Midterm Examination

Statistics 200C

Ferguson — Friday, May 8, 2009

1. Suppose that $X_1, X_2, \ldots$ are independent Bernoulli random variables with $\mathrm{P}(X_n = 1) = p_n$ for $n = 1, 2, \ldots$. Give necessary and sufficient conditions on the p_n for which is it true that

(a) $X_n \xrightarrow{\mathrm{P}} 0$.

(b) $X_n \xrightarrow{a.s.} 0$.

2. Let $X_1, X_2, \ldots$ be independent Bernoulli random variables with $\mathrm{P}(X_j = 1) = p_j$, and let $Z_n = \sum_{i=1}^{n}(X_j - p_j)$. Let $B_n^2 = \mathrm{Var}(Z_n)$

(a) Show that if $B_n \to \infty$, then $Z_n/B_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, by checking the Lindeberg conditions.

(b) Suppose all the p_j are equal but depend on n; in particular, assume $p_j = \dfrac{1}{2} - \dfrac{1}{2\sqrt{n}}$. Find the limiting distribution of $\sqrt{n}(\overline{X}_n - 1/2)$.

3. Consider a multinomial distribution with 4 cells, sample size n and probability vector $((1-p), p(1-p), p^2(1-p), p^3)$ 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 $H_0 : p = .6$? What is its asymptotic distribution under this hypothesis?

(b) What is the approximate large sample distribution of this chi-square if the true value of p is .5?

4. Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli with success probability $p = \mathrm{P}(X_i = 1)$. We wish to estimate p^2.

(a) We may estimate p^2 by $\overline{X}_n^2$. Find the asymptotic distribution of $\overline{X}_n^2$.

(b) Let $Y_j = X_j X_{j+1}$. Since $\mathrm{E}(Y_j) = p^2$, we may also estimate p^2 by $\overline{Y}_n$. Find the asymptotic distribution of $\overline{Y}_n$.

5. Let $S_N = \sum_1^N z_j a(R_j)$ be a linear rank statistic. Let $N = 3n$ and let

$$a(j) = \begin{cases} 1 & \text{for } j = 1, \ldots, n \\ -1 & \text{for } j = n+1, \ldots, 2n \\ 0 & \text{for } j = 2n+1, \ldots, N. \end{cases}$$

and $z_j = j$, for $j = 1, \ldots, N$.

(a) Find $\mathrm{E}(S_N)$ and $\mathrm{Var}(S_N)$.

(b) Find the asymptotic distribution of S_N.

Solutions to Midterm Examination
Statistics 200C

T. Ferguson Fri. May 8, 2009

1. (a) $X_n \xrightarrow{\mathrm{P}} 0$ if and only if $p_n \to 0$ as $n \to \infty$.

(b) From the Borel-Cantelli Lemma, $X_n \xrightarrow{a.s.} 0$ if and only if $\sum_1^\infty p_n < \infty$, since the X_i are independent.

2. (a) $B_n^2 = \mathrm{Var}(Z_n) = \sum_1^n p_j(1-p_j)$. Since $(X_j - p_j)^2 < 1$ for all j,

$$\frac{1}{B_n^2}\sum_{j=1}^n \mathrm{E}[(X_j - p_j)^2 \mathrm{I}((X_j - p_j)^2 > \epsilon^2 B_n^2)] \le \frac{1}{B_n^2}\sum_{j=1}^n \mathrm{E}[(X_j - p_j)^2 \mathrm{I}(1 > \epsilon^2 B_n^2)]$$
$$= \mathrm{I}(1 > \epsilon^2 B_n^2)] \to 0$$

for every $\epsilon > 0$ if $B_n \to \infty$.

(b) Thus we have $Z_n/B_n = n(\overline{X}_n - \bar{p}_n)/B_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. We have $\bar{p}_n = (\frac{1}{2} - \frac{1}{2\sqrt{n}})$ and $B_n^2 = \sum_1^n(\frac{1}{2} + \frac{1}{2\sqrt{n}})(\frac{1}{2} - \frac{1}{2\sqrt{n}}) = n(1-\frac{1}{n})/4 = (n-1)/4$. Since $B_n^2/n \to 1/4$, we have by Slutsky's Theorem, $\sqrt{n}(\overline{X} - \bar{p}_n) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1/4)$. Finally, $\sqrt{n}(\overline{X}_n - 1/2) = \sqrt{n}(\overline{X}_n - \bar{p}_n) - (1/2) \xrightarrow{\mathcal{L}} \mathcal{N}(-1/2, 1/4)$.

3. (a) $\chi_{\mathrm{P}}^2 = n\sum_{j=1}^4 (\hat{p}_j - p_j)^2/p_j$, where $p_1 = .4$, $p_2 = .24$, $p_3 = .144$, $p_4 = .216$, and $\hat{p}_j = n_j/n$ for $j = 1, \ldots, 4$. Under H_0, χ_{P}^2 has asymptotically a chi-square distribution with 3 degrees of freedom.

(b) If the true value of p is .5, then $p_1 = .5$, $p_2 = .25$, $p_3 = .125$ and $p_4 = .125$. Then for large n, χ_{P}^2 has approximately a noncentral chi-square distribution with 3 degrees of freedom and noncentrality parameter λ, where

$$\lambda = n\left[\frac{(.5-.4)^2}{.4} + \frac{(.24-.25)^2}{.24} + \frac{(.125-.144)^2}{.144} + \frac{(.125-.216)^2}{.216}\right].$$

4. (a) From the Central Limit Theorem, $\sqrt{n}(\overline{X}_n - p) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p(1-p))$. Then using Cramér's Theorem with $g(x) = x^2$ and $g'(x) = 2x$, we have $\sqrt{n}(\overline{X}_n^2 - p^2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, g'(p)^2 p(1-p)) = \mathcal{N}(0, 4p^3(1-p))$.

(b) The Y_j are 1-dependent stationary, with $\mathrm{E}(Y_j) = p^2$, $\mathrm{Var}(Y_j) = p^2 - p^4$, and $\mathrm{Cov}(Y_j, Y_{j+1}) = p^3 - p^4$. Therefore, $\sqrt{n}(\overline{Y}_n - p^2) \xrightarrow{\mathcal{L}} \mathcal{C}(0, \sigma^2)$, where $\sigma^2 = p^2 - p^4 + 2p^3 - 2p^4 = p^2 + 2p^3 - 3p^4$.

5. (a) We have $\bar{z}_N = (N+1)/2$ and $\bar{a}_N = 0$, so $\mathrm{E}(S_N) = 0$. We have $\sigma_z^2 = (N+1)(N-1)/12$ and $\sigma_a^2 = 2n/N = 2/3$, so $\mathrm{Var}(S_N) = (N^2/(N-1))((N+1)(N-1)/12)(2/3) = N^2(N+1)/18$.

(b) Since $\max(z_j - \bar{z}_N)^2 = (N-1)^2/4$ and $\max(a_j - \bar{a}_N)^2 = 1$, we have

$$\delta_N = \frac{1}{N} \cdot \frac{(N-1)^2/4}{(N+1)(N-1)/12} \cdot \frac{1}{2/3} \to 0$$

as $N \to \infty$. Therefore,

$$\frac{S_N}{\sqrt{N^2(N+1)/18}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \qquad \text{or} \qquad \frac{S_N}{N^{3/2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1/18)$$