# Statistics 581, Final Exam

Wellner; 12/14/2009

**This exam is to be taken without the use of any books or notes.**

1. (40) points) **Define** *each* of the following terms. In each case, briefly provide an appropriate context for your definition.
   (a) The information matrix for $\theta$ in a regular parametric model $\mathcal{P} = \{P_\theta : \ \theta \in \Theta \subset R^d\}$.
   (b) The efficient score function for a parameter $\theta_1$ when $\theta = (\theta_1, \theta_2)$.
   (c) The efficient influence function $\tilde{l}_1$ for a parameter $\theta_1$ when $\theta = (\theta_1, \theta_2)$.
   (d) The efficient influence function for $\tilde{l}_\nu$ for a differentiable parameter $q(\theta) = \nu(P_\theta)$ in a regular parametric model $\mathcal{P}$.
   (e) An asymptotically linear estimator of a parameter $\nu(P)$ with influence function $\psi$.

2. (32 points) **State** *four* of the following five results, providing the appropriate (brief) context for your statement:
   (a) The (elementary) Skorokhod theorem.
   (b) The Cramér - Wold device.
   (c) A result about the finite-dimensional limiting distributions of the sample quantile process $\{\sqrt{n}(\mathbb{F}_n^{-1}(t) - F^{-1}(t)) : \ 0 < t < 1\}$ specifying the assumption(s) carefully.
   (d) Le Cam's third lemma.
   (e) The Glivenko-Cantelli theorem.

**Do either problem 3 or problem 4:**

3. (50 points). (A two-sample model.) Suppose that $Z \sim$ Bernoulli($\eta$), and conditional on $Z$, a random variable $X$ has conditional distributions given $Z$ described by $(X|Z) \sim$ Exponential($\mu Z + \nu(1-Z)$) where $\mu, \nu > 0$. Thus

$$P(X > x|Z = 1) = \exp(-\mu x), \quad x \ge 0,$$
$$P(X > x|Z = 0) = \exp(-\nu x), \quad x \ge 0.$$

(a) Let $\theta = (\mu, \nu, \eta) \equiv (\theta_1, \theta_2, \theta_3) \in (0,\infty) \times (0,\infty) \times (0,1)$, and write $(X, Z) \sim P_\theta$. Show that the joint density $p_\theta$ of $(X, Z)$ (with respect to the product of Lebesgue measure $\lambda$ on $(0,\infty)$ and counting measure $\#$ on $\{0,1\}$ is given by

$$p_\theta(x, z) = (\eta\mu e^{-\mu x})^z((1-\eta)\nu e^{-\nu x})^{1-z} 1_{(0,\infty)}(x) 1_{\{0,1\}}(z).$$

(b) Find the score(s) for $\theta = (\mu, \nu, \eta)$ and the information matrix for $\theta$ (when $n = 1$).
(c) If $(X_1, Z_1), \ldots, (X_n, Z_n)$ are i.i.d. as $(X, Z) \sim P_{\theta_0}$, what is the Cramér - Rao bond for estimation of $\nu(P_\theta) \equiv \theta_1 - \theta_2 = \mu - \nu$?
(d) For a sample of $(X, Z)$ pairs as in (c), find the maximum likelihood estimator

$\widehat{\theta}_n$ of $\theta$.

(e) Find the limiting distribution of $\sqrt{n}(q(\widehat{\theta}_n) - q(\theta_0))$ when $\theta = \theta_0$ is true.

(f) Propose a test statistic for testing $H : \mu = \nu$ versus $K : \mu \neq \nu$ based on the result of (e). What is the limiting distribution of your test statistic under $H$? What is the limiting distribution of your test statistic under local alternatives of the form $\mu = \nu + tn^{-1/2}$?

4. (48 points).

Suppose that $X_1, \ldots, X_n$ are independent and identically distributed real-valued random variables with distribution function $F$ and density $f$.

(a) Consider the sample median $\mathbb{F}_n^{-1}(1/2)$ and the sample mean $\overline{X}_n$. Give conditions under which the sample median $\mathbb{F}_n^{-1}(1/2)$ is an asymptotically linear estimator of the population median $F^{-1}(1/2)$. Identify the influence function $\psi(x)$ and the limiting distribution of $\sqrt{n}(\mathbb{F}_n^{-1}(1/2) - F^{-1}(1/2))$.

