Statistics 581, Midterm Exam Solutions

Wellner; 11/05/2008

1. (24 points) **Define** any **three** of the following five terms.
   (a) A *Brownian bridge process* $\mathbb{U}$.
   (b) *Convergence in distribution* of a sequence of random vectors X_n in $\mathbb{R}^k$.
   (c) *A normal random vector* $Y = (Y_1, \ldots, Y_n)$.
   (d) A *non-central chi-square distribution* with m degrees of freedom and non-centrality parameter δ.
   (e) The *Hellinger distance* between two probability measures P and Q on a measurable space $(\mathcal{X}, \mathcal{A})$.

   Solution: See notes, Chapters 1 and 2.

2. (36 points). **State** any **three** of the following:
   (a) The Lindeberg-Feller central limit theorem.
   (b) The inverse transformation theorem.
   (c) The multivariate delta method or g' theorem.
   (d) The Mann-Wald or continuous mapping theorem.
   (e) The Glivenko-Cantelli theorem.
   (f) A limit theorem about the joint limiting distribution of $(\sqrt{n}(\mathbb{F}_n^{-1}(s) - F^{-1}(s)), \sqrt{n}(\mathbb{F}_n^{-1}(t) - F^{-1}(t)))$ where $0 < s < t < 1$ are fixed.

   Solution: See notes, Chapters 1 and 2.

   Do **either** problem 3 **or** problem 4.

3. (40 points) Suppose that $\underline{N} = (N_1, \ldots, N_k) \sim \text{Mult}_k(n, \underline{p})$ where $\underline{p} = (p_1, \ldots, p_k)$. In class and homework problems we have discussed the chi-square statistic Q_n and the Hellinger distance statistic $4nH_n^2$ as test statistics for testing $H : \underline{p} = \underline{p_0}$ versus $K : \underline{p} \neq \underline{p}_0$. An alternative statistic for testing H versus K is the likelihood ratio statistic $2\log\lambda_n$ where

$$\lambda_n \equiv \frac{\sup_{\underline{p}} L_n(\underline{p})}{L_n(\underline{p}_0)} = \frac{\prod_{j=1}^k \widehat{p}_j^{N_j}}{\prod_{j=1}^k p_{0j}^{N_j}} = \prod_{j=1}^k \left\{ \frac{\widehat{p}_j}{p_{0j}} \right\}^{N_j} .$$

   (a) Show that

$$2\log\lambda_n = 2n \sum_{j=1}^k \widehat{p}_j \log\left(\frac{\widehat{p}_j}{p_{0j}} \right) .$$

(b) If the alternative hypothesis K is true, so $\underline{p} \neq \underline{p}_0$, show that

$$n^{-1} 2 \log \lambda_n = g(\widehat{\underline{p}}) \rightarrow_p g(\underline{p}) ,$$

and identify $g(\underline{p})$ as a function of $\underline{p}$ and $\underline{p_0}$.
(c) If the alternative hypothesis K is true, so $\underline{p} \neq \underline{p}_0$, show that

$$\sqrt{n}(2n^{-1} \log \lambda_n - g(\underline{p})) = \sqrt{n}(g(\widehat{\underline{p}}) - g(\underline{p})) \rightarrow_d N(0, V^2(\underline{p})) ,$$

and compute $V^2(\underline{p})$. Could you use this to approximate the power of the likelihood-ratio test? How?

Solution: (a) We have

$$\begin{aligned} 2 \log \lambda_n &= 2 \log \left(\prod_{j=1}^{k} \left(\frac{\widehat{p}_j}{p_{0,j}} \right)^{N_j} \right) = 2 \sum_{j=1}^{k} \log \left(\frac{\widehat{p}_j}{p_{0,j}} \right)^{N_j} \\ &= 2 \sum_{j=1}^{k} N_j \log \left(\frac{\widehat{p}_j}{p_{0,j}} \right) = 2n \sum_{j=1}^{k} \widehat{p}_j \log \left(\frac{\widehat{p}_j}{p_{0,j}} \right) \end{aligned}$$

using $\widehat{p}_j = n^{-1} N_j$ in the last line.
(b) It follows immediately from (a) that $2n^{-1} \log \lambda_n = g(\widehat{p})$ where $g : [0,1]^k \rightarrow \mathbb{R}$ is given by

