Math 37600 PR (19364) Test 2 Solutions Fall 2024

In each of the following problems, $X_1, X_2, \ldots$ is an iid sequence of random variables each with mean μ and variance σ^2.

1- (i) State Chebyshev's Inequality for a random variable Y with mean μ and variance σ^2.

(ii) Assume that the the rv's in the iid sequence $X_1, X_2 \ldots,$ have the pdf $f(x, \theta)$. Let $T_n(X_1, \ldots, X_n)$ be a statistic for each $n = 1, 2, \ldots$.

Define what it means for the sequence $\{T_n\}$ to be a *consistent sequence* of estimators for θ.

(iii) Prove: If each T_n is an unbiased estimator and the variance of T_n is σ_n^2 with the sequence $\{\sigma_n^2\}$ tending to zero, then $\{T_n\}$ is a consistent sequence of estimators for θ.

(i) $P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}$.

(ii) $Lim_{n\to\infty} P_\theta(|T_n - \theta| \geq \epsilon) = 0$ for any $\epsilon > 0$. Equivalently, the sequence $\{T_n\}$ converges to the constant θ in probability (i.e. converges (P)) with respect to the pdf $f(x, \theta)$.

(iii) $E_\theta(T_n) = \theta$ because each T_n is unbiased. So for $\epsilon > 0$

$$P_\theta(|T_n - \theta| \geq \epsilon) = P_\theta(|T_n - \theta| \geq (\frac{\epsilon}{\sigma_n})\sigma_n) \leq \frac{\sigma_n^2}{\epsilon^2}.$$

So the limit is 0 as $n \to \infty$.

2- Assume that the the rv's in the iid sequence $X_1, X_2 \ldots,$ have the uniform distribution $Unif(0, \theta)$.

(i) For each positive n we use $X_1, \ldots, X_n$ as a sample. Explain why $M_n = \max(X_1, \ldots, X_n)$ is the MLE for θ, and compute the MLE for the mean μ of the distribution $Unif(0, \theta)$.

(ii) Compute the cdf, pdf and expected value of M_n. Show that M_n is a biased estimate of θ.

(iii) Prove that the sequence $\{M_1, M_2, \ldots\}$ is a consistent estimator sequence for θ (Hint: Use the cdf of M_n).

(i) $L(\theta; x_1, \ldots, x_n) = 1/\theta^n$ provided $x_1, \ldots, x_n \leq \theta$ and $= 0$ otherwise. Because $1/\theta^n$ is a decreasing function of θ, the maximum occurs at the smallest value with $\theta \geq x_1, \ldots, x_n$, or, equivalently, $\theta \geq \max(x_1, \ldots, x_n)$. This means that the MLE is $\hat{\theta} = M_n$.

The mean value μ of X is $\theta/2$ and so the MLE for the mean is $M_n/2$ and not $\bar{X}$.

(ii) The cdf is $F_X(t) = \frac{t}{\theta}, \quad 0 < t < \theta$. Since $M_n \leq t$ if and only if $X_1, \ldots, X_n \leq t$, the cdf of M_n is $F_{M_n}(t) = (\frac{t}{\theta})^n, \quad 0 < t < \theta$ and so the pdf of M_n is $f_{M_n}(t) = F'_{M_n}(t) = n\frac{t^{n-1}}{\theta^n}$ and so

$$E_\theta(M_n) = \int_0^\theta n\frac{t^n}{\theta^n}dt = \frac{n}{n+1}\theta.$$

So the estimate is biased.

(iii) For $0 \leq t \leq \theta$, $|t - \theta| \geq \epsilon$ exactly when $t \leq \theta - \epsilon$. So with θ and $\epsilon > 0$ fixed,

$$P_\theta(|M_n - \theta| \geq \epsilon) = P_\theta(M_n \leq \theta - \epsilon) = (\frac{\theta - \epsilon}{\theta})^n.$$

Since $0 \leq \frac{\theta-\epsilon}{\theta} < 1$ this tends to 0 as $n \to \infty$.

Technically, this needs $\epsilon \leq \theta$. If $\epsilon > \theta$, then $P_\theta(|M_n - \theta| \geq \epsilon) = 0$.

3- Assume that the the rv's in the iid sequence $X_1, X_2 \ldots,$ have the pdf $f(x, \theta)$ with $f(x, \theta) > 0$ for all $x \in \mathbb{R}$ and $\theta \in \Omega$. That is, the support of the rv's is the entire real line. Let C be a subset of $\mathbb{R}^n$.