(b) Give conditions under which the sample median $\mathbb{F}_n^{-1}(1/2)$ and the sample mean $\overline{X}_n$ have a joint limiting distribution; i.e. conditions which imply that the random vector

$$\begin{pmatrix} \sqrt{n}(\mathbb{F}_n^{-1}(1/2) - F^{-1}(1/2)) \\ \sqrt{n}(\overline{X}_n - \mu) \end{pmatrix}$$

converges in distribution where $\mu = \mu_F = E_F(X_1)$. Find the limiting distribution explicitly.

(c) A simple test for asymmetry of a distribution function is based on the difference of the mean and median: $\gamma(F) \equiv \mu_F - F^{-1}(1/2)$. Note that $\gamma_F = 0$ if $F$ is symmetric about some point, while $\gamma(F)$ is positive for $F$ skewed to the right, and negative for $F$ skewed to the left. Use the results of (b) to find the limiting distribution of

$$\sqrt{n}(\gamma(\mathbb{F}_n) - \gamma(F)) = \sqrt{n}(\overline{X}_n - \mathbb{F}_n^{-1}(1/2) - (\mu_F - F^{-1}(1/2))).$$

Compute the limiting variance in terms of expectations of functions of $F$. Is $\gamma(\mathbb{F}_n)$ asymptotically linear?

**Do either problem 5 or problem 6.**

5. (48 points).

Suppose that $\underline{N}_n \equiv \sum_{i=1}^n \underline{\Delta}_i \sim \text{Mult}_k(n, \underline{p})$ where $\underline{\Delta}_i \sim \text{Mult}_k(1, \underline{p})$ are independent. The usual MLE $\hat{\underline{p}}_n$ of $\underline{p}$ is $\hat{\underline{p}}_n = n^{-1}\underline{N}_n$. For $a \in [0, 1)$ and $\underline{p}_0 = (1/k, \ldots, 1/k)$, consider the alternative estimator $\widetilde{\underline{p}}_n$ of $\underline{p}$ defined by

$$\widetilde{\underline{p}}_n = \begin{cases} \widehat{\underline{p}}_n & \text{if } \|\widehat{\underline{p}}_n - \underline{p}_0\|_2 > n^{-1/4}, \\ a\widehat{\underline{p}}_n + (1-a)\underline{p}_0 & \text{if } \|\widehat{\underline{p}}_n - \underline{p}_0\|_2 \leq n^{-1/4}. \end{cases}$$

Let $\underline{Z} \sim N_k(\underline{0}, \text{diag}(\underline{p}) - \underline{p}\underline{p}^T)$, $\underline{Z}_0 \sim N_k(\underline{0}, \text{diag}(\underline{p}_0) - \underline{p}_0\underline{p}_0^T)$.

(a) Find the limiting distribution of $\sqrt{n}(\widetilde{\underline{p}}_n - \underline{p}_0)$ when $\underline{p} \neq \underline{p}_0$ and when $\underline{p} = \underline{p}_0$. (Express these in terms of $\underline{Z}$ and $\underline{Z}_0$.)

(b) Let $\underline{p}_n = \underline{p}_0 + n^{-1/2}\underline{c}$ where $\underline{1}^T\underline{c} = 0$. Find the limiting distribution of $\sqrt{n}(\widetilde{\underline{p}}_n - \underline{p}_n)$ under $\underline{p}_n$.

(c) Is $\widetilde{\underline{p}}_n$ a locally regular estimator of $\underline{p}$ at $\underline{p} = \underline{p}_0$? Explain why or why not.

(d) Find the limit of $E_{\underline{p}_n} n\|\widetilde{\underline{p}}_n - \underline{p}_n\|^2$ where $\|\underline{v}\|^2 \equiv \sum_{j=1}^k v_j^2$ for $\underline{v} \in \mathbb{R}^k$. How does this compare to the limit of $E_{\underline{p}_n} n\|\widehat{\underline{p}}_n - \underline{p}_n\|^2$?

