Final Examination

Statistics 200C

T. Ferguson June 9, 2008

1. (a) Give an example of independent two-valued random variables, X_n, such that $X_n \xrightarrow{a.s.} 0$, but not $X_n \xrightarrow{q.m.} 0$.

(b) Give an example of independent Bernoulli random variables, X_n, such that $X_n \xrightarrow{q.m.} 0$ but not $X_n \xrightarrow{a.s.} 0$.

2. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample from a bivariate population such that $\mathrm{E}X = 1$, $\mathrm{Var}(X) = 1$, $\mathrm{E}(Y|X) = \theta X$, and $\mathrm{E}(Y^2|X) = 1 + \theta^2 X^2$, for some number θ.

(a) Show $\mathrm{E}Y = \theta$, $\mathrm{Var}(Y) = 1 + \theta^2$, and $\mathrm{Cov}(X, Y) = \theta$.

(b) Assuming this, what is the asymptotic distribution of $(\overline{X}_n, \overline{Y}_n)$.

(c) Take $\theta = 1$ and find the asymptotic joint distribution of $(\overline{X}_n - \overline{Y}_n, \overline{X}_n^2 + \overline{Y}_n^2)$.

3. A population consists of $N = 3m$ objects of which m have value 1, m have value 0, and m have value -1. From this population a sample of size n is taken and the sum of the sampled values is denoted by S_N.

(a) Show that S_N is a linear rank statistic of the form $S_N = \sum_1^N z_j a(R_j)$ where $a(j) = 1$ if $1 \le j \le n$ and $a(j) = 0$ if $n < j < N$; that is, find z_j for $j = 1, \ldots, N$.

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

(c) Show that $(S_N - \mathrm{E}S_N)/\sqrt{\mathrm{Var}(S_N)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ as $N \to \infty$, provided also that $N/(n(N - n)) \to 0$.

4. Suppose $X_1, \ldots, X_n$ are i.i.d. from a distribution on $(0, x_0)$ with density $f(x)$ continuous at x_0 and $f(x_0) = c$, where $0 < c < \infty$.

(a) Show that $1 - F(x) \sim (x_0 - x)c$ as $x \to x_0$.

(b) Find the asymptotic distribution of $M_n = \max\{X_1, \ldots, X_n\}$.

5. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample from a bivariate population with unknown unknown means, μ_x, μ_y, variances σ_x^2, σ_y^2, and covariance σ_{xy}. Let $\theta = \mu_y/\mu_x$ and $\tilde{\theta}_n = \overline{Y}_n/\overline{X}_n$.

(a) Find the asymptotic distribution of $\tilde{\theta}_n$, assuming $\mu_x \ne 0$.

(b) Find the asymptotic efficiency of $\tilde{\theta}_n$ when the true population distribution has density

$$f(x, y) = \frac{1}{\sqrt{2\pi}} e^{-(y-\theta x)^2/2} e^{-x} \qquad \text{for } x > 0 \text{ and } -\infty < y < \infty$$

where θ is unknown. (Note: the distribution of X is exponential mean 1 and variance 1, and the conditional distribution of Y given $X = x$ is $\mathcal{N}(\theta x, 1)$, so $\theta = \mathrm{E}Y/\mathrm{E}X$. You may use problem 2(a) above.)

6. Suppose $X_1, .., X_n$ is a sample from a Poisson distribution with mean λ (density $e^{-\lambda}\lambda^x/x!$ for $x = 0, 1, \ldots$) and $Y_1, \ldots, Y_n$ is an independent sample from a Poisson distribution with mean μ. We wish to test the hypothesis $H_0 : \mu = 2\lambda$.

(a) Find the likelihood ratio statistic for testing H_0.

(b) Describe how to carry out the test of H_0 for large samples.

7. Consider a multinomial distribution with c cells and probabilities $(p_1(\theta), \ldots, p_c(\theta))$, for $\theta \in \Theta$, some open set on the real line. Assume $p_i'(\theta)$ and $p_i''(\theta)$ exist and are continuous for all $\theta \in \Theta$.

(a) Find Fisher information, $\mathcal{I}(\theta)$. (Take sample size 1, so that $f(n_1, \ldots, n_c|\theta) = \prod_1^c p_i(\theta)^{n_i}$, for exactly one of the n_i equal to one and the rest equal to zero.)

(b) What is the approximate distribution of the minimum chi-square estimate of θ, given a large sample from this distribution? (Assume you answered part (a) correctly.)

8. A sample of size n is taken in a multinomial experiment with IJ cells denoted (i, j), $i = 1, \ldots, I$ and $j = 1, \ldots, J$. Let p_{ij} denote the probability of cell (i, j), and let n_{ij} denote the number falling in cell (i, j), so that $\sum\sum p_{ij} = 1$ and $\sum\sum n_{ij} = n$.

