This problem has two separate parts.

(a) Suppose $n \geq 2$ is some fixed integer. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with probability function

$$P_\theta(X = 0) = 1 - \theta, \quad P_\theta(X = j) = \theta/5, \quad \text{for } j = 1, \ldots, 5.$$

Here, $0 \leq \theta \leq 1$.

i. Prove or disprove that $T = \sum_{i=1}^n X_i$ is a sufficient statistic for θ. If T is not sufficient, find a non-trivial sufficient statistic (i.e., not $X_1, \ldots, X_n$) for θ and prove it.

ii. Is there a UMVU estimator for θ? If so, find and prove it. If not, explain why.

(b) Let S be a collection of an unknown number, p, of colored balls where each ball has a different color. A sample of n balls is randomly drawn from S with replacement.

i. Find a non-trivial sufficient statistic for p. Prove that it is indeed sufficient.

ii. Find an unbiased estimator for p or prove that no unbiased estimator exists. If an unbiased estimator exists, find out whether it is UMVU.