

Stat 643 Final Exam

December 13, 1995

Prof. Vardeman

Below are 8 questions. Write answers to **exactly 5** of them. Each question will be worth 20 points. When you are finished, staple your answers together in numerical order.

1. Consider the two distributions P_0 and P_1 on $\mathcal{X} = (-1, 1)$ with densities with respect to Lebesgue measure on \mathcal{X}

$$f_0(x) = \frac{1}{2} \quad \text{and} \quad f_1(x) = |x| \quad .$$

a) Find a most powerful level $\alpha = .10$ test of $H_0: X \sim P_0$ vs $H_1: X \sim P_1$.

b) Consider the family of (mixture) distributions $\mathcal{P} = \{P_\theta\}$ with densities with respect to Lebesgue measure

$$f_\theta(x) = \theta f_1(x) + (1 - \theta) f_0(x) \quad \text{for } \theta \in [0, 1] \quad .$$

In this family, find a UMP level $\alpha = .10$ test of $H_0: \theta \leq .5$ vs $H_1: \theta > .5$ based on a single observation X . Argue carefully that the test you invent is UMP size .10 for this problem.

2. Suppose that X_1, X_2, \dots, X_n are iid with probability mass function given below for $\theta \in [0, 1]$.

x	0	1	2
$f_\theta(x)$	θ^2	$2\theta(1 - \theta)$	$(1 - \theta)^2$

a) Find an explicit form for the MLE of θ and argue that the MLE is consistent.

b) Note that if $\theta = 0$ or $\theta = 1$, the MLE of θ is not asymptotically normal. Point out a basic part of the hypotheses of Theorem 49 that is violated in the cases of $\theta = 0$ and $\theta = 1$.

c) Give explicitly two different 90% large sample confidence interval formulas for θ based on the asymptotic distribution of the MLE when $\theta \in (0, 1)$.

3. Suppose that for $i = 1, 2, \dots, n$, the random vectors $\mathbf{X}_i = (Y_i, Z_i)$ are iid with distribution P_p described as follows. Y_1 and Z_1 are independent, $Y_1 \sim \text{binomial}(m, p)$ and $Z_1 \sim \text{geometric}(p)$ (i.e. with pmf $g_p(z) = p(1 - p)^{z-1}$ for $z = 1, 2, 3, \dots$).

a) Give the C-R lower bound on the variance of an unbiased estimator of p based on $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$.

b) Describe as explicitly as possible how you would go about testing $H_0: p = .2$ vs $H_1: p \neq .2$ based on $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ with an approximate level of $\alpha = .1$.

c) Describe how you would make an approximately 90% confidence set for p based on $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ and the inversion of appropriate tests.

4. Miscellaneous Short Answers

a) Consider an estimation problem with $X \sim P_\theta$ with density f_θ with respect to Lebesgue measure on \mathcal{R}^1 . Suppose however, that measurement precision is such that in truth one can only observe X "to the nearest integer." What is an appropriate loglikelihood to use in this problem if n independent X_i 's are to be observed?

b) Suppose that in a "regular" estimation problem with bivariate parameter $\theta = (\theta_1, \theta_2)$ the 2×2 Fisher information matrix based on a single observation is $I_1(\theta)$. Suppose further that based on n independent observations one may compute the MLE, $\hat{\theta}_n = (\hat{\theta}_{1n}, \hat{\theta}_{2n})$. Describe how you would expect to compute a large sample (joint) confidence set for θ .

c) Argue that in a simple versus simple testing problem, a test minimax with "0-1 loss" must be an equalizer rule.

d) Continuity of risk functions is useful in showing the admissibility of Bayes decision rules. Power functions of tests are usually continuous and are related to risk functions for 0-1 loss. Do these facts provide the risk function continuity needed for simple admissibility arguments for Bayes tests under 0-1 loss (when, for example, $\Theta \subset \mathcal{R}^1$)? Explain.

