1998 Honors Examination in Statistics
(for students having taken Math 41 and Stat 111)

Department of Mathematics and Statistics
Swarthmore College

Instructions: This exam consists of a total of four questions. Number the questions clearly in your work and start each question on a new page. You must show your work to make it clear how you obtained your answers. Answers without any work may lose credit even if they are correct, and will receive no credit if incorrect.

This is a closed-book three-hour exam. You may not refer to your notes or textbooks.

1. You plan in the near future to eat uncooked scallops, raw pork, and Hudson Foods hamburgers for dinners. The probability of becoming violently ill from the scallops, from the pork, and from the hamburgers is 1/3, 1/2, and 3/4, respectively. Assume violent illness on different meals occurs independently.

   (a) Suppose you have each of the three dinners on three consecutive days, eating the scallops on the first night, the pork on the second night, and the hamburgers on the third night. Let X be the number of days (out of the three) in which you become violently ill. Why does X not have a binomial distribution (please do not carry out any calculations for your answer)?

   (b) Determine the probability distribution of X.

   (c) If you choose to eat the uncooked scallops night after night, what is the expected number of days until you become violently ill from the scallops?

   (d) Suppose you decide to eat the same dinner on two consecutive days. Also suppose the probability of eating the scallops on both days is 1/2, of eating the pork is 1/4, and of eating the hamburgers is 1/4. If you become violently ill on both days, what is the probability you ate the scallops?

2. In this day and age, it's difficult enough for a man to find the woman of his dreams, but it's that much more difficult if he is overweight. At the "National Heffer Dating Service," potential dates are chosen in an unusual manner. Two fair coins labeled A and B are flipped. Let X denote the total number of heads appearing, and let $Y = 1$ if heads appears on coin A, and $Y = 0$ otherwise. The choice of a date will depend on the coin flips, as described below.

   (a) Display the joint probability mass function of X and Y in a suitable table, and determine the marginal distributions of X and Y.

   (b) Determine $\text{Cov}(X, Y)$.

   (c) Let W be the weight (in pounds) of the potential date. If both $X = 1$ and $Y = 1$, then a date is chosen such that $W \sim \text{N}(175, 25^2)$. Otherwise, a date is chosen such that $W \sim \text{N}(230, 15^2)$. Write down an expression for $\text{P}(W > 200)$, the (unconditional) probability that the man is set up with a woman who weighs over 200 pounds, in terms of unconditional probabilities involving a standard normal random variable, $Z \sim \text{N}(0, 1)$.

   (d) Let p denote the probability of being set up with a woman weighing over 200 pounds. In terms of p, what is the probability that, on the 8th visit to the dating service, a man is set up with a fifth woman weighing over 200 pounds?

3. A random variable X has a probability mass function

$$f(x;\theta) = \begin{cases} \frac{1}{4}(1-\theta) & \text{if } x = 0, 1, 2, 3 \\ \theta & \text{if } x = 4 \\ 0 & \text{otherwise,} \end{cases}$$

where θ, $0 \le \theta \le 1$, is an unknown parameter. We wish to estimate θ based on a sample of $n = 50$ independent observations $X_1, X_2, \ldots, X_{50}$.

(a) Let

$$Y = \text{number of } X_i \text{ out of 50 that are equal to 4.}$$

What is the exact distribution of Y, given θ?

(b) Show that $\hat{\theta}_1 = Y/50$ is an unbiased estimator of θ.

(c) Let

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_{50}}{50}$$

be the sample mean of the 50 observations. Find numbers a and b so that the estimator $\hat{\theta}_2 = a\bar{X} + b$ is an unbiased estimator of θ.

(d) Find the variances of $\hat{\theta}_1$ and $\hat{\theta}_2$. Based on the variances of these unbiased estimators of θ, which estimator would you recommend using? Briefly justify.

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

$$f(x;\theta) = \begin{cases} (\theta+1)x^\theta & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

for $\theta > -1$.

(a) Derive the method-of-moments estimator of θ.

(b) Sometimes method-of-moments estimators can produce nonsense estimates, depending on the observed data (for example, in normal models, method-of-moment estimates of variances are sometimes negative). Does the estimator you derived in part (a) ever produce nonsense estimates of θ? If so, give an example of data coming from the above distribution that would produce a nonsense estimate. If not, show why not.

(c) Derive the maximum likelihood estimator of θ.

(d) Determine the minimum variance of an unbiased estimator of θ given by the Cramér-Rao lower bound.