Statistics 582, Problem Set 9, Corrected

Wellner; 3/2/2011

Reading: Chapter 6, sections 6.1 and 6.2 (through page 19).
Due: Wednesday, March 8, 2011.
Reminder: Final Exam, Monday, March 14, 8:30 - 10:30 (MGH 287)

1. Consider the Locally Most Powerful test ϕ for testing $H : \theta \le 0 \equiv \theta_0$ versus $K : \theta > 0 = \theta_0$ in Example 6.1.5.
   (a) Suggest two different approximations to the power of this test, one for local alternatives (of the form $\theta_n = t/\sqrt{n}$ with $t > 0$), and the other for fixed alternatives, $\theta > 0$.
   (b) What is the behavior of each of these two approximations for large values of θ? Which of them shows that the power function decreases to 0 as $\theta \to \infty$? Why?
   (c) Find a test ϕ of H versus K which has monotone increasing power function $\beta_\phi(\theta)$.

2. Let X and Y be random variables with joint density

$$p_{X,Y}(x, y) = \lambda\mu \exp(-\lambda x - \mu y) 1_{(0,\infty)}(x) 1_{(0,\infty)}(y).$$

   (a) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda \le \mu + 1$ versus $H_1 : \lambda > \mu + 1$.
   (b) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda = \mu$ versus $H_1 : \lambda \ne \mu$.
   (c) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda \ge 2\mu$ versus $H_1 : \lambda < 2\mu$.
   (d) What happens when $X_1, \ldots, X_m$ are i.i.d. Exponential(λ) and $Y_1, \ldots, Y_n$ are i.i.d. Exponential(μ)?

3. Lehmann and Romano, TSH, problem 4.3, page 139.

4. (From Wasserman, *All of Statistics*, page 171.) In 1961, 10 essays appeared in the *New Orleans Daily Crescent*. They were signed “Quintus Curtius Snodgrass” and some people suspected they were actually written by Mark Twain. To investigate this, we will consider the proportion of three letter words founds in an author’s work. From eight Twain essays we have

   .225, .262, .217, .240, .230, .229, .235, .217

From 10 Snodgrass essays we have:

.209, .205, .196, .210, .202, .207, .224, .223, .220, .201

(a) Perform a Wald test for equality of the means. Give a $p-$value and a 95% confidence interval for the difference of means. What conclusion do you reach?
(b) Now use a permutation test to avoid the use of large - sample methods. What is your conclusion?

5. For observations $\underline{X} = (X_1, \ldots, X_n)$, let $X_{(1)} \le \ldots \le X_{(n)}$ denote the *order statistics* of the X_i's ($X_{(i)} \equiv \mathbb{F}_n^{-1}(i/n)$, $i = 1, \ldots, n$) and let $\underline{R} = (R_1, \ldots, R_n)$ denote the *ranks*; defined by $X_i = X_{(R_i)}$, $i = 1, \ldots, n$ (if $X_i = X_j$ for some $i < j$, define the ranks by $R_i < R_j$ and $X_i = X_{(R_i)}$).
(a) Suppose that $X_1, \ldots, X_n$ are i.i.d. $F \in \mathcal{F}_{ac}$ (the absolutely continuous df's F on R) with density f. Show that the order statistics $\underline{X}_{(\cdot)} \equiv (X_{(1)}, \ldots, X_{(n)})$ are independent of the ranks $\underline{R}$ and that the order statistics have joint density $\bar{p}$ given by

$$\bar{p}(\underline{x}_{(\cdot)}) = n! \prod_{i=1}^{n} f(x_{(i)}) \,, \qquad -\infty < x_{(1)} < \ldots < x_{(n)} < \infty$$

while

$$P(\underline{R} = \underline{r}) = \frac{1}{n!} \,, \qquad \underline{r} \in \Pi \equiv \{ \text{ all permutations of } \{1, \ldots, n\}\} \,.$$

(b) Show that if the density f of the X_i's is log-concave, then the joint density $\bar{p}$ of the order statistics $\underline{X}_{(\cdot)}$ is log-concave; i.e. show that if $f((x+y)/2)^2 \ge f(x)f(y)$ for all $x, y \in \mathbb{R}$, then $\bar{p}((\underline{x}+\underline{y})/2)^2 \ge \bar{p}(\underline{x})\bar{p}(\underline{y})$ for all $\underline{x}, \underline{y} \in \mathcal{O}_n \equiv \{\underline{x} \in \mathbb{R}^n : x_1 \le x_2 \le \cdots \le x_n\}$.
(c) Show that (a) continues to hold for any joint distribution p of the $\underline{X}$ which is symmetric with respect to permutation of its coordinates: $p(\pi\underline{x}) = p(\underline{x})$ for all $\underline{x}$ and $\pi \in \Pi$ where $\pi x \equiv (x_{\pi(1)}, \ldots, x_{\pi(n)})$.
(d) If the joint distribution p of $\underline{X}$ is general (not permutation symmetric), show that the joint density $\bar{p}$ of the order statistics is given by

$$\bar{p}(\underline{x}_{(\cdot)}) = \sum_{\pi \in \Pi} p(\pi \underline{x}_{(\cdot)}) \ ,$$

and

$$P(\underline{R} = \underline{r} | \underline{X}_{(\cdot)} = \underline{x}_{(\cdot)}) = \frac{p(r\underline{x}_{(\cdot)})}{\bar{p}(\underline{x}_{(\cdot)})} \ .$$

