Master's Exam - Spring 2005

March 17, 2005
1:00 – 5:00 pm

NAME:____________________________________

A. The number of points for each problem is given.

B. There are 12 problems with varying numbers of parts.

1 – 6	Probability (97 points)
7 – 12	Statistics (115 points)

C. Write your answers on the exam paper itself. If you need more room you may use the extra sheets provided. Answer as many questions as you can on each part.

D. Professors Sakhanenko and Stapleton will be in their offices (441 and 429) during most of the exam, with at least one being available at all times. One of them will come to the exam room every 15 minutes to see if there are any questions.

PROBABILITY PART.

1. For each of the following experiments and events determine the probability of the event. Assume all the dice are equally likely to result in any of the numbers 1, 2, ..., 6 and are thrown independently.

(a) (5 pts.) *Experiment:* Five dice are thrown. *Event:* At least one of the five dice is a 6.

(b) (6 pts.) *Experiment:* Five dice are thrown. *Event:* At least two 6's occur.

(c) (7 pts.) *Experiment:* Six hundred dice are thrown. *Event:* At least 120 of the tosses result in 6.

(d) (8 pts.) *Experiment:* Six hundred dice are thrown. *Event:* The total of all the numbers obtained is less than 2000. Hint: The variance of the uniform distribution on 1, 2, ..., 6 is $(36 - 1)/12$.

(e) (5 pts.) *Experiment:* Two dice are thrown until 6-6 occurs. *Event:* The first 6-6 occurs on throw number 5.

(f) (5 pts.) *Experiment:* Two dice are thrown 100 times. *Event:* Fewer than two 6-6's occur. Use an approximation.

2. (7 pts.) A test for the HIV virus has probability 0.99 of being positive if the person really has the virus. It has probability 0.005 of being positive if the person does not have the virus. Among all people taking the test 0.02% (1/50 of 1%) actually have the virus. If one of these people is chosen randomly and the test is positive, what is the conditional probability that the person has the virus?

3. Let $f(x) = x/2$ for $0 \le x \le 2$ and 0 elsewhere.

(a) (7 pts.) Let X have density f, and let $Y = 1/X$. Find the density of Y.

(b) (6 pts.) Let X_1 and X_2 be independent, each with the density f. Find $P(X_1 + X_2 > 1)$.

(c) (7 pts.) Let X_1, X_2, and X_3 be independent, each with the density f. Find the covariance $\text{cov}(X_1 + X_2, X_1 + X_2 + X_3)$ and the correlation coefficient $\rho(X_1 + X_2, X_1 + X_2 + X_3)$.

4. (10 pts.) Let X_1, X_2, X_3 be independent discrete random variables such that

x	$P(X_1 = x)$	$P(X_2 = x)$	$P(X_3 = x)$
1	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{3}{5}$
2	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
3	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$

Find the cumulative distribution function and probability mass function of $Y = \max(X_1, X_2, X_3)$.

5. (10 pts.) Consider the following gambling game. Each time you play the game you must pay \$19. From a box containing two \$1 bills, one \$10 bill and one \$20 bill, you must draw 2 bills randomly without replacement. You then keep the bills you draw. Suppose you play this game independently 100 times. What is the probability that you will have lost more than \$200?

6. Let X, Y be random variables with the joint density $f(x, y) = 5y + 2xy,\ 0 \le y \le x^2, -1 \le x \le 1$, and 0 elsewhere.

(a) (7 pts.) Find the density of X.

(b) (10 pts.) Find $Z = \mathbf{E}(Y|X)$.

(c) (7 pts.) Given a random variable U, uniformly distributed on $[0, 1]$, how could a random variable Z be generated from U?

(d) (10 pts.) Find the density of Y. Suggestion: Check that it is a density.

STATISTICS PART.

