Statistics 581, Final Exam

Wellner; 12/14/2001

1. (40) points) **Define** *each* of the following terms. In each case, provide an appropriate context for your definition.
   (a) The information matrix for θ in a regular parametric model $\mathcal{P} = \{P_\theta : \ \theta \in \Theta \subset R^k\}$.
   (b) The efficient score function for a parameter θ_1 when $\theta = (\theta_1, \theta_2)$.
   (c) The efficient influence function $\tilde{l}_1$ for a parameter θ_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 with influence function ψ.

2. (30 points) **State** *each* of the following results, providing the appropriate (brief) context for your statement:
   (a) The Glivenko-Cantelli theorem.
   (b) Donsker's theorem for the empirical process.
   (c) A result about the finite-dimensional limiting distributions of the sample quantile process.

3. (45 points). Suppose that $X, X_1, \ldots, X_n$ are i.i.d. positive random variables. Let

$$A_n = \overline{X}_n \ = \ n^{-1} \sum_{i=1}^n X_i \,,$$

$$G_n = \left\{ \prod_{i=1}^n X_i \right\}^{1/n}, \qquad \text{and} \qquad H_n = \frac{1}{\frac{1}{n}\sum_{i=1}^n X_i^{-1}}$$

be the arithmetic, geometric, and harmonic means, respectively.

A. Use the weak law of large numbers and the continuous mapping theorem to show that $A_n \to_p a$, $G_n \to_p g$, and $H_n \to_p h$ for some constants a, g, and h depending on the distribution of X under appropriate integrability assumptions. Specify these assumptions precisely and identify the constants a, g, and h.

B. Compute the constants a, g, and h when $X \sim (1/2)$ Uniform$(1,2) + (1/2)\delta_1$; i.e. when the distribution function F of the X's is given by

$$F(x) = \left\{ \begin{array}{rl} 0\,, & x < 1 \\ x/2\,, & 1 \le x \le 2 \\ 1\,, & 2 \le x < \infty \end{array} \right\} .$$

C. Use the multivariate central limit theorem and the delta-method to show that under appropriate assumptions on the distribution of X

$$\sqrt{n} \begin{pmatrix} A_n - a \\ G_n - g \\ H_n - h \end{pmatrix} \to_d N_3(0, V)$$

for some covariance matrix V. Identify the assumptions you need explicitly. Compute V as explicitly as possible.

Do **either** problem 4 **or** problem 5.

4. (40 points).
For a regular parametric model satisfying the Cramér hypotheses of our Theorem 4.1.5 of Chapter 4, the local likelihood ratios satisfy the LAN condition: under P_{θ_0}

$$\log \frac{L_n(\theta_0 + tn^{-1/2})}{L_n(\theta_0)} = tZ_n - \frac{1}{2}t^T I(\theta_0)t + o_p(1) \to_d t^T Z - \frac{1}{2}t^T I(\theta_0)t \qquad (1)$$

where $Z \sim N_k(0, I(\theta_0))$.
A. Show that the right side of (1) can be written as

$$-\frac{1}{2}(t - W)^T I(\theta_0)(t - W) + \frac{1}{2}W^T I(\theta_0)W$$

where W is some linear transformation of Z. Identify W explicitly in terms of $I(\theta_0)$ and Z.
B. The displayed equation in part A shows that the limit is maximized as a function of t by $t = W$ and the maxmimum value is $(1/2)W^T I(\theta_0)W$. Rewrite both of these relations in terms of $I(\theta_0)$ and Z and interpret the result.
Hint: If we interpret the maximum likelihood estimator $\hat{\theta}_n$ of θ (assuming it exists) as $\text{argmax}_\theta \log L_n(\theta)$, then

$$\text{argmax}_t \left(\log L_n(\theta_0 + tn^{-1/2})/L_n(\theta_0)\right) = \sqrt{n}(\hat{\theta}_n - \theta_0)$$

while the argmax of the right side of (1) is just W; when W is rewritten in terms of $I(\theta_0)$ and Z this becomes quite natural!

5. (40 points). Suppose that $Z \sim P_{\mu,\Sigma} = N_k(\mu, \Sigma)$ where Σ is positive definite.
A. What is the density of Z with respect to Lebesgue measure on R^k?
B. Is $P_{\mu,\Sigma} << P_{0,\Sigma}$ as measures? If your answer is yes, explain why and compute the Radon-Nikodym derivative

$$\frac{dP_{\mu,\Sigma}}{dP_{0,\Sigma}}(z)\,.$$

Hint: if $P << \mu$, $Q << \mu$, and $Q << P$, then

$$\frac{dQ}{dP} = \frac{dQ/d\mu}{dP/d\mu}\,.$$

C. Suppose $\mu = I(\theta_0)t$, $\Sigma = I(\theta_0)$, so that $Z \sim P_{I(\theta_0)t, I(\theta_0)}$. Use the result of B to compute

$$\log\left\{\frac{dP_{I(\theta_0)t, I(\theta_0)}}{dP_{0, I(\theta_0)}}\right\}(Z)\,.$$

Explain the meaning of this computation for the Local Asymptotic Normality (LAN) part of Theorem 4.1.5.

