Stat 643 Spring 2010
Assignment 08 - 10

Information Inequalities for Estimation

76. First, $\mathcal{P}$ is FI regular at θ_0. Let g_η be the R-N derivative of Q_η with respect to μ. Then $g_\eta = f_\theta$, $a.s.(\mu)$. Further,
$$\frac{dg_\eta}{d\eta} = \frac{df_\theta}{d\theta}\frac{1}{dh(\theta)/d\theta} = \frac{df_\theta}{d\theta}\frac{1}{h'(\theta)}. \tag{1}$$
The problem then is solved by observing the numerator and denominator of C-R lower bound ($\{1/h'(\theta)\}^2$ cancels out. Also note that $h'(\theta)$ is not zero because of the non-vanishing assumption).

77. Notice that in this case, for any $P_{\theta'}, P_\theta$ from the family $\{U(0,\theta) \mid \theta > 0\}$, $P_{\theta'} \ll P_\theta$ if and only if $\theta' \leq \theta$. Thus the Chapman-Robbins lower bound is
$$\sup_{\theta'\leq\theta} \frac{(\theta-\theta')^2}{E_\theta\left[1-\frac{\theta}{\theta'}I\{0<x<\theta'\}\right]^2} = \sup_{\theta'\leq\theta}(\theta-\theta')\theta' = \left\{\frac{(\theta-\theta')+\theta'}{2}\right\}^2 = \theta^2/4, \tag{2}$$
with equality if and only if $\theta' = \theta/2$. One unbiased estimator can be $2X$ and $Var_\theta(2X) = \theta^2/3$ which is larger than the Champan-Robbins lower bound.

78. The log-likelihood function is
$$L_n(\mathbf{p}) = \sum_{i=1}^{k} N_i \log p_i + \log\left(\frac{n!}{\prod_{i=1}^k N_k!}\right) = \sum_{i=1}^{k-1} N_1 \log p_i + N_k \log(1-p_1-\cdots p_{k-1}) + \log\left(\frac{n!}{\prod_{i=1}^k N_k!}\right). \tag{3}$$
Then
$$I(\mathbf{p}) = E_{\mathbf{p}}\left\{\frac{\partial L_n(\mathbf{p})}{\partial \mathbf{p}} \cdot \frac{\partial' L_n(\mathbf{p})}{\partial \mathbf{p}}\right\} = n\{diag(1/p_1,\ldots,1/p_{k-1}) + 1/p_k\mathbf{11'}\}, \tag{4}$$
where $\mathbf{1}$ is the vector full of 1's, and $I(\mathbf{p})$ is a $(k-1)\times(k-1)$ matrix.

The result for the second problem is straightforward by verifying conditions in the C-R lower bound theorem.

Testing

Before proceeding, review the relationship between randomized decision rule ϕ_x and randomized test ϕ. Especially, in the hypothesis test $H_0 : \theta \in \Theta_0\, vs\, \theta \in \Theta_1$ framework, we are considering the $0-1$ loss function
$$L(\theta, a) = I\{\theta \in \Theta_0\}I\{a = 1\} + I\{\theta \in \Theta_1\}I\{a = 0\}. \tag{5}$$
The action space $\mathcal{A} = \{0, 1\}$. A randomized test is actually $\phi(x) = \phi_x(\{1\})$ (recall that ϕ_x is a measure on $\mathcal{A}$ given a specific x). That's why we say, when the randomized test $\phi(x) = \pi$, one rejects H_0 (take action 1) with probability π. Consequently, when $\theta \in \Theta_0$, $R(\theta, \phi) = E_\theta I\{a = 1\} = \int_\mathcal{X}\int_\mathcal{A} I\{a = 1\}d\phi_x(a)dP_\theta(x) = E_\theta\{\phi_x(\{1\})\} = E_\theta(\phi(X))$. We do use the same notation for both randomized decision rule and randomized test, but they should not cause any confusion by carefully looking into them.

