Statistics 582, Problem Set 9

Wellner; 3/1/2006

Reading: Chapter 6, sections 6.1 and 6.2 (through page 19)
Due: Wednesday, March 8, 2006.

1. From problem set 1, 583, 1997 let X and Y be independent random variables with geometric distributions

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

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 $H_1 : \varphi(\theta_1, \theta_2) \ne 0$?

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. (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?

4. For observations $\underline{X} = (X_1, \ldots, X_n)$, let $X_{(1)} \leq \ldots \leq X_{(n)}$ 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 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 $\pi x \equiv (x_{\pi(1)}, \ldots, x_{\pi(n)})$.
C. 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
$$P(\underline{R} = \underline{r} | \underline{X}_{(\cdot)} = \underline{x}_{(\cdot)}) = \frac{p(r\underline{x}_{(\cdot)})}{\bar{p}(\underline{x}_{(\cdot)})} .$$

5. **Optional bonus problem:** Consider the Locally Most Powerful test ϕ for testing $H : \theta = 0$ versus $K : \ \theta > \theta_0$ in the Cauchy location Example 1.5, pages 6 and 7, chapter 6.
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 does one approximation succeed in this respect while the other fails?