NAME:

18.466 final exam, Wednesday, May 21, 2003, 9 A.M.-noon

Closed book exam. No books or notes may be consulted during this exam.

There are 13 questions on the exam. Answer any TEN of the 13 for full credit. Please indicate which three you omit.

Explanations should be given only where requested, or as time permits.

1. Let $X = \{0,1,2\}$, $P(0) = 0.8$, $P(1) = 0.05$, $P(2) = 0.15$, $Q(0) = 0.008$, $Q(1) = 0.002$, $Q(2) = 0.99$.

(a) What is the most powerful non-randomized test of P vs. Q with size ≤ 0.05?

(b) What is the power of the test in (a)?

(c) Find the likelihood ratio $R_{Q/P}(x)$ at each x.

(d) What are the admissible non-randomized tests of P vs. Q?

(e) What is the power of each test in part (d)?

2. (a) Define sequential probability ratio tests (SPRTs).

(b) State the main optimality theorem about SPRTs.

3. (a) Define exponential families.

(b) Define the order of an exponential family.

(c) For n i.i.d. observations from an exponential family as defined in part (a), give a sufficient statistic whose dimension equals the order and so doesn't depend on n.

(d) For the family of gamma distributions with densities

$$f((\alpha, \lambda), x) = \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x} / \Gamma(\alpha)$$

for $x > 0$ and 0 for $x \leq 0$, where $\alpha > 0$ and $\lambda > 0$, show that the family is exponential and find its order.

(e) What is the statistic as in part (c) for n i.i.d. observations?

4. Let a parameter θ range over an interval $a < \theta < b$. Let T be an estimator of a function $g(\theta)$ with a bias $b(\theta)$, so that $E_\theta T = g(\theta) + b(\theta)$ for $a < \theta < b$. Here T, g, and b are real-valued. Suppose sufficient conditions for the information inequality hold for T as an unbiased estimator of $(g + b)(\theta)$. Give a lower bound for the mean-squared error of T as an estimator of $g(\theta)$, $E_\theta((T - g(\theta))^2)$.

5. Let x have a $N(\mu, I)$ distribution in $\mathbb{R}^3$ where I is the 3×3 identity matrix and the unknown μ can be any vector in $\mathbb{R}^3$.
(a) What is $E_\mu|x - \mu|^2$?

(b) What is a function $J(x)$ such that $E_\mu|J(x) - \mu|^2 < E_\mu|x - \mu|^2$ for all μ in $\mathbb{R}^3$?

6. Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$. Let $V := \sum_{j=1}^{n}(X_j - \overline{X})^2$.
(a) For what constants $c_1(n)$ depending on n is $c_1(n)V$ an unbiased estimator of σ^2?

(b) For what constants $c_2(n)$ depending on n is the mean-square error $E_{\sigma^2}[(c_1(n)V - \sigma^2)^2]$ minimized for all $\sigma > 0$? *Hints*: V/σ^2 has a χ^2_{n-1} distribution. A χ^2_d distribution is $\Gamma(\alpha, \lambda)$ as in problem 3(d) with $\alpha = d/2$ and $\lambda = 1/2$. If Y has a $\Gamma(\alpha, \lambda)$ distribution then $EY = \alpha/\lambda$ and Y has variance α/λ^2.

7. Let Θ be a parameter space and $(X, \mathcal{B})$ a sample space. Given a function $h(\theta, x)$ on $\Theta \times X$,

(a) Define M-estimator (of ρ type) based on h.

(b) Define what is a sequence $T_n = T_n(X_1, \ldots, X_n)$ of approximate M-estimators based on h.

(c) Define what it means for h to be *adjusted*.

(d) Define what it means for h to be *adjustable*.

(e) If observations have a distribution P, define what it means for a $\theta_0 \in \Theta$ to be pseudo-true.

(f) What is the relationship between a sequence T_n of approximate M-estimators and a pseudo-true θ_0 under sufficient regularity conditions?