$$g(u) = 2 \sum_{j=1}^{k} u_j \log(u_j / p_{0,j}).$$

(c) By the continuous mapping theorem applied to $g(\widehat{p}_n)$ where $\widehat{p}_n \rightarrow_p p$, it follows immediately that $2n^{-1} \log \lambda_n \rightarrow_p g(p) = 2K(P, P_0)$ where $K(P, P_0)$ is the Kullback-Leibler divergence between P and P_0 determined by p and p_0 respectively.
(d) By (b) we can write

$$\begin{aligned} \sqrt{n}(2n^{-1} \log \lambda_n - g(p)) &= \sqrt{n}(g(\widehat{p}_n) - g(p)) \\ &\rightarrow_d g'(p) Z \sim N(0, g'(p) \Sigma g'(p)^T) \end{aligned}$$

by the delta-method or $g'-$ theorem where the components of the vector of derivatives $g'(p)$ are given by

$$\frac{\partial}{\partial u_j} g(u) \Big|_{u=p} = 2(\log(p_j / p_{0,j}) + 1) \equiv d_j$$

for $j = 1, \ldots, k$, and $\Sigma = \text{diag}(p) - pp^T$. Thus $V^2 = d\Sigma d^T = \sum_{j=1}^k (d_j^2 p_j) - (d^T p)^2 = Var_p(D)$ where $P(D = d_j) = p_j$, $j = 1, \ldots, k$. This can be used to approximate the power of the likelihood ratio test:

$$\begin{aligned} P_p(2\log\lambda_n > \chi^2_{k-1,\alpha}) &= P_p(\sqrt{n}(n^{-1}2\log\lambda_n - g(p)) > \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - g(p))) \\ &\doteq P(N(0, V^2) > \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - g(p))). \end{aligned}$$

A non-central chi-square approximation of the power based on local alternatives is also possible.

Beyond the exam question: A question not addressed by the exam problem concerns the asymptotic distribution of $2\log\lambda_n$ under the null hypothesis. Here we briefly sketch a proof of $2\log\lambda_n \to_d \chi^2_{k-1}$. Note that by Taylor expansion we can write

$$g(u) = g(p_0) + g'(p_0)^T(u - p_0) + \frac{1}{2}(u - p_0)^T g''(u^*)(u - p_0)$$

where $|p^* - p_0| \le |u - p_0|$; here g'' is a $k \times k$ matrix. Now $g(p_0) = 0$, and from the calculation in (d) above, $g'(p_0) = 2\underline{1} = 2(1, \ldots, 1)^T$, and $g''(u) = \text{diag}(2/u)$. Thus we find, using $\underline{1}^T\widehat{p}_n = 1 = \underline{1}^T p_0$,

$$\begin{aligned} g(\widehat{p}_n) &= 2\underline{1}^T(\widehat{p}_n - p_0) + (\widehat{p}_n - p_0)^T \text{diag}(1/p_n^*)(\widehat{p}_n - p_0) \\ &= (\widehat{p}_n - p_0)^T \text{diag}(1/p_n^*)(\widehat{p}_n - p_0) \end{aligned}$$

and hence

$$2\log\lambda_n = ng(\widehat{p}_n) = n(\widehat{p}_n - p_0)^T \text{diag}(1/p_n^*)(\widehat{p}_n - p_0)$$

where p_n^* satisfies $|p_n^* - p_0| \le |\widehat{p}_n - p_0| \to 0$. Thus $2\log\lambda_n = Q_n + o_p(1) \to_d \chi^2_{k-1}$ under the null hypothesis $p = p_0$.

