UNIVERSITY OF MASSACHUSETTS
DEPARTMENT OF MATHEMATICS AND STATISTICS
ADVANCED EXAM - STATISTICS (II)
10:00 AM - 1:00 PM, August 27, 2021

Work all problems and show all work. Explain your answers. State the theorems used whenever possible. 70 points are required to pass.

1. Let $\{X_n\}_{n\geq 1}$ be a sequence of real-valued random variables with distribution function $F_n(x)$ for each n and X be another real-valued random variable with distribution function $F(x)$.
   (a) (3 points) State the definition of almost sure convergence (denoted as $X_n \xrightarrow{a.s.} X$).
   (b) (3 points) State the definition of convergence in ath mean (denoted as $X_n \xrightarrow{a} X$).
   (c) (3 points) State the definition of convergence in probability (denoted as $X_n \xrightarrow{P} X$).
   (d) (3 points) State the definition of convergence in distribution (denoted as $X_n \xrightarrow{d} X$).
   (e) (5 points) Prove that for fixed $a > 0$, $X_n \xrightarrow{a} X$ implies $X_n \xrightarrow{P} X$.
   (f) (6 points) Show that $\bar{X}_n = \mu + O_p\left(\frac{1}{\sqrt{n}}\right)$ as $n \to \infty$ where $\bar{X}_n$ is the sample mean of n independent and identically distributed random variables with mean μ and finite variance σ^2. Note that the definition of O_p is as follows: $X_n = O_p(Y_n)$ if for every $\epsilon > 0$, there exist M and N such that $P(|X_n/Y_n| < M) > 1 - \epsilon$ for all $n > N$ where $\{Y_n\}_{n\geq 1}$ is a sequence of real-valued random variables with distribution function $G_n(y)$ for each n.

2. Suppose $Y_1, Y_2, \ldots,$ is a simple random sample from an exponential distribution with density $g(y) = \theta \exp(-\theta y)$ where $\theta > 0$ and $y > 0$. Consider the estimator of $h(\theta) = 1/\theta$, $\hat{h}_n = \sum_{i=1}^{n} Y_i/(n+2)$.
   (a) (4 points) Compute the bias of $\hat{h}_n$, denoted as $B(\hat{h}_n)$.
   (b) (4 points) Compute the variance of $\hat{h}_n$, denoted as $Var(\hat{h}_n)$.
   (c) (4 points) Show that $B(\hat{h}_n) \sim k_1 Var(\hat{h}_n) \sim k_2(1/n)$ as $n \to \infty$ for some constants k_1 and k_2 depending on θ. Note that the definition of $\sim$ is as follows: the sequence of real numbers $\{c_n\}_{n\geq 1}$ is asymptotically equivalent to the sequence $\{d_n\}_{n\geq 1}$, written as $c_n \sim d_n$ if $(c_n/d_n) \to 1$ as $n \to \infty$.

3. Suppose $X_1, X_2, \ldots,$ are independent and identically distributed with mean μ and variance σ^2. Let $Y_i = \bar{X}_i = (\sum_{j=1}^{i} X_j)/i$.
   (a) (3 points) Are $Y_1, Y_2, \ldots$ independent? Verify your answer.
   (b) (7 points) Show that $\bar{Y}_n = (\sum_{i=1}^{n} Y_i)/n$ is a consistent estimator of μ.

   [Hint] Use $Var(\bar{Y}_n) = \frac{\sigma^2}{n^2}\left(2n - \sum_{i=1}^{n} \frac{1}{i}\right)$.

(c) (8 points) Compute the limit of the relative efficiency of $\bar{Y}_n$ to $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, defined to be $\frac{Var(\bar{X}_n)}{Var(\bar{Y}_n)}$ as $n \to \infty$.

4. Answer the following questions.

(a) (10 points) Suppose that X is a random variable with $E(X) = 0$ and $Var(X) = \sigma^2 < \infty$. Let Z_n denote the random variable $X^2 I\{\mid X \mid \geq \sigma\sqrt{n}\}$. Prove that $E(Z_n) \to 0$ as $n \to \infty$.

(b) (15 points) Suppose that $X_1, X_2, \ldots,$ are independent and identically distributed with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Let $k_{n1}, \ldots, k_{nn}$ be constants satisfying

$$\frac{\max_{i \leq n} k_{ni}^2}{\sum_{j=1}^n k_{nj}^2} \to 0 \quad \text{as} \quad n \to \infty.$$

Let $T_n = \sum_{i=1}^n k_{ni} X_i$. Prove that

$$\frac{T_n - E(T_n)}{\sqrt{Var(T_n)}} \xrightarrow{d} N(0,1).$$

[Hint] Use Lindeberg-Feller Central Limit Theorem by checking the Lindeberg condition.

5. Suppose that $X_1, \ldots, X_n$ are independent and identically distributed with the unknown distribution function $P(X_i \leq x) = F(x - \mu)$ for the mean parameter μ and the variance parameter σ^2 where $f(x) = F'(x)$ exists and is symmetric about 0. Note that σ^2 is not a function of μ. We want to test $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$. Consider the statistic $\bar{X}_n = \sum_{i=1}^n X_i/n$.

(a) (3 points) Prove that under $H_0 : \mu = 0$,

$$\sqrt{n}\bar{X}_n/\sigma \xrightarrow{d} Z,$$

where Z is the standard normal distribution with mean 0 and variance 1.

(b) (3 points) Assume that one rejects $H_0 : \mu = 0$ whenever

$$\bar{X}_n > u_\alpha \sigma/\sqrt{n},$$

where u_α is the $1 - \alpha$ quantile of the standard normal distribution. Show that this test has asymptotic level α using (a).

(c) (10 points) Show that the following asymptotic result holds for the alternatives $\{\mu_n\}$ satisfying $\mu_n > 0$ for all n:

$$\sqrt{n}\frac{(\bar{X}_n - \mu_n)}{\sigma} \xrightarrow{d} Z,$$

where μ_n means that the mean parameter depends on n.

(d) (6 points) Suppose that $\mu_n > 0$ for all n and $\sqrt{n}\mu_n \to \delta > 0$. Let $Power_n(\mu_n)$ be the power of this test against the alternative μ_n. Show that

$$Power_n(\mu_n) \to \Phi\left(\frac{\delta}{\sigma} - u_\alpha\right) \quad \text{as } n \to \infty$$

where $\Phi(x)$ denotes the cumulative distribution function of the standard normal distribution.