(a) Let H_1 denote the hypothesis of independence, that $p_{ij} = p_i\pi_j$ for some probability vectors, $(p_1, \ldots, p_I)$ and $(\pi_1, \ldots, \pi_J)$. What is the chi-square test of H_1 against all alternatives? How many degrees of freedom does it have?

(b) Let H_0 denote the hypothesis that p_{ij} is independent of j (that is, that $p_{ij} = \theta_i$ for some vector, $(\theta_1, \ldots, \theta_I)$, such that $\sum\theta_i = 1/J$). Find the chi-square test of H_0 against all alternatives. How many degrees of freedom?

(c) What, then, is the chi-square test of H_0 against H_1, and how many degrees of freedom does it have?

Solutions to the Final Examination, Stat 200C, Spring 2008.

1. (a) $X_n = n$ with probability $1/n^2$ and $X_n = 0$ otherwise. Then $\mathrm{E}X_n^2 = 1 \not\to 0$, but $X_n \xrightarrow{a.s.} 0$ since $\sum_1^\infty 1/n^2 < \infty$.

(b) $X_n = 1$ with probability $1/n$ and zero otherwise. Then $\mathrm{E}X_n^2 = 1/n \to 0$ and X_n does not converge to zero alnost surely since $\sum_1^\infty 1/n = \infty$.

2. (a) $\mathrm{E}(Y) = \mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(\theta X) = \theta$, $\mathrm{E}(Y^2) = \mathrm{E}(\mathrm{E}(Y^2|X)) = \mathrm{E}(1+\theta^2X^2) = 1+2\theta^2$, so $\mathrm{Var}(Y) = 1+2\theta^2-\theta^2 = 1+\theta^2$, and $\mathrm{E}(XY) = \mathrm{E}(\mathrm{E}(XY|X)) = \mathrm{E}(X\mathrm{E}(Y|X)) = \mathrm{E}(\theta X^2) = 2\theta$, so $\mathrm{Cov}(X,Y) = 2\theta - \theta = \theta$.

(b) By the CLT $\sqrt{n}((\overline{X}_n, \overline{Y}_n) - (1,\theta)) \xrightarrow{\mathcal{L}} \mathcal{N}((0,0), \begin{pmatrix} 1 & \theta \\ \theta & 1+\theta^2 \end{pmatrix})$.

(c) Make the change of variable $g(x,y) = (x-y, x^2+y^2)$, with $\dot g(x,y) = \begin{pmatrix} 1 & -1 \\ 2x & 2y \end{pmatrix}$, for which $g(1,1) = (0,2)$ and $\dot g(1,1) = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}$. Therefore,

$$\begin{aligned}\sqrt{n}((\overline{X}_n - \overline{Y}_n, \overline{X}_n^2 + \overline{Y}_n^2) - (0,2)) &\xrightarrow{\mathcal{L}} \mathcal{N}((0,0), \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}) \\ &= \mathcal{N}((0,0), \begin{pmatrix} 1 & -2 \\ -2 & 20 \end{pmatrix})\end{aligned}$$

3. (a) We may take $z_j = \begin{cases} -1 & \text{for } 1 \le j \le m \\ 0 & \text{for } m < j \le 2m\,. \\ 1 & \text{for } 2m < j \le N \end{cases}$

(b) $\bar z_N = 0$ and $\bar a_N = n/N$, so $\mathrm{E}S_N = 0$. $\sigma_z^2 = 2m/N = 2/3$ and $\sigma_a^2 = n(N-n)/N^2$, so $\mathrm{Var}(S_N) = (N^2/(N-1))(2/3)(n(N-n)/N^2) = 2n(N-n)/(3(N-1))$.

(c) $\max_j\{(a(j)-\bar a_N)^2\} \le 1$, and $\max_j\{(z_j - \bar z_N)^2\} = 1$, so $(S_N - \mathrm{E}S_N)/\sqrt{\mathrm{Var}(S_N)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ provided

$$\frac{1}{N}\frac{N^2}{n(N-n)}\frac{3}{2} \to 0$$

that is, provided $N/(n(N-n)) \to 0$, or equivalently, both $n \to \infty$ and $N - n \to \infty$.

4. (a) $(1-F(x))/(x-x_0) = (F(x_0) - F(x)/(x_0 - x) \to F'(x_0) = f(x_0) = c$.

(b) $nc(M_n - x_0) \xrightarrow{\mathcal{L}} G_{2,1}$ the exponential distribution on the negative axis.

