COMPREHENSIVE EXAM

Statistical Inference

August 26, 1986

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.

2. Students seeking a Ph.D. Level Pass should attempt problems 6 through 10 and one other problem.

3. Students seeking a pass at both levels must clearly designate those problems to be considered for the Master's Pass.

Part I: Students seeking a M.S. Level Pass Should First Attempt Problems 1-5.

1. [20 Points]

(a) Let X_i be a normal random variable with mean i and variance i^2, $i=1,2,3$. Assume that X_1, X_2, and X_3 are independent. Using only the three random variables X_1, X_2, and X_3:

(i) Give an example of a statistic that has a chi-square distribution with three degrees of freedom.

(ii) Give an example of a statistic that has an F distribution with two and one degrees of freedom.

(iii) Give an example of a statistic that has a t-distribution with two degrees of freedom.

(b) Let $X_1, \ldots, X_n$ be independent Bernoulli random variables with unknown parameter $p = P\{X_i=1\} = 1-P\{X_i=0\}>0$ for $i=1, \ldots, n$. Use the Rao-Blackwell theorem to find the UMVUE of $\theta = p^3$.

2. [20 Points]

Let $X_1, \ldots, X_n$ be a random sample from a random variable having the p.d.f.

$$f(x) = \begin{cases} e^{-(x-\theta)} & \text{for } x>\theta \\ 0 & \text{otherwise} \end{cases}.$$

(a) Find the maximum likelihood estimator (m.l.e.) of θ.

(b) Find the moment generating function (m.g.f.) of the m.l.e.

(c) Using your result in part (b) show that if $M(t)$ is the m.g.f. of the m.l.e. then

$$M'(t) = M(t)\{\theta + 1/(n-t)\}.$$

(d) Is the m.l.e. an unbiased estimator of θ? Is it a consistent estimator of θ? Defend your answers by using part (c).

(e) Find the method of moments estimator (m.o.m.) of θ. {Hint: set $n=1$ in $M(t)$ to find the m.g.f. of X.}

(f) Which estimator, the m.l.e. or the m.o.m., is the best estimator of θ? Defend your answer.

3. [20 Points]

Let $X_1, \ldots, X_n$ be a random sample from a Poisson population with mean θ.

(a) Construct the UMP test for testing H_0: $\theta \leq \theta_0$ vs. H_1: $\theta > \theta_0$.

(b) Describe how to find the critical value for a given Type I error size α.

(c) Assume n is large and then find an expression for the power function of the UMP test in terms of a well-known probability distribution.

(d) Using your result in part (c), find an expression for the minimal sample size needed to test H_0 versus H_1: $\theta = \theta_1 > \theta_0$ when the Type I and II errors are specified to be α and β, respectively.

4. [20 Points]

Let B and A refer to the measurement of a person's systolic blood pressure before and after the administration of a certain anti-hypertensive drug. Assume we have selected a random sample of n hypertensive subjects for our study and further assume that for person i the vector (B_i,A_i) has a bivariate normal distribution with mean vector (μ_B,μ_A) and covariance matrix

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

Here μ_B,μ_A, σ^2 and ρ represent unknown parameters. In our study the drug will be considered to be ineffective unless it reduces the average systolic blood pressure a minimum of 20 mm so that we are interested in testing H_0: $\Delta=20$ versus H_1: $\Delta>20$, where $\Delta = \mu_B-\mu_A$. Show that the likelihood ratio test of H_0 versus H_1 reduces to the familiar paired t statistic. {Hint: first consider the transformation $(B,A) \rightarrow (D,S)$ where $D = B-A$ and $S = B+A$}.

5. [20 Points]

Suppose that $\underset{\sim}{X} = (X_1, \ldots, X_n)$ is a random sample from a uniform distribution $U(0,\theta)$, $\theta>0$. Let θ have the prior density

$$\pi(\theta) = \begin{cases} \dfrac{\alpha}{\theta_0}\left(\dfrac{\theta_0}{\theta}\right)^{\alpha+1}, & \text{if } \theta>\theta_0, \\ 0, & \text{otherwise} \end{cases}$$

where $1<\alpha<\infty$, $0<\theta_0<\infty$.