4. (40 points). Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F having a continuous density function f. Let $\mathbb{F}_n$ be the empirical distribution function of the X_i's, and suppose that b_n is a sequence of positive numbers, and let

$$\hat{f}_n(x) = \frac{\mathbb{F}_n(x + b_n) - \mathbb{F}_n(x - b_n)}{2b_n}.$$

(a) Show that $E\hat{f}_n(x) \to f(x)$ if $b_n \to 0$.
(b) Show that $Var(\hat{f}_n(x)) \to 0$ if $b_n \to 0$ and $nb_n \to \infty$.
(c) Use some appropriate central limit theorem to show that (perhaps under some suitable further conditions that you might need to specify)

$$\sqrt{2nb_n}(\hat{f}_n(x) - E\hat{f}_n(x)) \to_d N(0, f(x)).$$

Hint: Write $\hat{f}_n(x)$ in terms of some Bernoulli random variables and identify $p = p_n$.
Note: This estimator $\hat{f}_n$ is a *kernel density estimator* based on the uniform kernel $k(x) = 1_{[-1,1]}(x)/2$, and can be rewritten as

$$\hat{f}_n(x) = \int_{-\infty}^{\infty} \frac{1}{b_n} k\left(\frac{x-y}{b_n}\right) d\mathbb{F}_n(y);$$

other kernel density estimators result when the uniform kernel is replaced by some other density function.

Solution: (a) First note that $2nb_n\widehat{f}_n(x) = n(\mathbb{F}_n(x+b_n) - \mathbb{F}_n(x-b_n))$ is a Binomial(n, p_n) random variable with $p_n = F(x+b_n) - F(x-b_n)$. Hence if $b_n \to 0$

$$\begin{aligned} E\widehat{f}_n(x) &= \frac{F(x+b_n) - F(x-b_n)}{2b_n} = \frac{p_n}{2b_n} \\ &= \frac{1}{2}\left\{\frac{F(x+b_n) - F(x)}{b_n} + \frac{F(x) - F(x-b_n)}{b_n}\right\} \\ &\to \frac{1}{2}\{f(x) + f(x)\} = f(x). \end{aligned}$$

(b) Furthermore

$$\begin{aligned} Var(\widehat{f}_n(x)) &= \frac{np_n(1-p_n)}{(2nb_n)^2} \\ &= \frac{1}{2nb_n}\frac{p_n}{2b_n}(1-p_n) \\ &\to 0 \cdot f(x) \cdot 1 = 0 \end{aligned}$$

if $nb_n \to \infty$ and $b_n \to 0$.
(c) Since $2nb_n\widehat{f}_n(x) = \sum_{i=1}^n X_{ni}$ where $X_{ni} \sim$ Bernoulli(p_n), it follows that $\sigma_{ni}^2 = p_n(1-p_n)$ so that $\sigma_n^2 = \text{Var}(\sum_{i=1}^n X_{ni}) = np_n(1-p_n)$, and

$$\begin{aligned} \gamma_n \equiv \sum_{i=1}^n \gamma_{ni} &= \sum_{i=1}^n E|X_{ni} - \mu_{ni}|^3 \\ &= np_n(1-p_n)\{(1-p_n)^2 + p_n^2\} \\ &\le 2np_n(1-p_n) \end{aligned}$$

so that

$$\gamma_n/\sigma^3 \le \frac{2}{\sqrt{np_n(1-p_n)}} = \frac{2}{\sqrt{nb_n(p_n/b_n)(1-p_n)}} \to 0$$

if $b_n \to 0$ and $nb_n \to \infty$. Thus, by the Liapunov CLT,

$$\frac{2nb_n(\widehat{f}_n(x) - E\widehat{f}(x))}{\sqrt{np_n(1-p_n)}} \to N(0,1)$$

if $b_n \to 0$ and $nb_n \to \infty$. Thus

$$\begin{aligned}\sqrt{2nb_n}(\widehat{f}_n(x) - E\widehat{f}_n(x)) &= \frac{2nb_n(\widehat{f}_n(x) - E\widehat{f}_n(x))}{\sqrt{np_n(1-p_n)}}\sqrt{\frac{np_n(1-p_n)}{2nb_n}} \\ &\to N(0,1)\sqrt{f(x)} = N(0, f(x)).\end{aligned}$$