We will use $\{X_1, \ldots, X_n\}$ as a sample to choose between the *Null Hypothesis* $\theta = \theta_0$ and the *Alternative Hypothesis* $\theta = \theta_1$.

(i) What does it mean to use C as a critical region for the choice? That is, define what a *critical region* is.

(ii) Given $\alpha > 0$, what does it mean to say that C has size α? That is, define the *size* of a critical region.

(iii) If C has size α, what does it mean to say that C is a *best critical region* of size α?

(i) We use C as a critical region when we choose by rejecting the null hypothesis if the realized value $(x_1, \ldots, x_n)$ lies in C,and so choose $\theta = \theta_1$. Otherwise, we accept the null hypothesis, choosing $\theta = \theta_0$.

(ii) The size of the critical region C is the probability of landing in C assuming the null hypothesis is true. That is, it is $P_{\theta_0}((X_1, \ldots, X_n) \in C)$. This is the probability of a Type I error. So size α means $\alpha = P_{\theta_0}((X_1, \ldots, X_n) \in C)$.

(iii) C is a best critical region when for any other critical region A of size α, the power of C is greater than or equal to that of A. That is,

$$P_{\theta_0}((X_1, \ldots, X_n) \in C) = P_{\theta_0}((X_1, \ldots, X_n) \in A) \implies$$

$$P_{\theta_1}((X_1, \ldots, X_n) \in C) \geq P_{\theta_1}((X_1, \ldots, X_n) \in A).$$

4- (i) Recall that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is the *sample mean*. Compute the expected value $E(\bar{X}_n)$, the variance $Var(\bar{X}_n)$ and the second moment $E(\bar{X}_n^2)$ (Hint:For an rv Y how are the variance $Var(Y)$ and the second moment $E(Y^2)$ related?).

(ii) Verify that $\bar{X}_n$ is an unbiased estimate for the mean μ. Verify the Weak Law of Large Numbers which says that the the sequence $\{\bar{X}_n\}$ converges in probability to the constant μ and so $\{\bar{X}_n\}$ is a consistent sequence of estimators for the mean μ (Hint: you may use the result quoted in Problem 1).

(iii) The sample variance S_n^2 is given by the equivalent formulae:

$$S_n^2 \quad = \quad \frac{1}{n-1}[\sum_{i=1}^n (X_i - \bar{X}_n)^2] \quad = \quad \frac{n}{n-1}[(\frac{1}{n}\sum_{i=1}^n X_i^2) - \bar{X}_n^2].$$

Note that one of the terms on the right is the sample mean for $\{X_1^2, \ldots, X_n^2\}$.

Compute the expected value $E(S_n^2)$ and the limit in probability of the sequence $\{S_n^2\}$. Verify that S_n^2 is an unbiased estimate of σ^2 and that the sequence $\{S_n^2\}$ is a consistent estimate of σ^2.

(iv) Explain why the sequence $\sqrt{S_n^2}$ is a consistent sequence of estimators for the standard deviation σ. Use Jensen's Inequality to show that $\sqrt{S_n^2}$ is a biased estimator of σ, that is, it is <u>not</u> an unbiased estimator.

(i) In general,

$$E(\bar{X}_n) = E(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n}\sum_{i=1}^n E(X_i) = \frac{1}{n}n\mu = \mu.$$

Because the rv's are independent

$$Var(\bar{X}_n) = Var(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n^2}\sum_{i=1}^n Var(X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

In general, $Var(Y) = E(Y^2) - E(Y)^2$ and so $E(Y^2) = E(Y)^2 + Var(Y)$. So

$$E(\bar{X}_n^2) \;=\; \mu^2 + \frac{\sigma^2}{n}.$$

(ii) Since $E(\bar{X}_n) = \mu$, $\bar{X}_n$ is an unbiased estimate for the mean. Since the variance $Var(\bar{X}_n) = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$, it follows from (iii) of problem 1, that the sequence $\{\bar{X}_n\}$ is consistent, and so $\{\bar{X}_n\} \to \mu$ in probability.

(iii)Because $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ is the sample mean for the iid sequence $\{X_i^2\}$ it follows that

$$E(\frac{1}{n}\sum_{i=1}^{n} X_i^2) = E(X^2) = E(X)^2 + Var(X) = \mu^2 + \sigma^2$$

Furthermore,

$$\{\frac{1}{n}\sum_{i=1}^{n} X_i^2)\} \to E(\frac{1}{n}\sum_{i=1}^{n} X_i^2)) = \mu^2 + \sigma^2 \quad (P).$$

Since, $E(\bar{X}_n^2) = \mu^2 + \frac{\sigma^2}{n}$ it follows that

$$E(S_n^2) = \frac{n}{n-1}[E(\frac{1}{n}\sum_{i=1}^{n} X_i^2) - E(\bar{X}_n^2)] = \frac{n}{n-1}[(\mu^2+\sigma^2)-(\mu^2+\frac{\sigma^2}{n})] = \sigma^2.$$

So S_n^2 is an unbiased estimate of σ^2.