7. An experiment has probability p of succeeding, for $0 \le p \le 1$, unknown. The experiment is performed independently until it succeeds. Let X_1 be the number of experiments necessary. Then the experiment is repeated until a success has occurred. Let X_2 be the number of experiments needed this time.

(a) (10 pts.) Find the maximum likelihood estimator of p.

(b) (8 pts.) Find the method of moments estimator of p.

8. In order to determine the effect of an additive to the gasoline designed to increase miles per gallon in automobiles, 14 autos were chosen for an experiment. These were paired according to the size and age of the autos, those in the same pair being similar. One of the two in each pair was chosen randomly to receive the additive. Then each of the 14 autos was driven over a 650 kilometer course and the number of liters used was determined. These were:

Pair	1	2	3	4	5	6	7
With Additive	72	83	83	67	82	70	94
Without Additive	80	88	92	64	83	76	96

(a) (12 pts.) State a parametric model, define hypotheses, and carry out a test at level $\alpha = 0.05$ which will enable you to decide whether the additive reduces gasoline consumption.

(b) (8 pts.) Perform a nonparametric test of the same hypotheses. Find the p-value for this test.

9. In order to estimate the change $\Delta = p_1 - p_2$ in the proportion of a population of 1 million voters who preferred Kerry to Bush between time t_1 and time t_2, random samples of sizes $n_1 = 1000$ and $n_2 = 800$ were taken independently at the two times. Let X_1 and X_2 be the numbers favoring Kerry for the two samples.

(a) (8 pts.) Define notation and give a formula for a 95% confidence interval on Δ.

(b) (5 pts.) Apply the formula for the case $X_1 = 431, X_2 = 312$.

(c) (7 pts.) Explain the meaning of your interval in such a way that someone who had not ever studied statistics would understand.

(d) (8 pts.) If equal sample sizes, $n = n_1 = n_2$ were to be used how large must n be in order that the estimator $\hat{\Delta}$ have probability at least 0.90 of being within 0.03 of Δ?

10. (10 pts.) Let $X_1, ..., X_n$ be independent random variables, each with mean μ and variance σ^2. Consider the two estimators $\mu^* = (\sum_{i=1}^n iX_i)/(\sum_{i=1}^n i)$ and $\bar{X} = (\sum_{i=1}^n X_i)/n$.

(a) Show that μ^* is an unbiased estimator of μ.

(b) Determine the relative efficiency (relative size of variances) of μ^* to $\bar{X}$ and find its limiting value as $n \to \infty$. Hint: $\sum_{i=1}^n i = n(n+1)/2$ and $\sum_{i=1}^n i^2 = n(n+1)(2n+1)/6$.

11. Let (X_1, X_2) be a random sample from the Poisson distribution with unknown parameter $\lambda > 0$. Suppose that we wish to test $H_0 : \lambda = 3$ vs $H_a : \lambda < 3$.

(a) (5 pts.) Consider the test which rejects for $T = X_1 + X_2 < 2$. What is the level of significance α?

(b) (10 pts.) Use a theorem to prove that this test is uniformly most powerful for this α level.

(c) (6 pts.) Find the power of this test for $\lambda = 2$.

12. Let (x_i, Y_i) be observed for $i = 1, ..., n$. Suppose that the x_i are constants, and that $Y_i = \beta x_i + \varepsilon_i$, where β is an unknown parameter, and the ε_i are independent, each with the $N(0, \sigma^2)$ distribution.

(a) (8 pts.) Show that the least squares estimator of β is $\hat{\beta} = (\sum_{i=1}^n x_i Y_i)/(\sum_{i=1}^n x_i^2)$. Do **not** use matrix or vector space methods.

(b) (5 pts.) Show that $\hat{\beta}$ is an unbiased estimator of β.

(c) (5 pts.) Find $\text{Var}(\hat{\beta})$.

(d) (8 pts.) For the following (x_i, Y_i) pairs find a 95% confidence interval on β: (1, 2), (2, 5), (3, 10).