M. S. Comprehensive Exam (Part I of Written Exam)
Monday, January 30, 1995

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. Consider a data set containing 300 observations on $Y, X_1, X_2, \ldots, X_{40}$. The X's are explanatory variables to be used in explaining variation in Y.

(a) Describe a procedure for identifying good multiple regression equations here. Include in your description comments on the criterion or criteria you are using, what computer software you would use at various stages in this process, and how you would choose among the millions of potential regression equations here.

(b) Discuss briefly the consequences of having too few or too many variables in the final model you recommend.

Problem 2. A competition is held between two boats, A and B. The competition involves two rounds. The first round consists of six races. If one of the boats has won more races than the other at the end of the first round, then that boat is declared winner of the competition. Otherwise a second round is held in which races are run until one of the boats has won two *consecutive* races. Suppose that races are independent of one another and boat A wins a race with probability 1/2, boat B wins with probability 1/2. Let X denote the number of races in the competition. Find the probability mass function of X.

Problem 3. Tom Bayes has purchased a coin from a novelty shop and wishes to estimate the probability p of getting a head when this coin is tossed. Tom suspects the coin is biased, but knows nothing about the direction of the bias. He decides to use a mixture of Beta distributions as his prior distribution for p, with density $\pi(p) = 2(p^3 + (1-p)^3)$, $0 < p < 1$. He then tosses the coin 4 times and observes 2 heads and 2 tails. Find Tom's posterior distribution for p and compute the mean and variance of this posterior distribution.

Problem 4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with density

$$f(x|\theta) = \frac{\theta}{(1+x)^{\theta+1}} \quad \text{for } 0 < x < \infty,\ \theta > 0\,.$$

(a) Find the MLE $\hat{\theta}$ of θ.

(b) Show that $\hat{\theta}$ is a complete sufficient statistic for θ.

(c) Is $\hat{\theta}$ unbiased?

Problem 5. (An $F_{\alpha=0.05}$ table is needed for this problem) A pharmaceutical company would like to examine the potency of a liquid medication mixed in large vats. To do this, a random sample of three vats from a month's production was obtained and five separate samples were selected from each vat. The results are listed in the following table:

Vat 1	3.2	3.8	3.5	3.0	3.4
Vat 2	2.6	2.9	2.8	2.0	2.5
Vat 3	3.4	3.9	3.3	3.1	3.6

A random-effects model for this experiment has the form:

$$Y_{ij} \;=\; \mu + v_i + \epsilon_{ij}, \qquad i = 1, 2, 3;\;\; j = 1, \ldots, 5,$$

where μ is the overall mean, the v_i's are iid $N(0, \sigma_v^2)$ random variables and the ϵ_{ij}'s are iid $N(0, \sigma_\epsilon^2)$ variables. Furthermore, $\{v_i\}$ and $\{\epsilon_{ij}\}$ are assumed to be independent.

(a) Given $\bar{Y}_{1\cdot} = 3.38$, $\bar{Y}_{2\cdot} = 2.56$, $\bar{Y}_{3\cdot} = 3.46$ and $\bar{Y}_{\cdot\cdot} = 3.13$. Test the null hypothesis $H_0 : \sigma_v^2 = 0$ vs the alternative hypothesis $H_a : \sigma_v^2 > 0$ by using $\alpha = 0.05$.

(b) Let MSA be the mean squares among groups and MSW the mean squares within groups. Show that $\hat{\sigma}_\epsilon^2 = MSW$ and $\hat{\sigma}_v^2 = (MSA - MSW)/5$ are unbiased estimates for σ_ϵ^2 and σ_v^2, respectively.

Problem 6. Let X and Y be i.i.d. integer valued random variables. Show that

$$P(X + Y \text{ is even}) \geq 1/2.$$

Problem 7. Let X and Y be nonnegative random variables such that $XY \geq 1$ (a.s.) Show that $EX \cdot EY \geq 1$.