Beyond the exam question: There are at least two further questions of interest here:
(a) Under what conditions is the limiting distribution of $\sqrt{2nb_n}(\hat{f}_n(x) - f(x))$ the same as the limiting distribution derived in (c) for $\sqrt{2nb_n}(\hat{f}_n(x) - E\hat{f}_n(x))$?
(b) What is the joint limiting distribution of

$$\sqrt{2nb_n}\begin{pmatrix} \hat{f}_n(x) - E\hat{f}_n(x) \\ \hat{f}_n(y) - E\hat{f}_n(y) \end{pmatrix}$$

for $x, y \in \mathbb{R}$ with $x < y$? To answer (a) we need to find conditions on F or f so that $\sqrt{2nb_n}(E\hat{f}_n(x) - f(x)) \to 0$. To do this, suppose that f has a continuous derivative f' in a neighborhood of x. Then

$$F(x+h) = F(x) + hf(x) + \frac{1}{2}h^2 f'(x_h^*)$$

where $|x_h^* - x| \le |h|$. Therefore

$$\frac{F(x+h) - F(x-h)}{2h} - f(x) = \frac{1}{2}h\left\{f'(x_h^*) - f'(x_{-h}^*)\right\},$$

so that

$$|\sqrt{nb_n}\{E\hat{f}_n(x) - f(x)\}| \le \sqrt{nb_n}\frac{1}{2}b_n|f'(x_{b_n}^*) - f'(x_{-b_n}^*)| \to 0$$

if $n^{1/2}b_n^{3/2} \to C \in [0,\infty)$ since $f'(x_{b_n}^*) - f'(x_{-b_n}^*) \to f'(x) - f'(x) = 0$ as $n \to \infty$. Thus the bias vanishes asymptotically if $b_n = cn^{-1/3}$ for some $c > 0$ and f' exists and is continuous in a neighborhood of x.

5. (35 points). Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. Uniform$(0, \theta)$. Consider $\hat{\theta}_n = \max_{1 \le i \le n} X_i = X_{(n)}$ as an estimator of θ.

(a) Compute $E(\hat{\theta}_n)$.

(b) Show that $Y_n = n(\theta - \hat{\theta}_n) \to_d Y$ (c) Show that $\hat{\theta}_n \to_p \theta$.
and find the distribution of Y.

(d) Consider the function $g(x) = (1-x)^{-2}$. Does $g(Y_n) \to_d$ something? If the answer is yes, what is the limit (expressed in terms of the random variable Y)?

(e) Consider the function $g(x) = \log x$. Find the limiting distribution of $n(g(\hat{\theta}_n) - g(\theta))$.

Solution: (a) Note that with $\xi_1, \ldots, \xi_n$ i.i.d. Uniform$(0,1)$ we have $X_i \stackrel{d}{=} \theta \xi_i$, $1 \le i \le n$, and hence $X_{(n)} \stackrel{d}{=} \theta \xi_{(n)}$. Therefore

$$E(\hat{\theta}_n) = E(X_{(n)}) = E(\theta \xi_{(n)}) = \theta \frac{n}{n+1}.$$

(b) Now

$$P(\hat{\theta}_n \le t) = P(X_{(n)} \le t) = P(X_1 \le t, \ldots, X_n \le t) = (t/\theta)^n$$

for $0 \le t \le \theta$. Therefore

$$\begin{aligned} P(n(\theta - \hat{\theta}_n) \ge t) &= P(\hat{\theta}_n \le \theta - t/n) = \left\{ \frac{\theta - t/n}{\theta} \right\}^n \\ &= \left(1 - \frac{t}{n\theta} \right)^n \to \exp(-t/\theta). \end{aligned}$$

Thus

$$Y_n \equiv n(\theta - \hat{\theta}_n) \to_d Y \sim \text{ Exponential}(1/\theta).$$

(c) From (b) we conclude that

$$\hat{\theta}_n - \theta = n^{-1} n(\hat{\theta}_n - \theta) = o(1) O_p(1) = o_p(1),$$

so $\hat{\theta}_n - \theta \to_p 0$, or $\hat{\theta}_n \to_p \theta$.

(d) Now $Y_n \to_d Y \sim$ Exponential$(1/\theta)$, and $g(x) = 1/(1-x)^2$ is continuous a.e. P_Y since $P(Y = 1) = 0$. Thus $g(Y_n) \to_d g(Y)$ by the Mann-Wald theorem.