5. (a) $\sqrt{n}(\overline{X}_n, \overline{Y}_n) - (\mu_x, \mu_y) \xrightarrow{\mathcal{L}} \mathcal{N}((0,0), \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix})$. Let $g(x,y) = y/x$; then $\dot g(x,y) = (-y/x^2, 1/x)$ and $\dot g(\mu_x, \mu_y) = (-\mu_y/\mu_x^2, 1/\mu_x)$. Therefore,

$$\sqrt{n}(\tilde\theta_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \frac{\mu_y^2\sigma_x^2}{\mu_x^4} - 2\frac{\mu_y\sigma_{xy}}{\mu_x^3} + \frac{\sigma_y^2}{\mu_x^2})$$

(b) When $\mu_x = 1$, $\mu_y = \theta$, $\sigma_x^2 = 1$, $\sigma_{xy} = \theta$, and $\sigma_y^2 = 1 + \theta^2$, this reduces to $\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. The best we can do asymptotically is use the MLE with limiting variance $1/\mathcal{I}(\theta)$. Since $\Psi(x,y,\theta) = \frac{d}{d\theta}\log f(x,y|\theta) = -x(y-\theta x)$ and $\mathrm{E}(\dot{\Psi}(X,Y,\theta)) = \mathrm{E}(-X^2) = -2$, Fisher information is $\mathcal{I}(\theta) = 2$. The estimate $\tilde{\theta}_n$ is 50% efficient at this distribution.

6. The likelihood is $L_n = e^{-n\lambda - n\mu}\lambda^{\sum x_i}\mu^{\sum y_i}/\prod x_i!y_i!$.

(a) Under the general hypothesis, the MLE's are $\hat{\lambda}_n = \overline{X}_n$ and $\hat{\mu}_n = \overline{Y}_n$. So $\max_H L_n = e^{-n(\bar{X}_n+\bar{Y}_n)}\overline{X}_n^{n\bar{X}_n}\overline{Y}_n^{n\bar{Y}_n}/\prod x_i!y_i!$. Under H_0, the log likelihood is proportional to $-3n\lambda + (\sum X_i + \sum Y_i)\log(\lambda)$. Setting the derivative to zero gives as the MLE, $\tilde{\lambda} = (\overline{X}_n + \overline{Y}_n)/3$. So $\max_{H_0} L_n = e^{-n(\bar{X}_n+\bar{Y}_n)}((\overline{X}_n + \overline{Y}_n)/3)^{n(\bar{X}_n+\bar{Y}_n)}2^{n\bar{Y}_n}/\prod x_i!y_i!$. The likelihood ratio is

$$\Lambda_n = \frac{((\overline{X}_n + \overline{Y}_n)/3)^{n(\bar{X}_n+\bar{Y}_n)}2^{n\bar{Y}_n}}{\overline{X}_n^{n\bar{X}_n}\overline{Y}_n^{n\bar{Y}_n}}$$

(b) We reject H_0 if $-2\log\Lambda_n$ is too large for the chi-square distribution with 1 d.f.

7. (a) $\Psi(\boldsymbol{n},\theta) = \frac{\partial}{\partial\theta}\log f(\boldsymbol{n}|\theta) = \sum_1^c n_i p_i'(\theta)/p_i(\theta)$, and $\dot{\Psi}(\boldsymbol{n},\theta) = \sum_1^c n_i[p_i''(\theta)/p_i(\theta) - p_i'(\theta)^2/p_i(\theta)^2]$. Using $\mathrm{E}(n_i) = p_i(\theta)$, we find

$$\begin{aligned}\mathcal{I}(\theta) = -\mathrm{E}\dot{\Psi}(\boldsymbol{n},\theta) &= -\sum_1^c \mathrm{E}(n_i)[p_i''(\theta)/p_i(\theta) - p_i'(\theta)^2/p_i(\theta)^2] \\ &= -\sum_1^c p_i''(\theta) + \sum_1^c p_i'(\theta)^2/p_i(\theta) = \sum_1^c p_i'(\theta)^2/p_i(\theta)\end{aligned}$$

where $\sum_1^c p_i''(\theta) = 0$ because $\sum_1^c p_i(\theta) = 1$ for all θ.

(b) Since minimum chi-square estimates are asymptotically efficient, we have for the minimum chi-square estimate, $\tilde{\theta}_n$,

$$\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (\sum_1^c p_i'(\theta)^2/p_i(\theta))^{-1}).$$

8. (a) This is the chi-square test of independence. We reject H_1 if $\chi^2(H_1) = \sum_i\sum_j(n_{ij} - \hat{n}_{ij})^2/\hat{n}_{ij}$ is too large for a chi-square distribution with $(I-1)(J-1)$ d.f., where $\hat{n}_{ij} = n_{i.}n_{.j}/n_{..}$.

(b) The likelihood function is proportional to $\prod_{ij}\theta_i^{n_{ij}} = \prod_i \theta_i^{n_{i.}}$. Since the sum of the θ_i is $1/J$, we have $\hat{\theta}_i = n_{i.}/(Jn_{..})$ as the MLE's. We reject H_0 if $\chi^2(H_0) = \sum_i\sum_j(n_{ij} - n_{..}\hat{\theta}_i)^2/(n_{..}\hat{\theta}_i)$ is too large for a chi-square distribution with $I(J-1)$ d.f.

(c)We reject H_0 against H_1 if $\chi^2(H_0) - \chi^2(H_1)$ is too large for a chi-square distribution with $J-1$ d.f.