79. Just keep x axis the same and do a symmetric rotation of y with respect to the line $y = 1/2$. Because $\mathcal{V} = \{(x, 1-y) \mid (x,y) \in \mathcal{S}\}$.

80. (a) The answer is true because the Bayes risk would be 0 in those two cases.

(b) The Bayes test against G would be the Bayes rule with respect to G. Thinking about minimizing the condition expectation of the 0-1 loss function. One can see that the Bayes rule respect to G would be a non-randomized decision rule

$$\delta(x) = I\{P_{\theta|X=x}(\theta \in \Theta_0) \leq P_{\theta|X=x}(\theta \in \Theta_1)\}. \tag{6}$$

In terms of g_0, g_1, it would be

$$\delta(x) = I\{g_1(x)/g_0(x) \geq G(\Theta_0)/G(\Theta_1)\} \tag{7}$$

(c) We would just apply the result from above. When G is a normal distribution, $G(\{0\}) = 0$, $\phi(x) = 1$ would then be a Bayes test. When $G = (N + \Delta)/2$, $G(\Theta_0) = 1/2 = G(\Theta_1)$. After some disaster-like algebra, we can show that the Bayes test is

$$\phi(x) = I\left\{x^2 \geq \left(1 + \frac{1}{\sigma^2}\right) \log\left(1 + \frac{1}{\sigma^2}\right)\right\} \tag{8}$$

81. (a) By Neyman-Pearson Lemma, the test $\phi(x) = I\{x > 0.8\}$ is the most powerful test at level $\alpha = 0.2$.

(b) Consider $\mathcal{V} = \{(\beta_\phi(0), \beta_\phi(1)) \mid \phi : \mathcal{X} \to [0,1] \text{ is a randomized test}\}$. Notice that $\beta_\phi(0) = \int_0^1 \phi(x)dx$ and $\beta_\phi(1) = \int_0^1 3x^2\phi(x)dx$. Now the question is, what's the shape of $\mathcal{V}$, can $\mathcal{V}$ just be $[0,1]^2$? One can see that $(0,0), (1,1) \in \mathcal{V}$ and $\mathcal{V}$ is symmetric about $(1/2, 1/2)$. Let $\int_0^1 \phi(x)dx = a$, $\int_0^1 3x^2\phi(x)dx = b$. Then $b \leq 3\int_0^1 \phi(x)dx = 3a$. As $\mathcal{V}$ is symmetric about $(1/2, 1/2)$, we have $3a - b \leq 2$. Honestly, I don't have a clear map of $\mathcal{V}$ in my head.

82. Since $\alpha \in (0, 0.5), c \in (\alpha/(2-2\alpha), \alpha)$, $2c/\alpha > 1, (1-c)/(1-\alpha) > 1$, then the likelihood ratio is

$$\Lambda(x) = \frac{\sup_{\theta\in\Theta} f_\theta(x)}{\sup_{\theta\in\Theta_0} f_\theta(x)} = \max\left(1, \frac{\sup_{\theta\in\Theta_1} f_\theta(x)}{\sup_{\theta\in\Theta_0} f_\theta(x)}\right) = \frac{2c}{\alpha} I\{|X| = 2\} + \frac{1-c}{1-\alpha} I\{|X| < 2\}. \tag{9}$$

Further, $c/(1-c) > c > \alpha/(2-2\alpha) \Rightarrow 2c/\alpha > (1-c)/(1-\alpha) > 1$. The size α likelihood ratio test could be

$$\phi_{LRT}(x) = I\{\Lambda(x) > (2c/\alpha + (1-c)/(1-\alpha))/2\} = I\{|X| = 2\}. \tag{10}$$

Consider the power function of the two tests of size α, ϕ_{LRT} and $\phi(x) = I\{x = 0\}$, for any $\theta \neq -1$, as $\beta_{\phi_{LRT}}(\theta) = \theta c + (1-\theta)c = c < \alpha = \beta_\phi(\theta)$, which indicates that ϕ is strictly more powerful than LRT whatever be θ.