M. S. Comprehensive Exam (Part I of Written Exam)
Thursday, January 2, 1997

You have three hours. Do as many problems as you can. No one is expected to answer all the problems correctly. Partial credit will be given. All problems are worth an equal amount of credit.

Put your solution to each problem on a separate sheet of paper.

Problem 1. Let X be a non-negative random variable with density function $f(x)$ and distribution function $F(x) = P(X \leq x)$. The survival function $S(x)$ is $S(x) = 1 - F(x) = P(X > x)$. The failure rate function is

$$\lambda(x) = f(x)/S(x).$$

Show that $S(x) = \exp\{-\int_0^x \lambda(u)du\}$.

Problem 2. Let $\{X_1, X_2, \ldots, X_n\}$ be mutually independent random variables with

$$P(X_k \leq t) = \begin{cases} 0, & \text{if } t < 0, \\ t^k, & \text{if } 0 \leq t < 1, \\ 1, & \text{if } 1 \leq t \end{cases} \qquad k = 1, 2, \ldots, n$$

(a) Let $Y = \max(X_1, X_2, \ldots, X_n)$. Find $\mathrm{E}Y$.

(b) Let $Z = \prod_{i=1}^n X_i$. Find $\mathrm{E}Z$.

Problem 3. While rolling a balanced die successively, the first 6 occurred on the third roll. Conditional on this, what is the expected number of rolls until the first 1 appears?

Problem 4. Let $\{X_1, X_2, X_3, X_4, X_5\}$ be a random sample from an exponential distribution having density function

$$f(x|\theta) = \begin{cases} \theta \exp\{-\theta x\}, & \text{for } x > 0,\ \theta > 0, \\ 0, & \text{Otherwise.} \end{cases}$$

Define the order statistics $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq X_{(4)} \leq X_{(5)}$.

(a) Derive the density function of $X_{(3)}$, the median.

(b) Derive $\mathrm{E}(X_{(3)})$.

Problem 5. Suppose $X_1, \ldots, X_n$ are i.i.d. and have a uniform distribution on the interval $[0, \theta]$, $\theta > 0$.

(a) Find the probability density function of $\max_i X_i$.

(b) Show that $n(\theta - \max_i X_i)$ converges in distribution to an exponential random variable with mean θ.

Problem 6. Let $X_1, \ldots, X_n$ be independent and identically distributed Bernoulli random variables with parameter $\theta \in (0, 1)$.

(a) Show that $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is an unbiased estimation of $\mathrm{Var}(X_1)$ for $n \geq 2$.

(b) Show that, for $n = 1$, there does not exist any unbiased estimator for $\mathrm{Var}(X_1)$.

Problem 7. An attorney has called you, saying:

My client is contesting the results of last November's election which he lost by a narrow margin to the incumbent. In this election contest, 27,000 absentee ballots were cast. The official count showed that 16,132 were cast for the incumbent.

However, not all the absentee voters used a No. 2 pencil to darken the circles on the mark sense form on which they recorded their vote. The Supervisor of Elections went through the ballots and darkened with pencil those that appeared to be filled in with pen.

After the election results were announced, my assistant and I went through the 27,000 absentee ballots and identified all those that appeared to have been darkened by the Supervisor of Elections. We found a sample of 6498 such ballots, 3996 of which were marked for the incumbent. Is this tendency for more of the darkened ballots to be for the incumbent surprising?

How would you answer the attorney's question?

Problem 8. In the standard regression model

$$Y = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where X is an $n \times p$ matrix with rank p, the first column of X is $(1, 1, \ldots, 1)'$, and $\text{Cov}(\epsilon) = \sigma^2 I_n$.

(a) Let $\hat{\boldsymbol{\beta}}$ be the least squares estimate of $\boldsymbol{\beta}$ and $\hat{Y} = X\hat{\boldsymbol{\beta}}$. Show that $\sum_{i=1}^n \text{Var}(\hat{y}_i) = p\sigma^2$.

(b) Show that $\text{Var}(\sum_{i=1}^n \hat{y}_i) = n\sigma^2$.

Problem 9. A two-way factorial experiment has been done. Factor A has three levels, and Factor B has four levels. There are two replications, that is, there are two observations in each of the 12 cells. Factor A is a fixed effect and Factor B is a random effect. Denote the variance of the experiment error by σ^2, the variance of the random effect for Factor B by σ_b^2, and the variance of the interaction effect by σ_{ab}^2. The expected mean squares are $\text{E}[MSA] = \sigma^2 + 2\sigma_{ab}^2 + c$, $\text{E}[MSB] = \sigma^2 + 6\sigma_b^2$, $\text{E}[MS(AB)] = \sigma^2 + 2\sigma_{ab}^2$, and $\text{E}[MSE] = \sigma^2$.

(a) Construct the appropriate ANOVA table, giving the sources of variation, degrees of freedom.

(b) How would you test the hypothesis that Factor A has no effect? Factor B? Why are the denominators of the test statistics for these two tests different?

(c) Suppose you did the F-statistic for the significance of Factor A using a computer package and that the output said "p-value = .0003. " Interpret this output.

(d) Suppose you did the F-statistic for the significance of Factor B using a computer package and that the output said "p-value = .0000. " Interpret this output.