Statistics and Biostatistics Ph.D. Theory Qualifier 2006

Total time: 3 hours.

Instructions: This is a closed book closed notes exam; however, you can use copies of the pages of your textbook containing the definitions of various pdfs and their moments. You need to answer any four problems. The problems are not arranged in any particular order, so I suggest that you read all six problems before deciding on the four that you wish to answer. If you answer more than four questions, only the first four will be graded. If you don't want to grade a part of your answer, please cross out that part of your answer book completely.

1. Consider the two pdfs given by:

$$f_0(x) = \tfrac{1}{2}, \ \ -1 < x < 1$$

and

$$f_1(x) = |x|\,, -1 < x < 1.$$

Suppose an observation X is made such that X is distributed as $f(x), \ -1 < x < 1$.

a. Find the most powerful level $\alpha = 0.1$ test for $H_o : f = f_0$ vs.$H_a : f = f_1$.

b. Consider the mixture distribution $f_\theta(x) = \theta f_1(x) + (1-\theta) f_0(x)$ for $0 \le \theta \le 1$.

In this family, find a UMP level $\alpha = 0.1$ test of $H_o : \theta \le 0.5$ vs $H_a : \theta > 0.5$ based on a single observation X. Argue carefully that the test you propose is UMP of size 0.1 for this problem.

2. Let X and Y be independent random variables with geometric distributions

$$f_{XY}(x, y|\theta_1, \theta_2) = (1-\theta_1)(1-\theta_2)\theta_1^x \theta_2^y, \qquad x = 0.1.2, ...; y = 0, 1, 2, ...$$

where $0 < \theta_1 < 1, 0 < \theta_2 < 1$.

Find a UMP unbiased test of size $\alpha = 0.2$ for testing

$H_o : \theta_1 \le \theta_2$ vs $H_a : \theta_1 > \theta_2$.

Be specific in terms of the critical value of the test.

3. a. Carefully and precisely define the terms:

i. complete and sufficient statistic.

ii. ancillary statistics.

b. Let $X_1, ..., X_n$ be an $i.i.d.$ sample from the pdf $f(x, \theta)$. Suppose $T(X) = T(X_1, ..., X_n)$ be a complete and sufficient statistics for θ, and $S(X)$ be an ancillary statistics. State the theorem that describes the statistical relationship between $T(X)$ and $S(X)$. Provide a proof of the theorem in the special case where both $T(X)$ and $S(X)$ are discrete random variables.

c. Let $X_1, ..., X_n$ $i.i.d.$Bernoulli(θ). Find the distribution of $W_k = \prod_{i=1}^{k} X_i$, for $k = 1, 2, .., n$. What is $E(W_k)$?

d. Find the UMVUE of θ^k, for $k = 1, 2...n$.

e. Find the UMVUE of the polynomial

$$a_0 + a_1\theta + a_2\theta^2 + ... + a_n\theta^n$$

where a_j are known constants. Justify your answer.

4. Let $X_1, ..., X_n$ be a random sample from he population with density

$$g(x; \theta, \mu_1, \mu_2) = \theta f(x, \mu_1) + (1-\theta) f(x, \mu_2)$$

where $f(x, \mu)$ is the density of $N(\mu, 1)$, and $0 < \theta < 1$.

Carefully derive a procedure to compute the MLE of (θ, μ_1, μ_2) by using EM-algorithm.

5. Let $X_1, ..., X_n$ be a random sample from a population with probability function

$$f(x, \theta) = \theta(1-\theta)^x,\ x = 0, 1, 2, 3....$$

where θ is an unknown parameter in $(0, 1)$, under the classical approach to inference. One can also use a Bayesian approach by assuming that θ has a prior distribution $\pi(\theta)$.

a. Compute the Cramer-Rao lower bound for an unbiased estimator of θ.

b. Find UMVUE for θ, if it exists.

c. Define the term "conjugate prior". What is the conjugate prior for θ for the family of the pdf $f(x, \theta)$?

d. If the loss function $L(\theta, a) = (\theta - a)^2$, find the Bayes estimator of θ with respect to the conjugate prior as defined in part (c).

e. If the loss function $L(\theta, a) = \frac{(\theta-a)^2}{\theta(1-\theta)}$, find the Bayes estimator under the conjugate prior.

6. Suppose the primary end point of a clinical trial is survival time T, and the distribution of T is given by:

$$f(t, \theta) = \frac{1}{\theta}\exp(-\frac{t}{\theta}),\ t > 0.$$

Suppose further that the primary objective of the study is formulated in testing

$$H_0 : \theta = 3 \quad \text{vs} \quad H_a : \theta = 1.$$

a. Suppose $T_1, ..., T_n$ be the observed survival times from a random sample of n patients. Derive the most powerful test for testing H_0 against H_a.

b. Determine the sample size for the clinical trial to obtain a level 5% test with power 80% using

i. the exact method.

ii. the normal approximation.