On the other hand, since the limit of $\{\bar{X}_n\} = \mu$, it follows that the limit of $\{\bar{X}_n^2\} = \mu^2$

$$Lim\{S_n^2\} = Lim(\frac{n}{n-1})[Lim\{\frac{1}{n}\sum_{i=1}^{n} X_i^2\} - Lim\{\bar{X}_n^2\}] = 1\cdot[(\mu^2+\sigma^2)-\mu^2)] = \sigma^2,$$

and the sequence is consistent.

(iv) Because $\{S_n^2\}$ converges to σ^2 (P), it follows that $\{\sqrt{S_n^2}\}$ converges to σ (P) and so the sequence $\{\sqrt{S_n^2}\}$ is consistent for the standard deviation σ.

However, because the function $x \mapsto \sqrt{x}$ is concave down (bent down) for $x > 0$, it follows that for any non-constant, nonnegative rv Y, that $\sqrt{E(Y)} > E(\sqrt{Y})$. Therefore,

$$\sigma = \sqrt{E(S_n^2)} > E(\sqrt{S_n^2}\ ),$$

and so the estimate $\sqrt{S_n^2}$ for σ is biased.

5- Assume that the the rv's in the iid sequence $X_1, X_2 \dots,$ have the pdf $f(x, \theta) = \theta x^{\theta-1}$ with $0 < x < 1$ and $\theta > 0$.

(i) Compute the MLE for θ using the sample $X_1, \dots, X_n$.

(ii) In choosing between the Null Hypothesis of $\theta_0 = 1$ against the alternatives $\theta > 1$, explain why $C_n = \{(x_1, \dots, x_n) : x_1 \cdot \dots \cdot x_n \geq k\}$ is a UMP (= uniformly most powerful) critical region of its size for any k between 0 and 1.

(iii) With $n = 2$, so that $C_2 = \{(x_1, x_2) : x_1 \cdot x_2 \geq k\}$, compute the size of the test.

(i) The likelihood function is given by

$$L(\theta; x_1, \dots, x_n) = (\theta^n)(x_1 \cdot \dots x_n)^{\theta-1}$$

and so the log-likelihood is

$$\ell(\theta; x_1, \dots, x_n) = n \ln(\theta) + (\theta - 1) \ln(x_1 \cdot \dots x_n).$$

Taking the derivative with respect to θ and setting it equal to zero we obtain $\hat{\theta} = -n/\ln(x_1 \cdot \dots x_n)$. So that MLE is the statistic $-n/\ln(X_1 \cdot \dots X_n)$.

(ii) The ratio Λ for the likelihoods for θ_0 and an alternative θ_1 is given by

$$\Lambda = (\frac{\theta_0}{\theta_1})^n (x_1 \cdot \dots x_n)^{\theta_0 - \theta_1}.$$

If $\theta_1 > \theta_0$ then the exponent is negative and so $g(\theta_0, \theta_1, t) = (\frac{\theta_0}{\theta_1})^n t^{\theta_0 - \theta_1}$ with $0 < t < 1$ is a decreasing function of t. This is the Monotone Likelihood Condition.

Hence, for any k with $0 < k < 1$

$$x_1 \cdot \dots x_n \geq k \quad \Longleftrightarrow \quad \Lambda \leq (\frac{\theta_0}{\theta_1})^n k^{\theta_0 - \theta_1}.$$

So by the Neyman-Pearson Theorem, $C_k = \{(x_1, \dots, x_n) : x_1 \cdot \dots x_n \geq k\}$ is a best critical region of its size for any choice of the null hypothesis $\theta = \theta_0$ against any alternative $\theta = \theta_1$ with $\theta_1 > \theta_0$. So it is a Uniformly Most Powerful Test.

(iii) With $\theta_0 = 1$, the density is uniform on $[0, 1]$. So with $n = 2$ and $0 < k < 1$, the size is

$$\int_k^1 \int_{k/x_1}^1 dx_2 dx_1 = \int_k^1 1 - \frac{k}{x_1} \, dx_1 = 1 - k + k \ln(k).$$