Statistics 581, Midterm Exam

Wellner; 11/06/2006

This exam is to be taken without any books or notes.

1. (24 points) **Define any three of the following five terms**.
(a) A *uniformly integrable* sequence of random variables.
(b) *Convergence in rth mean* of a sequence of random variables.
(c) *A normal random vector* $Y = (Y_1, \ldots, Y_n)$.
(d) The *inverse or quantile function* F^{-1} of a distribution function F.
(e) The *total variation distance* between two probability measures P and Q on a measurable space $(\mathcal{X}, \mathcal{A})$.

Do **either** problem 2 **or** problem 3.

2. (40 points).
(a) State the ordinary (univariate) central limit theorem.
(b) State the Cramér-Wold device.
(c) State the multivariate central limit theorem.
(d) Use the ordinary (univariate) central limit theorem (a) and the Cramér-Wold device (b) to prove the multivariate central limit theorem (c).

3. (40 points). State and prove the Glivenko-Cantelli theorem.

4. (36 points). Let $X_1, \ldots, X_n$ be i.i.d. with exponential density $p_\theta(x) = \theta \exp(-\theta x) 1_{[0,\infty)}(x)$.
(a) Find a constant c so that $c\mathbb{F}_n^{-1}(p) \to_p \theta^{-1}$.
(b) For the $c = c_p$ you found in (a), show that $\sqrt{n}(c\mathbb{F}_n^{-1}(p) - \theta^{-1}) \to_d N(0, \sigma^2)$ and find $\sigma^2 = \sigma^2(p)$.
(c) Show that the asymptotic variance $\sigma^2(p)$ is minimized by p satisfying $2p = -\log(1-p)$.

5. (36 points)
Suppose that $X, X_1, \ldots, X_n$ are i.i.d. Exponential(θ) random variables so that $P_\theta(X > x) = \exp(-\theta x) = 1 - F_\theta(x)$ for $x > 0$.
(a) Fix $x_0 > 0$ and let $\mathbb{F}_n(x) = n^{-1}\sum_{i=1}^n 1_{[X_i \le x]} = n^{-1}\sum_{i=1}^n 1_{(-\infty,x]}(X_i)$ denote the empirical distribution function. Show that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - 1/\theta \\ \mathbb{F}_n(x_0) - F_\theta(x_0) \end{pmatrix} \to_d Y \sim N_2(0, \Sigma)$$

and find Σ.

(b) Let $g(\theta) \equiv F_\theta(x_0) = 1 - \exp(-\theta x_0)$, and consider the two estimators of $F = F_\theta$ given by $T_{n,1} \equiv g(\widehat{\theta}_n)$ and $T_{n,2} \equiv \mathbb{F}_n(x_0)$ where $\widehat{\theta}_n \equiv 1/\overline{X}_n$. Show that

$$\sqrt{n}\left(\begin{array}{c} T_{n,1} - F_\theta(x_0) \\ T_{n,2} - F_\theta(x_0) \end{array} \right) \rightarrow_d \tilde{Y}$$

and find the distribution of $\tilde{Y}$.

(c) What is the advantage of $T_{n,2} = \mathbb{F}_n(x_0)$ as an estimator even though it is inefficient when the exponential model holds?

6. (36 points).
Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with distribution function F given by $P(X > x) = 1 - F(x) = 1/x^5$, $x \geq 1$, $F(x) = 0$, $x \leq 1$.
(a) For what values of $r > 0$ is $E|X|^r < \infty$? If they are finite compute $\mu = E(X)$ and $\sigma^2 = Var(X)$.
(b) Compute $F^{-1}(t) = Q(t)$, the quantile function corresponding to F.
(c) Which of the following are true? (Briefly indicate why or why not.)

(i) $\sum_{i=1}^n X_i = O_p(n^{1/2})$.

(ii) $n^{1/3}(\overline{X}_n - \mu) = o_p(1)$.

(iii) $n^{3/4}(\overline{X}_n - \mu) = O_p(1)$.

(iv) $g(n^{1/3}(\overline{X}_n - \mu)) \rightarrow_p 1/2$ where $g(x) = 1/(1 + e^{-x})$.

(v) $h(n^{1/2}(\overline{X}_n - \mu)) = O_p(1)$ with $h(x) = \log|x|$.

(vi) $\sqrt{n}(\mathbb{F}_n^{-1}(1/2) - F^{-1}(1/2)) \rightarrow_d N(0, (1/4)/[5(1/2)^{6/5}]^2)$.