COMPREHENSIVE EXAMINATION

STATISTICAL INFERENCE

Wednesday, August 15, 1990
9:00 a.m. – 11:00 a.m.

This is a closed-book, closed-notes exam. *Please start all problems on a new sheet of paper.*

1. Students seeking a Master's Level Pass may attempt any five problems from problems 1–6.
2. Students seeking a Ph.D. Level Pass must attempt any five problems from problems 6–12.

1. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with probability density function

$$f(x;\theta) = \begin{cases} \frac{2\theta^2}{x^3}, & x > \theta, \\ 0, & \text{otherwise,} \end{cases}$$

for some $\theta > 0$.

(a) Show that $\min(X_1, \ldots, X_n)$ is the maximum likelihood estimator of θ.

(b) Construct a $\gamma \times 100\%$ confidence interval for θ, $0 < \gamma < 1$.

2. Let $X_1, \ldots, X_n$ be independent and identically distributed with probability mass function

$$f(x|\theta) = \frac{e^{-\theta}\theta^x}{x!}, \quad x = 0, 1, 2, \ldots, .$$

(a) Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for θ. Is it minimal sufficient? Why?

(b) Find a UMVUE for θ^2. Justify.

(c) Argue that the test

$$\phi(\mathbf{X}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i > k, \\ 0, & \text{otherwise,} \end{cases}$$

where k is a constant, is UMP of its size for testing H_0: $\theta \le 1$ vs. H_1: $\theta > 1$.

3. Let

$$g(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}e^{-\beta/\theta}, & \theta > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where α and β are known, positive constants.

(a) Verify $g(\theta)$ is a probability density function.

(b) Let X be a normal random variable with mean zero and variance θ. Consider the estimation problem with $\Theta = \mathcal{A} = (0, \infty)$ and loss function $l(\theta, a) = (\theta - a)^2$. Compute the Bayes rule for the prior distribution $g(\theta)$ with $\alpha > 1$ and $\beta > 0$.

4. X_1, X_2 and X_3 are i.i.d. normal variables with mean 0 and variance σ_1^2. State the distributions of the following variables.

   (a) $2X_1 + 3X_2 - 5X_3$

   (b) $(X_1 - X_2)^2/2\sigma_1^2$

   (c) $2X_3^2/(X_1^2 + X_2^2)$

   (d) $X_3^2/(X_1^2 + X_2^2 + X_3^2)$

   (e) $\sqrt{3}\,\bar{X}/S_1$ where $\bar{X} = (X_1 + X_2 + X_3)/3$ and $S_1^2 = \Sigma(X_i - \bar{X})^2/2$.

   Furthermore, Y_1 and Y_2 are i.i.d. normal variables with mean 0 and variance σ_2^2, independent of X_i's. State also the distributions of the following variables:

   (f) $\bar{X} - \bar{Y}$, where $\bar{Y} = (Y_1 + Y_2)/2$

   (g) $V = \frac{2S_1^2}{\sigma_1^2} + S_2^2\sigma_2^2$, where $S_2^2 = \Sigma(Y_j - \bar{Y})^2$.

   (h) $T = k(\bar{X} - \bar{Y})/\sqrt{V}$, where $k = \sqrt{(18/5)}/\sigma$ when $\sigma_1 = \sigma_2 = \sigma$.

   (i) Determine k as a function of σ_1 and σ_2, if $\sigma_1 \neq \sigma_2$, so that T would have the same distribution as in (h).

5. (a) Suppose the random variable X has the pdf $p(x, \theta)$ where θ is an unknown, real-valued parameter. If $T = T(X)$ is an unbiased estimator of $g(\theta)$, establish the inequality

   $$\mathrm{var}_\theta(T) \geq \frac{[g'(\theta)]^2}{I(\theta)},$$

   where $I(\theta)$ is the Fisher information, stating the regularity conditions.

   (b) If X has binomial (n, θ) distribution, obtain the Cramer-Rao lower bound for unbiased estimation of $\theta(1 - \theta)$.

6. $Y_1, Y_2, \ldots, Y_n$ are independent normal variables with means

$$E(Y_i) = \alpha + \beta x_i\,, \quad i = 1, \ldots, n\,,$$

and variance σ^2, where x_i's are known constants such that $\Sigma x_i = 0$ and $\Sigma x_i^2 > 0$, and α, β, σ are unknown parameters.

(a) Show that $T = (\Sigma Y_i, \Sigma Y_i^2, \Sigma Y_i x_i)$ is minimal sufficient.

(b) Verify that $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \Sigma Y_i x_i / \Sigma x_i^2$ and $\hat{\sigma}^2 = S^2 = \Sigma(Y_i - \hat{\alpha} - \hat{\beta} x_i)^2/n$ are the maximum likelihood estimators.