6. (48 points).
Consider a parametric model $\mathcal{P} = \{P_\theta : \ \theta \in \Theta \subset \mathbb{R}^d\}$ satisfying the hypotheses A0-A4 of section 4.1 of the Chapter 4 notes. Suppose that we are using the Rao (or score statistic) $R_n = \underline{Z}_n(\theta_0)^T I(\theta_0)^{-1} \underline{Z}_n(\theta_0)$ for testing the null hypothesis $H : \theta = \theta_0$ versus $K : \theta \ne \theta_0$ where $\underline{Z}_n(\theta_0) \equiv n^{-1/2} \sum_{i=1}^n \dot{\underline{l}}_\theta(\theta_0|X_i)$.
(a) Suppose that $\theta_n = \theta_0 + tn^{-1/2}$ and consider $l_n(\theta_n) - l_n(\theta_0) = \sum_{i=1}^n \log\{p_{\theta_n}(X_i)/p_{\theta_0}(X_i)$. Use the expansion developed in HW (and stated in part (vi) of Theorem 4.1.2) which implies LAN at $\theta_0$, together with the linearity of $\underline{Z}_n(\theta_0)$, to find the joint limit distribution of

$$\begin{pmatrix} \underline{c}^T \underline{Z}_n(\theta_0) \\ l_n(\theta_n) - l_n(\theta_0) \end{pmatrix} = \begin{pmatrix} \underline{c}^T \underline{Z}_n(\theta_0) \\ \log\left(\frac{dP_{\theta_n}^n}{dP_{\theta_0}^n}\right) \end{pmatrix}$$

under $P_{\theta_0}$.
(b) Use the result of (a) together with Le Cam's 3rd lemma to find the joint limiting distribution of

$$\begin{pmatrix} \underline{c}^T \underline{Z}_n(\theta_0) \\ l_n(\theta_n) - l_n(\theta_0) \end{pmatrix} = \begin{pmatrix} \underline{c}^T \underline{Z}_n(\theta_0) \\ \log\left(\frac{dP_{\theta_n}^n}{dP_{\theta_0}^n}\right) \end{pmatrix}$$

under $P_{\theta_n}$.
(c) Use the result of (b) to find the limiting distribution of $\underline{Z}_n(\theta_0)$ under $P_{\theta_n}$.
(d) Use the result of (c) to find the limiting distribution of the Rao statistic $R_n$ under $P_{\theta_n}$. What does this imply about the power of the Rao statistic?

7. (50 points).
Suppose that $P = P_0 = N(0,1)$, $Q = P_\theta = N(\theta, 1)$ on $(\mathbb{X}, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$.
(a) Compute $K(P,Q) = K(P_0, P_\theta)$.
(b) Compute $H^2(P,Q) = 1 - \rho(P,Q)$ and $\rho(P,Q) = \int \sqrt{p(x)q(x)}dx$. [It might be easiest to compute $\rho(P,Q)$ first recalling that if $Z \sim N(0,1)$ then $E\exp(tZ) = \exp(t^2/2)$.]
(c) Compute $d_{TV}(P,Q) = 1 - \eta(P,Q)$ and $\eta(P,Q) = \int p(x) \wedge q(x)dx$. [It might be easiest to compute $\eta(P,Q)$ first.]
(d) Show in general that $K(P,Q) \ge 2H^2(P,Q)$, thereby strengthening the fact $K(P,Q) \ge 0$ that we proved in class. [Hint: write both $K(P,Q)$ and $H^2(P,Q)$ in terms of $Y = (p(X)/q(X))^{1/2}$ and use the inequality $\log(1+x) \ge x/(1+x)$ for $x \ge 0$. You will need to relate $E_Q Y$ and $E_Q Y^2$ to $H^2(P,Q)$.]
(e) Use the results of (a) and (d) to find a lower bound for $K(P^n, Q^n)$ in terms of $H^2(P,Q)$ or $\rho(P,Q)$; here $P^n$ and $Q^n$ are the probability distributions of $X_1, \ldots, X_n$ i.i.d. as $P$ and $Q^n$ respectively.