Statistics 583, Problem Set 5

Wellner; 5/27/2016

Reading: Chapter 6, pages 1 - 50.
Van der Vaart, Chapters 13-15, pages 173 - 226.
Due: Wednesday, May 4, 2016
Reminder: Midterm exam, Monday, May 2.

1. A random variable X takes on the values $1,\ 2,\ 3,\ 4$ with probability distribution $p_0(x)$ or $p_1(x)$ as follows:

x	1	2	3	4
$p_0(x)$	.54	.08	.12	.26
$p_1(x)$	.22	.16	.36	.26

(a) Find a most powerful test of size $\alpha = .15$ for testing p_0 versus p_1 and determine its power.
(b) Find a test ϕ which minimizes the Bayes risk with $0-1$ loss and prior distribution $(\lambda, 1-\lambda) = (2/5, 3/5)$; i.e. a test ϕ which minimizes $\lambda a + (1-\lambda)b$ where $a = E_0\phi$ and $b = E_1(1-\phi)$.

2. Continuation of problem 1. For P_0 and P_1 as given in problem 1, compute $d_{TV}(P_0, P_1)$, $H(P_0, P_1)$, and the affinity $\rho(P_0, P_1) = \int \sqrt{p_0 p_1} d\mu$. For the product laws P_{0n} and P_{1n} (corresponding to observation of $X_1, \ldots, X_n$ i.i.d. P_0 or P_1 respectively) compute $\rho(P_{0n}, P_{1n})$ and $H(P_{0n}, P_{1n})$ for $n = 20,\ 50,\ 100$. What does this imply about the test, ϕ_n say, based on $X_1, \ldots, X_n$ which minimizes the sum of risks?

3. **Optional bonus problem 1:** 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 density 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 density 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

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

Hint: The easiest way might be to solve (d) first, then (c) followed by (a) and (b).

4. **Optional bonus problem 2:** Let $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ where p_θ is a family of densities with respect to a fixed dominating measure μ defined on a sample space $\mathcal{X}$ and $\Theta \subset \mathbb{R}$.
(a) Suppose that the densities $p_\theta(x) \equiv p(x, \theta)$ have a second mixed partial derivative and that

$$\frac{\partial^2}{\partial x \partial \theta} \log p(x, \theta) \ge 0$$

for all $x \in \mathbb{R}$ and $\theta \in \Theta$. Show that the inequality in the last display implies that $\mathcal{P}$ has monotone likelihood ratio. [Hint: use the fundamental theorem of calculus twice.]
(b) Show that the condition in (a) is equivalent to

$$p(x,\theta)\frac{\partial^2}{\partial x \partial \theta} p(x,\theta) \ge \frac{\partial}{\partial \theta} p(x,\theta) \frac{\partial}{\partial x} p(x,\theta) \quad \text{for all} \quad \theta \in \Theta, \; x \in \mathcal{X}.$$