(c) Show that the likelihood-ratio test of H_0: $\beta = 0$ against H: $\beta \neq 0$ is based on $\hat{\beta}^2/S^2$.

7. Let $X_1, X_2, \ldots, X_n$ be a random sample from the uniform distribution on the interval $(\theta, \theta + 1)$, θ unknown. To test

$$H_0\colon\ \theta = 0 \text{ vs. } H_A\colon\ \theta > 0\,,$$

the following procedure is used: Reject H_0 if and only if

$$\begin{array}{ll} \text{either} & \max\{x_1, \ldots, x_n\} > 1 \\ \text{or} & \max\{x_1, \ldots, x_n\} \leq 1 \text{ and } \min\{x_1, \ldots, x_n\} \geq C\,. \end{array}$$

(a) Determine the constant C so that the test will have size $\alpha = 0.05$.

(b) Prove or disprove: If C is chosen so that the test has size α, then it is U.M.P. among all tests of size α.

8. Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables with parameter p, i.e., $P(X_i = 1) = p = 1 - P(X_i = 0)$, $0 < p < 1$.

   (a) Show that no unbiased estimator of $\frac{1}{p}$ exists.

   (b) Propose an admissible estimator of the form $\frac{1}{aY+b}$ with $a \neq 0$ for $\frac{1}{p}$ under squared error loss.

9. Suppose F is Cauchy distribution with median θ, $-\infty < \theta < \infty$ unknown, i.e., $f(x, \theta) = \frac{1}{\pi[1+(x-\theta)^2]}$, $-\infty < x < \infty$. $X_1, X_2, \ldots, X_n$ are i.i.d. from F. Suppose we use the confidence interval

$$\left[X_{[r]}, X_{[s]}\right]$$

   to estimate θ, where $X_{[i]}$ is the i^{th} order statistic of the sample with $0 < r < \frac{n}{2} < s < n$ and r and s are integers. Find the coverage probability of this confidence interval, i.e., find

$$P(X_{[r]} \leq \theta \leq X_{[s]})\,.$$

10. $X_1, X_2, \ldots, X_n$ are i.i.d. Bernoulli random variables:

$$P(X_i = 1) = p = 1 - P(X_i = 0)\,, \quad 0 < p < 1\,.$$

   The parameter p is to be estimated with squared error loss.

   (a) Find the Bayes estimator of p if the prior distribution for p is beta, i.e.,

$$g(p) = \frac{p^{r-1}(1-p)^{s-1}}{B(r,s)}, \quad 0 < p < 1, \quad B(r,s) = \frac{\Gamma(r) \cdot \Gamma(s)}{\Gamma(r+s)}\,.$$

   (b) Show that the estimator $\hat{p} = \frac{\Sigma X_i + \frac{\sqrt{n}}{2}}{n+\sqrt{n}}$ is minimax.

11. Let $X \sim N(\theta, \sigma^2)$ with $-\infty < \theta < \infty$ and σ known. Consider estimating θ under the squared error loss. Show that the estimator $\delta_{a,b}(x) = ax + b$ is admissible if and only if one of the following holds:

    (a) $0 \leq a < 1$;

    (b) $a = 1$ and $b = 0$.

12. (a) Let $\mathcal{F}$ be a class of distributions and $\mathcal{F}_0 \subset \mathcal{F}$ be a subclass. Let $g(F)$, $F \in \mathcal{F}$ be a functional to be estimated under the loss $L\left(g(F), t\right)$. Suppose the estimator $\delta_0(X)$, $X \sim F$ is minimax for $g(F)$ when F is restricted to $\mathcal{F}_0$. Prove that a sufficient condition for $\delta_0(X)$ to be minimax for $F \in \mathcal{F}$ is

    $$\sup_{F \in \mathcal{F}_0} R(F, \delta_0) = \sup_{F \in \mathcal{F}} R(F, \delta_0)\,,$$

    where

    $$R(F, \delta) = \int L\left(g(F), \delta(x)\right) dF(x)\,.$$

    (b) Suppose $X_1, \ldots, X_n$ are i.i.d. $N(\theta, \sigma^2)$, $-\infty < \theta < \infty$ and $0 < \sigma^2 \leq C$. Consider estimating θ under the loss $L(\theta, t) = (\theta - t)^2$. Prove that $\bar{X} = \frac{1}{n} \sum_1^n X_i$ is minimax.

    Hint: Use (a).

    (c) Suppose $X_1, \ldots, X_n$ are i.i.d. $F \in \mathcal{F}$ where $\mathcal{F} = \{F\colon 0 < \sigma_F^2 \leq C\}$. Consider estimating the functional $\theta_F =$ mean of F. Under the squared error loss, prove that $\bar{X}$ is minimax.