Midterm Examination

Statistics 200C

Ferguson — Friday, May 7, 2010

1. Suppose that X has a Poisson distribution with mean λ, $\mathcal{P}(\lambda)$, and that X_n is a sequence of random variables such that $X_n \xrightarrow{\mathcal{L}} X$. Is it necessarily true that

(a) $X_n \mathrm{I}(X_n > 0) \xrightarrow{\mathcal{L}} X$?

(b) $\sqrt{n}(X_n - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda)$?

(c) $\mathrm{E}\left[\frac{e^{X_n}}{1+e^{X_n}}\right] \to \mathrm{E}\left[\frac{e^{X}}{1+e^{X}}\right]$?

(For the "yes" answers, state what theorem are you using.)

2. Let $X_1, X_2, \ldots$ be independent random variables and that X_j has a uniform distribution on the interval $(0, 2j)$. Let $S_n = \sum_{j=1}^{n} X_j$.

(a) Find the mean and variance of S_n.

(b) Show that $(S_n - \mathrm{E}(S_n))/\sqrt{\mathrm{Var}(S_n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$, by checking the Lindeberg conditions.

3. Consider a multinomial distribution with 6 cells, sample size n, and probability vector $\boldsymbol{p} = \frac{1}{3}(p, p, p, 1-p, 1-p, 1-p)$ for some unknown probability p, $0 < p < 1$. Let n_j denote the number of observations that fall in cell j.

(a) What is Pearson's chi-square for testing the hypothesis $H_0 : p = .5$? What is its asymptotic distribution under this hypothesis?

(b) What is Hellinger's chi-square for testing H_0?

(c) What is the approximate large sample distribution of Hellinger's chi-square if the true value of p is .4?

4. A waiting room contains a long line of n chairs. As customers enter the room, they are seated in the chairs in order of arrival. Male customers appear with probability p and female customers appear with probability $1-p$, independently. Let S_n denote the number of female customers that are seated between two male customers in the line.

(a) Find the mean and variance of S_n.

(b) Find is the asymptotic (for large n) distribution of S_n properly normalized.

5. Let $X_1, X_2, \ldots, X_n$ be a sample from the distribution with distribution function $F(x) = 1 - (1/x)$ on the interval $(1, \infty)$.

(a) Find the pth quantile, x_p for this distribution, $0 < p < 1$.

(b) Find the joint asymptotic distribution of the first and third sample terciles, $\hat{x}_{1/3}$ and $\hat{x}_{2/3}$, properly normalized.

Solutions to Midterm Examination
Statistics 200C
T. Ferguson Fri. May 7, 2010

1. (a) True. From Slutsky's Theorem, since $g(x) = x\mathrm{I}(x > 0)$ is continuous.

(b) Very, very false.

(c) True. From the Helly-Bray Theorem, since $g(x) = \exp(x)/(1+\exp(x))$ is bounded and continuous.

2. (a) Since X_j has the same distribution as $2jU_j$ where U_j has a uniform distributio on (0,1), we have $\mathrm{E}(X_j) = j$ and $\mathrm{Var}(X_j) = \mathrm{Var}(2jU_j) = 4j^2(1/12) = j^2/3$. So $\mathrm{E}(S_n) = \sum_1^n j = n(n+1)/2$ and $\mathrm{Var}(S_n) = \sum_1^n j^2/3 = n(n+1)(2n+1)/18$.

(b) Let $X_{nj} = X_j - j$, so that $\mathrm{E}X_{nj} = 0$. Then for $B_n^2 = \mathrm{Var}(S_n)$, we have $Z_n/B_n = (S_n - \mathrm{E}(S_n))/B_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ provided the Lindeberg condition holds. Using $X_{nj}^2 \le n^2$ for all $j \le n$, we have

$$\frac{1}{B_n^2}\sum_{j=1}^{\infty}\mathrm{E}[X_{nj}^2\mathrm{I}(X_{nj}^2 > \epsilon^2 B_n^2)] \le \frac{1}{B_n^2}\sum_{j=1}^{\infty}\mathrm{E}[X_{nj}^2\mathrm{I}(n^2 > \epsilon^2 B_n^2)]$$
$$= \mathrm{I}(n^2 > \epsilon^2 B_n^2) = 0 \quad \text{for all n sufficiently large}$$

3. (a) $\chi^2_{\mathrm{P}} = n\sum_{j=1}^6((n_i/n) - (1/6))^2/(1/6)$. Under H_0, χ^2_{P} has asymptotically a chi-square distribution with 5 degrees of freedom.

(b) $\chi^2_H = 4n\sum_1^6(\sqrt{n_i/n} - \sqrt{\frac{1}{6}})^2$.

(c) If the true value of p is .4, then for large n, χ^2_H has approximately a noncentral chi-square distribution with 5 degrees of freedom and noncentrality parameter λ, where we may take

$$\lambda = 4\,n\left[3\left(\sqrt{\frac{2}{15}} - \sqrt{\frac{1}{6}}\right)^2 + 3\left(\sqrt{\frac{1}{5}} - \sqrt{\frac{1}{6}}\right)^2\right] = .04051\,n,$$

or

$$\lambda = 6\,n\left[3\left(\frac{2}{15} - \frac{1}{6}\right)^2 + 3\left(\frac{1}{5} - \frac{1}{6}\right)^2\right] = \frac{n}{25}.$$

4. (a) Let $X_i = 1$ if the ith customer is female and $X_i = 0$ otherwise. The X_i are i.i.d. Bernoulli with probability of success $1-p$. Let $Y_i = \mathrm{I}(X_{i-1} = 0, X_i = 1, X_{i+1} = 0$ for $i = 2, \ldots, n-1$. The Y_i are 2-dependent and stationary, with $\mathrm{E}Y_i = p^2(1-p)$, $\sigma_{00} = p^2(1-p) - (p^2(1-p))^2$, $\sigma_{01} = -(p^2(1-p))^2$, and $\sigma_{02} = p^3(1-p)^2 - (p^2(1-p))^2$. The

number of females seated between two males is $S_n = \sum_2^{n-1} Y_i$. So $\mathrm{E}S_n = (n-2)p^2(1-p)$ and $\mathrm{Var}(S_n) = (n-2)\sigma_{00} + 2(n-3)\sigma_{01} + 2(n-4)\sigma_{02}$.

(b) $\sqrt{n}(S_n/(n-2) - p^2(1-p)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$, where

$$\sigma^2 = \sigma_{00} + 2\sigma_{01} + 2\sigma_{02} = p^2(1-p) + 2p^3(1-p)^2 - 5(p^2(1-p))^2.$$

5. (a) The pth quantile satisfies $F(x_p) = p$, or $1 - (1/x_p) = p$. So $x_p = 1/(1-p)$.

(b) $x_{1/3} = 3/2$ and $x_{2/3} = 3$. The density is $f(x) = 1/x^2$ for $x > 1$. So $f(x_{1/3}) = f(3/2) = 4/9$, and $f(x_{2/3} = f(3) = 1/9$. Thus

$$\sqrt{n}\begin{pmatrix} \hat{x}_{1/3} - 3/2 \\ \hat{x}_{2/3} - 3 \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{2/9}{(4/9)^2} & \frac{1/9}{(4/9)(1/9)} \\ \frac{1/9}{(4/9)(1/9)} & \frac{2/9}{(1/9)^2} \end{pmatrix}\right) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 9/8 & 9/4 \\ 9/4 & 18 \end{pmatrix}\right)$$