8. (a) Define the Kullback-Leibler information $I(P, Q)$ for two laws P and Q on a sample space.

(b) Under what conditions on laws P and Q is: $I(P, Q) < 0$?

$I(P, Q) = 0$?

$I(P, Q) > 0$?

(c) Let $h(\theta, x) = -\log f(\theta, x)$ for the likelihood function f of a family $\{P_\theta,\ \theta \in \Theta\}$, and let the distribution P of the data be P_{θ_1} for some $\theta_1 \in \Theta$. What is an adjustment function for h in this case?

(d) If the conditions in (c) hold, is the true θ_1 always, sometimes, or never equal to a pseudo-true value θ_0? Explain.

9. (a) Define the finite-sample breakdown point of a statistic $T_n = T_n(X_1, \ldots, X_n)$ having values in a parameter space Θ at a sample $(X_1, \ldots, X_n)$.

(b) For real observations $(X_1, \ldots, X_n)$ and their order statistics $X_{(1)} \leq \cdots \leq X_{(n)}$, what is the breakdown point of $X_{(j)}$?

(c) Define equivariance for location of a real-valued statistic $T_n = T_n((X_1, \ldots, X_n))$ for real $(X_1, \ldots, X_n)$.

(d) What can be said about the breakdown points of statistics equivariant for location?

(e) What is an example of a sequence of statistics T_n, each equivariant for location, having largest possible breakdown point for each n?

10. Suppose given a family $\{P_\theta,\ \theta \in \Theta\}$ where Θ is an open subset of a Euclidean space $\mathbb{R}^d$, with a likelihood function $f(\theta, x) > 0$ for all θ and x.
(a) Define the Fisher information matrix $I(\theta)$ of the family.

(b) If $f(\theta, x)$ is C^2 as a function of θ and other suitable conditions hold, give an alternate form of $I(\theta)$.

(c) Let $T_n = T_n((X_1, \ldots, X_n)\ )$ be a sequence of estimators of θ. When are T_n said to be efficient?

(d) What class of estimators were shown in the course to be efficient under some conditions?

11. In the previous problem, suppose $d = 1$ and we want to estimate a function $g(\theta)$.
(a) Under some regularity conditions, what is an asymptotic lower bound for the mean-square error $E_\theta[n(T_n - g(\theta))^2]$?

(b) Does the asymptotic lower bound hold for all θ or if not, what can be said about those θ for which it doesn't?

(c) Give an example where the regularity conditions needed for (a) fail and $g(\theta)$ can be estimated with a mean-square error that goes to 0 faster as $n \to \infty$.

12. (a) For what kinds of hypotheses and alternatives does the Wilks test provide a test?

(b) What is the Wilks test statistic?

(c) What is the asymptotic distribution of the statistic, under some regularity conditions, and if which hypothesis is true?

13. A beta(a, b) distribution has a density $x^{a-1}(1-x)^{b-1}/B(a, b)$ for $0 < x < 1$ and 0 elsewhere, where $a > 0$, $b > 0$, and $B(a, b)$ is the beta function $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$. The beta$(a, b)$ distribution has mean $a/(a+b)$ and variance $ab/[(a+b)^2(a+b+1)]$. Let $X_1, X_2, \ldots$, be i.i.d. with a Bernoulli distribution, $P(X_j = 1) = p = 1 - P(X_j = 0)$ where $0 \leq p \leq 1$. Let $S_n = X_1 + \cdots + X_n$. Let p have a beta(a, b) prior for some $a > 0$ and $b > 0$.
(a) After n observations, what is the likelihood function $f(p, ((X_1, \ldots, X_n)))$?

(b) What is the posterior distribution of p?

(c) What does it mean for posteriors to be consistent?

(d) Are the posteriors consistent in this case: for any p, $0 \leq p \leq 1$, or only if $0 < p < 1$, possibly depending on a and b?