Do **either** problem 6 **or** problem 7.

6. (40 points). Suppose that $\mathcal{P} = \{P_\theta : \theta \in \Theta \subset R^k\}$ is a regular parametric model with densities p_θ with respect to a dominating measure μ. Suppose that $\theta_0 \in \Theta$ is fixed.
   A. Express $K(P_{\theta_0}, P_\theta)$ in terms of the densities p_{θ_0} and p_θ.
   B. If $g(\theta) \equiv K(P_{\theta_0}, P_\theta)$, compute (assuming that the interchange of integration and differentiation is permissible)
$$\dot{g}(\theta) = (\frac{\partial}{\partial\theta_1} g(\theta), \ldots, \frac{\partial}{\partial\theta_k} g(\theta))^T$$
   and the $k \times k$ matrix
$$\ddot{g}(\theta) = \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} g(\theta) \right) .$$
   Interpret $\dot{g}(\theta)$ in terms of a particular constant involved in the limit behavior of the Rao (or score) statistic under fixed alternatives. Evaluate both $\dot{g}(\theta)$ and $\ddot{g}(\theta)$ at θ_0, and identify the special values involved.
   C. Carry out the computations in A and B explicitly when $p_\theta(x) = \theta^x e^{-\theta}/x!$ for $x = 0, 1, 2, \ldots$, $\theta > 0$.

7. (40 points)
   Suppose that $\underline{X}, \underline{X}_1, \ldots, \underline{X}_n$ are i.i.d. $\text{Mult}_k(1, \underline{p})$, so that $\underline{N}_n \equiv \sum_{i=1}^n \underline{X}_i \sim \text{Mult}_k(n, \underline{p})$. Thus
$$P_{\underline{p}}(\underline{X} = \underline{x}) = \prod_{j=1}^k p_j^{x_j} \quad \text{for } \; x_i \in \{0, 1\}, \;\; \sum_1^k x_i = 1\,,$$
$$P_{\underline{p},n}(\underline{N}_n = \underline{m}) = \frac{n!}{\prod_{j=1}^k m_j!} \prod_{j=1}^k p_j^{m_j} \quad \text{for } \; m_i \geq 0, \text{ integers } \; \sum_{j=1}^k m_j = n\,.$$
   A. Compute $K(P_{\underline{q}}, P_{\underline{p}}) \equiv K(\underline{q}, \underline{p})$ for vectors $\underline{q}$, $\underline{p}$ with $\sum p_j = \sum q_j = 1$.
   B. Evaluate $K(\widehat{\underline{p}}, \underline{p})$ where $\widehat{\underline{p}} = n^{-1}\underline{N}_n$. Relate this to the log-likelihood $\log L_n(\underline{p}|\underline{N}_n)$.
   C. Use the result of B to show, without using any calculus, that the MLE of $\underline{p}$ is $\widehat{\underline{p}} = \underline{N}/n$.

Do **either** problem 8 **or** problem 9.

8. (48 points). (Poisson regression). Much as in problem set 8, Suppose that $(Y|Z) \sim \text{Poisson}(\lambda e^{\gamma Z})$, and $Z \sim \text{Bernoulli}(\eta)$ on $\{0,1\}$ You may assume that η is known. Thus Z is a "covariate" or "predictor variable", γ is a "regression parameter" which affects the intensity of the (conditionally) Poisson variable Y, and $\theta = (\lambda, \gamma)$.
(a) Find the information matrix for θ.
(b) Find the information and information bound for estimation γ if the parameter λ is unknown.
(c) Find the efficient score function and the efficient influence function for estimation of γ when λ is unknown. Interpret these in terms of the scores for γ and λ
(d) If we observe $X_i = (Y_i, Z_i)$, $i = 1, \ldots, n$, i.i.d. P_θ, write down the likelihood equations for the maximum likelihood estimator $\hat{\theta}_n = (\hat{\lambda}_n, \hat{\gamma}_n)$. What do our theorems tell us about the asymptotic normality of $\hat{\theta}_n$?

9. (48 points). (Poisson regression, continued).
(a) Suggest three tests of the (composite!) null hypothesis $H : \gamma = 0$ versus $K : \gamma \neq 0$. What is the distribution of each of these three statistics under the null hypothesis and under local alternatives of the form $\gamma_n = tn^{-1/2}$?
(b) Consider estimation of the function

$$q(\theta) = \nu(P_\theta) = P_\theta(Y = 0)\,.$$

Compute $q(\theta)$ explicitly as a function of θ.
(c) Suggest a natural empirical estimator of this probability (which does not rely on the Poisson model). If this estimator is called $\tilde{\nu}_n$, show that $\tilde{\nu}_n$ is asymptotically linear and find its influence function ψ explicitly.
(d) Find the efficient influence function $\tilde{l}_\nu$ for estimation of $\nu(P_\theta)$ assuming the Poisson model.
(e) Describe the relationship between ψ and $\tilde{l}_\nu$ geometrically.