5. Consider a Bayesian decision theoretic approach to the testing of a point null hypothesis versus a composite alternative. For the sake of concreteness suppose that

$X \sim N(\theta, 1)$ corresponding to $H_0: \theta = 0$ and $H_1: \theta \neq 0$, take $\Theta = \mathcal{R}^1$, $\mathcal{A} = \{0, 1\}$ and

$$L(\theta, a) = I[\theta = 0]I[a = 1] + I[\theta \neq 0]I[a = 0] \quad .$$

a) Show that if G is a prior that is absolutely continuous with respect to Lebesgue measure on Θ , $\phi(x) \equiv 1$ is Bayes with respect to G . What does this suggest about suitable priors for this kind of problem?

b) Suppose now that $G = \frac{1}{2}\Delta_0 + \frac{1}{2}\Lambda$, where Δ_0 places mass 1 at the point $\{0\}$ and no mass elsewhere, and Λ is $N(0, \sigma^2)$ measure. Find the form of a Bayes decision rule for this prior.

6. Suppose that X_1, X_2, \dots, X_n are independent normal random variables, each with variance 1. Suppose further that each $\mu_i (= EX_i)$ is either 0 or 1, so that we have for this problem a parameter space $\Theta = \{\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n) \mid \text{each } \mu_i \text{ is 0 or 1}\} = \{0, 1\}^n$. Based on $\mathbf{X} = (X_1, X_2, \dots, X_n)$ we will consider making decisions about all n means μ_i . To this end, let $\mathcal{A} = \Theta = \{0, 1\}^n$ and for $\mathbf{a} = (a_1, a_2, \dots, a_n)$ define the (overall) loss function

$$L(\boldsymbol{\mu}, \mathbf{a}) = \sum_{i=1}^n I[a_i \neq \mu_i] \quad .$$

Then, for each permutation of the first n positive integers $\boldsymbol{\pi} = (\pi(1), \pi(2), \dots, \pi(n))$, define a transformation $g_{\boldsymbol{\pi}}: \mathcal{R}^n \rightarrow \mathcal{R}^n$ by $g_{\boldsymbol{\pi}}(\mathbf{y}) = \mathbf{y}_{\boldsymbol{\pi}}$, for $\mathbf{y} = (y_1, y_2, \dots, y_n)$ and $\mathbf{y}_{\boldsymbol{\pi}} = (y_{\pi(1)}, y_{\pi(2)}, \dots, y_{\pi(n)})$ and let $\mathcal{G} = \{g_{\boldsymbol{\pi}} \mid \boldsymbol{\pi} \text{ is a permutation of the first } n \text{ positive integers}\}$.

- Show that this decision problem is invariant under the group of transformations \mathcal{G} .
- Characterize as clearly/completely as possible what it means for a nonrandomized decision rule $\delta: \mathcal{R}^n \rightarrow \mathcal{A} = \{0, 1\}^n$ to be equivariant in this problem.
- Equivariant decision rules in invariant decision problems have risk functions that are constant on orbits. What are the orbits for this decision problem?

7. Consider the one parameter exponential family of distributions on $\mathcal{X} = (0, 1)$ with densities with respect to Lebesgue measure

$$f_{\eta}(x) = \begin{cases} \frac{\eta}{\exp\eta - 1} \exp\eta x & \eta \neq 0 \\ 1 & \eta = 0 \end{cases} .$$

- Suppose that one wishes to test $H_0: \eta = 0$ vs $H_1: \eta \neq 0$ based on a single observation from this distribution. Find a UMPU size $\alpha = .2$ test of these hypotheses and argue that your test is indeed UMPU.
- For ϕ your test from a), consider the test $1 - \phi$. What optimality properties does this test possess? (For example, you might show that it is "best" of some specific size for some specific pair of hypotheses.)

8. Consider a random observable $X = (N, Y)$ with distribution $P_{\boldsymbol{\theta}}$ parametrized by $\boldsymbol{\theta} = (p, \mu) \in (0, 1) \times \mathcal{R}^1$. Suppose that N is geometric with parameter p (i.e. has pmf $g_p(n) = p(1 - p)^{n-1}$ for $n = 1, 2, 3, \dots$) and that the conditional distribution of Y given N is normal with mean μ and variance N^{-1} .

- For testing $H_0: \mu \leq \mu_0$ vs $H_1: \mu > \mu_0$ based on X , p might well be termed a nuisance parameter. Find a UMPU size $\alpha = .1$ test of these hypotheses and argue that your test is indeed UMPU of this size.
- Invert your tests $\phi(x)$ to produce a 90% confidence bound for μ based on X .