original

STATISTICS PH.D. COMPREHENSIVE EXAM

August 19, 2005 2-5 PM

Directions: The exam consists of seven questions of equal point value. Make sure to write your ID number (last 4 digits of your Social Security number) on your answer sheets. Do not put your name on any of your answer sheets. All solutions should be complete and rigorous. If you cannot solve a problem, try to outline an approach that may lead to a solution.

1) In a linear model $Y = X\beta + e$, $E(e) = 0$, define what it means to be a least squares estimate of β. How do you actually find least squares estimates? Prove that your answer is correct.

2) Let $y_0, y_1, y_2, \ldots, y_n$ be independent $N(x_i'\beta, \sigma^2)$ random variables. y_0 is unobserved but think of a regression model for $y_1, \ldots, y_n$, say, $Y = X\beta + e$.

a) Show that

$$\frac{y_0 - x_0'\hat{\beta}}{\sqrt{MSE\,[1 + x_0'(X'X)^{-1}x_0]}}$$

has a t distribution and give the degrees of freedom. Here, $\hat{\beta}$ and MSE are the usual least squares based estimates from the regression model for $y_1, y_2, \ldots, y_n$ suggested in the problem description.

b) Give a $1 - \alpha$ prediction interval for y_0 and justify your answer.

3. Let $X_1, X_2, \ldots$ be independent identically distributed (i.i.d.) random variables. Formulate the Strong Law of Large Numbers (necessary and sufficient conditions). Give an example of random variables $\{X_n\}$ for which SLLN does not hold.

4. Let $\nu, X_1, X_2, \ldots$ be independent random variables. Suppose that $X_1, X_2, \ldots$ are identically distributed, $\mathrm{E}|X_1|^2 < +\infty$, and ν takes values in $\{1, 2, \ldots\}$ with $\mathrm{E}|\nu|^2 < +\infty$. Define

$$S = X_1 + \ldots + X_\nu.$$

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

5. State and prove Basu's Theorem. (Carefully define all statistical notation.)

6. Let $X_1, X_2, ..., X_n$ be independent random variables with a $N(\theta, \theta^2)$ distribution (here $\mathrm{Var}(X_i) = \theta^2$), where $\theta > 0$. For this family of distributions, $T = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is a sufficient statistic.

a) Let $\bar{X} = \sum_{i=1}^n X_i/n$ and $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2$. Show that $V = \bar{X}^2/S^2$ is ancillary.

b) Is V independent of T? Prove your claim. If V and T are not independent, then why is Basu's theorem not applicable here?

7. Let $Y_1, Y_2, ..., Y_n$ be independent random variables with a truncated Poisson(θ) distrution. That is, for $\theta > 0$

$$\Pr(Y_i = k) = \frac{e^{-\theta}\theta^k}{(1 - e^{-\theta})k!} \qquad k = 1, 2, \ldots.$$

Note that if X_i is Poisson(θ), then Y_i has the same distribution as X_i given that $X_i > 0$.

a) Let $c(\theta) = \theta/(1 - e^{-\theta})$. Show that

$$\mathrm{E}(Y_i) = c(\theta)$$

and that $c(\theta)$ is an increasing function for $\theta > 0$ with $\lim_{\theta \to 0^+} c(\theta) = 1$ and $\lim_{\theta \to \infty} c(\theta) = \infty$.

b) Let $\bar{Y} = \sum_i Y_i/n$. Show that the maximum likelihood estimate (MLE) of θ is the unique solution to the equation $c(\theta) = \bar{Y}$. In particular, show that the MLE satisfies $\hat{\theta} = c^{-1}(\bar{Y})$.

c) Writing the MLE as $\hat{\theta} = c^{-1}(\bar{Y})$, show directly (i.e. without applying limit theorems for MLEs) that $\hat{\theta} \to \theta$ almost surely as $n \to \infty$.

d) Find the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$ as $n \to \infty$.