(a) Show that the posterior density of θ is

$$g(\theta|\underset{\sim}{x}) = \begin{cases} \dfrac{\alpha+n}{\beta}\left(\dfrac{\beta}{\theta}\right)^{\alpha+n+1}, & \text{if } \theta>\beta \\ 0, & \text{otherwise,} \end{cases}$$

where $\beta = \max\{\theta_0, x_1, \ldots, x_n\}$.

(b) Find the Bayes estimate of θ if the loss function is $L(\theta,a) = (a-\theta)^2$.

(c) Find the Bayes estimate of θ if the loss function is $L(\theta,a) = (a-\theta)^2/\theta$.

Part II: Students seeking a Ph.D. pass should attempt Problems 6-10 inclusive.

6. [20 Points]

Let $X_1, \ldots, X_n$ be independent Bernoulli random variables with unknown parameter $p = P[X_i=1] = 1-P[X_i=0]>0$ for $i=1, \ldots, n$.

(i) Use the Rao-Blackwell Theorem to show that the minimum variance unbiased estimator of $\theta=p^2$ is

$$\hat{\theta}_1 = \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} X_i - 1)}{(n)(n-1)}.$$

(ii) Justify that $\hat{\theta}_1$ is minimum variance unbiased.

(iii) What is the maximum likelihood estimator of θ, say $\hat{\theta}_2$? Justify your answer.

(iv) Discuss the asymptotic relationship between $\hat{\theta}_1$ and $\hat{\theta}_2$.

(v) Derive the asymptotic distribution of $\hat{\theta}_1$, suitably normalized.

7. [20 Points]

Let $X \sim \mathrm{Bin}(n,p)$ $0 \leq p \leq 1$. Consider the problem of estimating p with squared error loss $L(p,\hat{p}) = (p-\hat{p})^2$.

(i) Obtain the Bayes estimator of p with respect to the prior Beta (a,b).

(ii) Show that the estimator $\delta^* = \dfrac{X+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}$ is minimax.

(iii) Prove that the m.l.e. $\dfrac{X}{n}$ is not minimax.

8. [20 Points]

Let $X \sim \mathrm{Bin}(n,p)$, $0<p<1$. Let $L(p,\hat{p}) = (p-\hat{p})^2/p(1-p)$ be the loss function. Show

(i) $\dfrac{X}{n}$ is minimax.

(ii) $\dfrac{X}{n}$ is admissible.

9. [20 Points]

Let T be a sufficient statistic for a family of distributions $\mathbb{F} = \{P_\theta, \theta \in \Theta\}$.

(a) Define

(i) T complete for $\mathbb{F}$

(ii) T boundedly complete for $\mathbb{F}$.

(b) Let

$$X = \{-1,0,1,2,...\},$$

$$P_\theta(x) = \begin{cases} \theta & , x = -1 \\ (1-\theta)^2\theta^n, & x = n \quad (n = 0,1,2,...) \end{cases},$$

$$\Theta = \{\theta \in \mathbb{R}: \ \theta > 0\},$$

and

$$T = X = \text{an observation with a distribution in } \mathbb{F}.$$

(a) Show that T is boundedly complete, but not complete. (Hint: Write $E_\theta g(T)$ as a power series.)

(b) Is the statistic $T^* = |X|$ sufficient for $\mathbb{F}$? (Justify your answer.)

10. [20 Points]

Let X_i be independent normal with $E(X_i) = \beta_0 + \beta_1 Z_i$, $\Sigma Z_i = 0$, $\Sigma Z_i^2 = 1$, i=1,2, and common variance $\sigma^2 = 1$.

(i) Show that $(\sqrt{2}\,\bar{X} - \sum_{i=1}^{2} X_i Z_i, \sqrt{2}\,\bar{X} + \sum_{i=1}^{2} X_i Z_i)$ is sufficient.

(ii) Derive the UMPU test of

$$H_0:\ \beta_1 = \sqrt{2}\beta_0$$
$$H_1:\ \beta_1 \neq \sqrt{2}\beta_0$$

(Please give all important steps in your argument.)

(iii) Is this test UMP invariant? Why?