6. **Optional bonus problem 1:** Let $X_1, \dots, X_n$ be a sample of size n from the uniform distribution $U(0, \theta)$. Sufficiency reduces the problem to $T = \max X_i$.
(a) Find the class of all Neyman-Pearson best tests of $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, where $\theta_1 > \theta_0$.
(b) Find the subclass of the tests that are independent of θ_1. These are UMP tests of H_0 versus $H_1' : \theta > \theta_0$.
(c) Show that the test $\phi(t) = 1\{t > \theta_0\} + \alpha 1\{t \le \theta_0\}$ is UMP of size α for testing $H_0' : \theta \le \theta_0$ versus $H_1' : \theta > \theta_0$ but that ϕ is not admissible.
(d) Show that $\phi(t) = 1\{[t > \theta_0] \cup [t \le b]\}$ where $b = \theta_0 \alpha^{1/n}$ is a UMP test of size α for testing $H_0 : \theta = \theta_0$ versus $\theta \ne \theta_0$.

7. **Optional bonus problem 2:** let X and Y be independent random variables with geometric distributions

$$p_{X,Y}(x, y|\theta_1, \theta_2) = (1 - \theta_1)(1 - \theta_2)\theta_1^x \theta_2^y, \qquad x, y \in \{0, 1, \dots\}.$$

where $0 < \theta_j < 1$, $j = 1, 2$. Find a UMP unbiased test of size $\alpha = .20$ for testing
(a) $H_0 : \theta_1 \le \theta_2$ versus $H_1 : \theta_1 > \theta_2$.
(b) $H_0 : \theta_1 = \theta_2$ versus $H_1 : \theta_1 \ne \theta_2$.
(c) For what functions $\varphi(\theta_1, \theta_2)$ do our methods guarantee existence of a UMP unbiased test of $H_0 : \varphi(\theta_1, \theta_2) = 0$ versus $H_1 : \varphi(\theta_1, \theta_2) \ne 0$?

8. **Optional bonus problem 3:** (From the material on consistency of Neyman Pearson tests, section 6.1; and Donoho and Jin, *Ann. Statist.* **32** (2004), 962-994)
Let $p_\mu(x) \equiv \phi(x - \mu)$ denote the density of $X \sim N(\mu, 1)$, and let P_μ denote the corresponding measure on $\mathbb{R}$.
(a) Consider testing $H : P = P_0$ versus $K : Q = P_\mu$ with $\mu > 0$ fixed. Compute $\rho(P, Q) = \rho(P_0, P_\mu) = \int \sqrt{p_0(x) p_\mu(x)} dx$ explicitly as a function of μ.
(b) Compare the power functions of the following two tests of H versus K when $\mu = 1$:
(i) The Neyman - Pearson test with $\alpha = .05$; (ii) The Neyman - Pearson type test with $k = k_n = 1$ (in the notation of Theorem 6.1.4, page 8, Chapter 6).
(c) Now suppose $\mu = \mu_n \equiv t/\sqrt{n}$ with $t > 0$ and consider testing $H : P = P_0$ versus $K_n : P = P_{\mu_n}$.
(i) Find the limit of $\rho(P_0, P_{\mu_n})^n$ as a function of t.
(ii) Compare the limiting power function of the two tests in (b).
(d) Now suppose that $Q = Q_n$ is given by the mixture

$$q(x) = q(x; \mu, \epsilon) = (1 - \epsilon) p_0(x) + \epsilon p_\mu(x)$$

where $\epsilon = \epsilon_n = n^{-\beta}$ with $1/2 < \beta < 1$ and $\mu = \mu_n = \sqrt{2r \log n}$ with $r > \rho^*(\beta)$

and where the function

$$\rho^*(\beta) = \begin{cases} \beta - 1/2, & 1/2 \le \beta < 3/4, \\ (1 - \sqrt{1-\beta})^2, & 3/4 \le \beta < 1. \end{cases}$$

Show that k_n can be chosen so that $E_{P_0^n}\phi_n(\underline{X}) \to 0$ and $E_{Q_n}^n(1 - \phi_n(\underline{X}_n)) \to 0$.