Master's Exam - Spring 2009

March 19, 2009
1:00 pm − 5:00 pm

NUMBER:———————————————————

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

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

1 − 5	Probability (132 points)
6 − 12	Statistics (134 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. Tables are provided.

1. A closet contains 13 different pairs of gloves. Twenty six gloves are randomly arranged into 13 pairs.

(a) (7 pts.) Find the probability that all left gloves are paired with right gloves.

(b) (7 pts.) Find the probability that the gloves are paired in the original order.

2. Suppose there are two boxes. The first box has 2 red balls and 18 white balls and the second box has 1 red ball and 19 white balls. The following experiment is performed. An unbiased coin is tossed. If head appears, a ball is drawn randomly from the first box and if tail appears, a ball is drawn randomly from the second box.

(a) (3 pts.) What is the chance that a red ball is drawn?

(b) (3 pts.) If a white ball is drawn, what is the chance that head appeared in the toss?

(c) (6 pts.) If the above experiment is repeated 40 times independently, what is the approximate probability that at most 2 red balls are drawn?

(d) (6 pts.) If the above experiment is repeated 4000 times independently, what is the approximate probability that at least 295 red balls are drawn?

.

3. (a) (3 pts.) Let X_n have the cdf F_n for $n = 1, 2, ...$ Let X have the cdf F. Define: X_n converges in distribution to X as $n \to \infty$.

(b) (3 pts.) Define: Y_n converges in probability to 5 as $n \to \infty$.

(c) (10 pts.) Let $X_1, ..., X_n, ...$ be independent random variables and X_n has density $f_n(x) = (1 + x/n)^{n+1}$ for $-n < x < 0$ and $f_n(x) = 0$ elsewhere. Show that X_n converges in distribution to some X as $n \to \infty$. Find the distribution of X.

4. Consider random variables X and Y, jointly distributed with the density function

$$f(x,y) = \begin{cases} 2y & , \ 0 \le x \le 1, x \le y \le 2x \\ 0 & , \ \text{otherwise.} \end{cases}$$

(a) (**7 pts.**) Find the marginal density of X. (Hint: graph the area of integration.)

(b) (**7 pts.**) Find the marginal density of Y. (Hint: graph the area of integration.)

(c) (**10 pts.**) Find the conditional density of Y given $X = x$, such that $0 < x < 1$.

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

(e) (**5 pts.**) Find $V = \mathbf{E}(\frac{Y}{1+X} + X|X)$.

5. Let X_1, X_2, X_3 be independent random variables with cdfs $F_1(x) = x, F_2(x) = x^2, F_3(x) = x^3$ for $0 \leq x \leq 1$ correspondingly. $F_i(x) = 0, i = 1, 2, 3$, for $x < 0$ and $F_i(x) = 1, i = 1, 2, 3$, for $x > 1$.

(a) (5 pts.) Given a random variable U, uniformly distributed on $[0, 1]$, how could such a random variable X_3 be generated from U?

(b) (10 pts.) Let $M = \max(X_1, X_2, X_3)$. Find the cdf and the density of M.

(c) (10 pts.) Let $S = X_2 - X_1$. Find the cdf of S. (Hint: graph the area of integration).

(d) (8 pts.) Find the cdf and the density for $Y = (X_3 - 1/3)^2$.

(e) (6 pts.) Find the correlation coefficient $\rho(3X_1 - 7X_2 + 2X_3, -4X_1 + 5X_2 + 9X_3)$.

(f) (6 pts.) Find $E(W)$ for

$$W = \frac{X_3^2 \sqrt{X_1}}{2 + X_2}$$

.

6. Let σ be a unknown parameter. Let $(X_1, \cdots, X_n)$ be a random sample from the density

$$f(x) = \frac{1}{\sigma} e^{1-x/\sigma}, x > \sigma.$$

(a) (7 pts.) Find the method of moments estimator of σ.

(b) (5 pts.) Find the likelihood function of σ. Graph it.

(c) (5 pts.) Find the maximum likelihood estimator (mle) of σ.

(d) (10 pts.) Find the bias of the mle of σ.

.

