M. S. Comprehensive Exam (Part I of Written Exam)
Sunday, January 2, 2000

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. The number, X, of defects per roll on wallpaper produced by a company has a Poisson distribution with mean λ, a positive number. The mean λ, however, changes from day to day and has an exponential distribution with mean equal to 2.

(a) Find the unconditional distribution of X, the number of defects per roll.

(b) If a roll has one defect, what is the probability that it was produced on a day for which $1 < \lambda < 3$?

Problem 2. An administrator in a governmental health care agency in which you want to work is interviewing you. Answer the administrator's questions (which are given below) in everyday English comprehensible to your interviewer who is very bright and has not taken any statistics courses.

> I have been reading some background documents on the agency's clinical trials policy. The documents state that a drug can only be considered for licensing if the p-value is less than 0.05.

(a) What is a "p-value"? Please do not be abstract with me. Pick a concrete example and explain p-values to me relative to this example.

(b) Why 0.05? What is the logic behind this number – why not 0.06 or 0.04?

(c) The documents also talk about type 1 and type 2 errors. This is most worrying. Why are they making errors in the trials and not correcting for them? The Government and drug companies could be in very serious trouble here.

(d) These documents also discuss significance levels, power, and sample size. How do these terms relate to each other and to p-values? Again, please be concrete and relate all this to the example you used in part (a).

Problem 3. A physician was interested in comparing the effects of four different antihistamines (A, B, C, D) in persons extremely sensitive to a ragweed skin allergy test. Four allergy patients were selected from the physician's private practice, with treatments (antihistamines) assigned to each patients according to a Balanced Incomplete Block (BIB) design. Each person then received injections of the assigned antihistamines in different sections of the right arm. The area of redness surrounding the point of injection was measured after a fixed period of time. The data are shown in the following table:

	Treatment			
Person	A	B	C	D
1	46	30	-	40
2	32	25	14	-
3	47	-	27	44
4	-	23	16	35

(a) In general, let t be the number of treatments, k the number of treatments per block, and b the number of blocks. Let Y_{ij} be the response variable in this experiment. Write an additive model for this experiment and state your assumptions.

(b) The sum of squares for treatments adjusted for blocks is

$$SST(Adj) = \frac{t-1}{nk(k-1)} \sum_i (kY_{i.} - B_{(i)})^2$$

where n is the total number of observations, $Y_{i.}$ is the sum of all observations on treatment i, and $B_{(i)}$ is the sum of all measurements for blocks that contain treatment i. Conduct an analysis of variance for this experiment.

Problem 4. Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n \geq 1$ from an exponential distribution with mean $\beta > 0$.

(a) Find the likelihood ratio test of $H_0 : \beta \leq \beta_0$ against $H_1 : \beta > \beta_0$ at significance level α. (β_0 is a known positive constant.)

(b) Is your test uniformly most powerful at the α level? Explain.

(c) Make an appropriate transformation on your test statistic so that a chi-square table can be used to perform the test.

Problem 5. Explain and justify mathematically the following observations:

Students who score high on the first test will, as a group, score closer to the average (lower) on the second test, while students who score low on the first test will, as a group, also score closer to the average (higher) on the second test.

Problem 6. Let $X_1, X_2, \ldots, X_n$ be independent with $X_i \sim \text{Poisson}(\lambda t_i)$ where $\lambda > 0$ is unknown and $t_1, t_2, \ldots, t_n$ are known positive constants.

(a) Find the maximum likelihood estimator of λ based on $X_1, \ldots, X_n$.

(b) Show that the estimator you found in part (a) is also the minimum variance unbiased estimate of λ.

Problem 7. Suppose $EX = 0$ and $\text{Var}(X) = \sigma^2$. Prove

$$P(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}, \quad a > 0$$

Hint: consider the random variable $X + (\sigma^2/a)$ and use the Markov inequality.

Problem 8. A deck of cards is shuffled and then divided into two halves of 26 cards each. A card is drawn from one of the halves; it turns out to be an ace. The ace is then placed in the second half-deck. This half is then shuffled and a card is drawn from it. Compute the probability that this drawn card is an ace.

Problem 9. Let $X_1, X_2, \ldots, X_n$ be i.i.d. with density

$$f(x|\theta) = \frac{1}{\pi\,(1 + (x - \theta)^2)}.$$

(a) What is the Cramér-Rao lower bound for the variance of an unbiased estimator of θ?

(b) Is there an estimator which exactly achieves this bound for finite values of n? Explain.

Hint for part (a): You may use (without proof) that

$$\int_{-\infty}^{+\infty} \frac{u^{2j}}{(1+u^2)^{j+k}}\, du = \frac{\Gamma(j+\frac{1}{2})\Gamma(k-\frac{1}{2})}{\Gamma(j+k)} \quad \text{and } \Gamma(1/2) = \sqrt{\pi}\,.$$