(e) The function $g(x) = \log x$ is differentiable at all $\theta > 0$ (since $g'(x) = 1/x$ for $x > 0$). Since $n(\hat{\theta}_n - \theta) \to_d -Y$, it follows from the delta method that

$$n(g(\hat{\theta}_n - g(\theta)) \to_d g'(\theta)(-Y) \stackrel{d}{=} -\frac{1}{\theta} Y \sim -\text{Exponential}(\theta).$$

6. (36 points).

Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with distribution function F given by $P(X > x) = 1 - F(x) = 1/(1+x)^3$, $x \ge 0$.

(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, and $f(x)$, the density function of F at x.

(c) Which of the following are true? (Briefly indicate why or why not.)

(i) $\sum_{i=1}^n X_i = O_p(n^{3/4})$.

(ii) $n^{1/5}(\overline{X}_n - \mu) = o_p(1)$.

(iii) $n^{2/3}(\overline{X}_n - \mu) = O_p(1)$.

(iv) $g(n^{1/3}(\overline{X}_n - \mu)) \to_p 1$ where $g(x) = \exp(3x)$.

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

(vi) $\sqrt{n}(\mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4)) \to_d N(0, 3^{-1}(1/4)^{14/3})$.

Solution: (a) Since $E|X|^r = EX^r = r\int_0^\infty x^{r-1}(1 - F(x))dx$, in this case we find that

$$E|X|^r = r\int_0^\infty \frac{x^{r-1}}{(1+x)^3}dx < \infty$$

if $3 - (r-1) > 1$, or if $r < 3$. Thus $E|X|$ and $E|X|^2$ are finite, and we comute

$$\begin{aligned} \mu &= E(X) = \int_0^\infty (1+x)^{-3}dx = -2^{-1}(1+x)^{-2}\Big|_0^\infty = 1/2, \\ E(X^2) &= 2\int_0^\infty \frac{x}{(1+x)^3}dx = 2\left\{\int_0^\infty \frac{1+x}{(1+x)^3}dx - \int_0^\infty \frac{1}{(1+x)^3}dx\right\} \\ &= 2\{1 - 1/2\} = 1, \end{aligned}$$

so that $\sigma^2 \equiv Var(X) = 1 - (1/2)^2 = 1 - 1/4 = 3/4$.

(b) Setting $F(x) = 1 - (1+x)^{-3} = u$, we find that $F^{-1}(u) = (1-u)^{-1/3} - 1$. The density function f is given by $f(x) = 3(1+x)^{-4}$.

(c) (i) is false since $n^{-1}\sum_1^n X_i \to_p \mu = 1/2$.

(ii) (ii) is true since $n^{1/5}(\overline{X}_n - \mu) = (n^{1/5}/n^{1/2}) \cdot \sqrt{n}(\overline{X}_n - \mu) = o(1)O_p(1) = o_p(1)$.

(iii) This is clearly False since $n^{2/3}(\overline{X}_n - \mu) = n^{1/6}n^{1/2}(\overline{X}_n - \mu)$.

(iv) Since $n^{1/3}(\overline{X}_n - \mu) \to_p 0$ and since $g(x) = exp(3x)$ is continuous at 0, the claim is true: $g(n^{1/3}(\overline{X}_n - \mu)) \to_p g(0) = 1$.

(v) True, since $\sqrt{n}(\overline{X}_n - \mu) \to_d Y \sim N(0, \sigma^2)$ and since h is discontinuous at the countable set of points $D = \{\pm k\pi : k \in \mathbb{Z}\}$ which has $P(Y \in D) = 0$. Thus $h(\sqrt{n}(\overline{X}_n - \mu)) \to_d h(Y)$ and it follows that $h(\sqrt{n}(\overline{X}_n - \mu)) = O_p(1)$.

(vi) True, since

$$\sqrt{n}(\mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4)) \to_d N(0, \frac{(1/4)(3/4)}{f^2(F^{-1}(3/4))})$$

where $f(F^{-1}(u)) = 3(1-u)^{4/3}$ so $f(F^{-1}(3/4)) = 3 \cdot (1/4)^{4/3}$, and hence

$$\frac{(3/4)(1/4)}{f^2(F^{-1}(3/4))} = \frac{3(1/4)^2}{\{3(1/4)^{4/3}\}^2} = \frac{1}{3}4^{2/3} = \frac{1}{3}2^{4/3}.$$