7. Let $f_0(x) = 4x^3, 0 \leq x \leq 1$ and $f_1(x) = 2x, 0 \leq x \leq 1$. Suppose that we wish to test $H_0 : X$ has density f_0 vs $H_a : X$ has density f_1.

(a) (8 pts.) Construct the most powerful test of $\alpha \in (0,1)$ significance level based on X.

(b) (3 pts.) Which theorem did you use in part (a)?

(c) (7 pts.) Find the power of this test.

8. (12 pts.) In an early study of the effects of a strong magnetic field on the development of mice, 7 cages, each containing 3 albino female mice were subjected for a period of ten days to a magnetic field. 21 other mice housed in 7 similar cages were not placed in a magnetic field and served as controls. The following table shows the weight gains, in grams, for each of the cages.

<div style="text-align:center">

Magnetic Field Present: 22.8 10.2 20.8 27.0 19.2 10.2 14.2
Magnetic Field Absent: 23.6 31.0 19.5 26.2 26.5 25.2 24.5

</div>

State a nonparametric model, define hypotheses, and carry out a test at level $\alpha = 0.05$ which will enable you to decide whether there is a significant difference in weight gain between these two groups.

9. Do twins have the same IQ? To study this issue, a random sample of 7 pairs of twins was taken and their full scale IQ were recorded. Below the results are summarized in a table:

Older twin	98	89	103	105	129	98	90
Younger twin	89	87	103	96	126	101	93

(a) (12 pts.) State a parametric model, define hypotheses, and carry out a test at level $\alpha = 0.1$ which will enable you to decide whether there is a significant difference between twins IQ. Find the p-value for this test.

(b) (8 pts.) Perform a nonparametric test of the same hypotheses with $\alpha = 0.1$. Find the p-value for this test.

.

10. Let sample $(x_i, Y_i); i = 1, \ldots, n$ satisfy the linear regression model $Y_i = \beta + \beta x_i + \varepsilon_i$, where ε_i; $i = 1, \ldots, n$, are i.i.d. with distribution $N(0, \sigma^2)$.

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

(b) (5 pts.) Show that $\hat{\beta}$ is an unbiased estimate of β. Do **NOT** use matrix or vector space methods.

(c) (6 pts.) Find $Var(\hat{\beta})$. Do **NOT** use matrix or vector space methods.

(d) (7 pts.) In a study, 10 pairs $(x_i, y_i); i = 1, \ldots, 10$ were observed, and the following statistic were calculated:

$$\sum_{i=1}^{10}(x_i+1)y_i = 8, \quad \sum_{i=1}^{10}(x_i+1)^2 = 16, \quad RSS = \sum_{i=1}^{10}(y_i - \hat{\beta} - \hat{\beta}x_i)^2 = 4.$$

Based on these results, carry out the 0.05 level t-test of $H_0 : \beta = 1$ versus $H_1 : \beta \neq 1$.

.

11. Suppose that n measurements are to be taken under a treatment condition (X) and $2n$ measurements are to be taken independently under a control condition (Y). It is thought that the standard deviation of a single observation is about 10 under both conditions.

(a) (7pts.) How large should n be so that a 95% confidence interval for $\mu_X - \mu_Y$ has a width of 2?

(b) (7pts.) How large should n be so that the test of $H_0 : \mu_X = \mu_Y$ against the one-sided alternative $H_1 : \mu_X > \mu_Y$ has a power of 0.5 if $\mu_X - \mu_Y = 2$ and $\alpha = 0.05$?

(c) (7 pts.) Suppose a 95% confidence interval for $\mu_X - \mu_Y$ is (-1, 3). Explain the meaning of the interval in such a way that someone who had not ever studied statistics would understand.

12. (10 pts.) A student has two dice and wants to know whether the dice are fair. He rolled two dice together for 1000 times and recorded the frequency of "1". The following table shows the result. Is there a significant discrepancy from the fairness assumption? State a model, define hypotheses, and carry out the Pearson's chi-square test to answer this question with the significance level of $\alpha = 0.05$.

Number of 1	0	1	2
Frequency	652	315	33