The written Master’s Examination

Option Probability and Statistics, Option II

FALL 2003

Full points may be obtained for correct answers to 8 questions. Each numbered question (which may have several parts) is worth the same number of points. All answers will be graded, but the score for the examination will be the sum of the scores of your best 8 solutions.

Use separate answer sheets for each question. DO NOT PUT YOUR NAME ON YOUR ANSWER SHEETS. When you have finished, insert all your answer sheets into the envelope provided, then seal and print your name on it.

Any student whose answers need clarification may be required to submit to an oral examination.
1. (Math 310)
For any matrix $A$ (may be rectangular) show that $\text{rank}(A^T A) = \text{rank}(A)$.

2. (Math 413)
For $x \in (-1,1)$ define $f(x) = \begin{cases} (1 + x)^{\frac{1}{2}}, & x \neq 0, \\ e, & x = 0. \end{cases}$

(a) Is this function continuous at zero? Justify your answer.
(b) Is this function differentiable at zero? If so find its derivative.

3. (Stat 401)
Let $(X, Y)$ be uniformly distributed over the semicircle $\{(x, y) : x^2 + y^2 \leq 1, y \geq 0\}$.

(a) Find the conditional expectation $E(1 + Y^2 | X = x)$.

(b) Find the probability density function of $Z = X^2 + Y^2$.

4. (Stat 411)
Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with parameter $\theta$, where $\theta > 0$.
Suppose we are interested in the function $h(\theta) = P(X = 0) = e^{-\theta}$.

(1) Find the maximum likelihood estimator for $h(\theta)$.

(2) Find the minimum variance unbiased estimator for $h(\theta)$.

[You have to justify your answers. If you are using a theorem, you have to mention it.]

5. (Stat 461)
The life of brand A bulbs is exponential with mean 1000 hrs. The life of brand B bulbs is also exponential with mean 800 hours. Assume that the two brands are independent. A dark room is lighted with both bulbs simultaneously. Find the average number of hours until the room is turning dark again. (Hint: $X + Y = \min(X, Y) + \max(X, Y)$.)
6. Stat 461. The waiting time at a bus stop between successive arrivals of buses is uniform in (0, 1) measured in hours. Passengers arrive at a Poisson rate of 10 per hour. When I arrived at 8 AM, I just missed the bus. Find the expected number of persons getting into the next bus from this bus stop.

7. Stat 461. $N(t)$ is a Poisson process that denotes the number accidents in Dan Ryan from 0 hours to $t$ hours in a day. On the average there are 2 accidents per day in Dan Ryan. Independent of the number of accidents, the car involved in an accident is of Japanese make with probability .3, American make with probability .5, and German make with probability .1, and other makes with probability .10.
   i. Find the probability that by 12 noon there are 2 accidents and one is American make and the other Japanese make.
   ii. Find the expected number of American makes that are involved in accidents by 12 noon at Dan Ryan.

8. Stat 471 Miyako Auto Manufacturers are eyeing on potential customers during live TV coverage of the
   i. Basketball finals: Chicago Monkeys vs New York Donkeys
   ii. Soap Opera: Another World
      The manufacturers have 3 goals.
      Goal 1: The advertisement should reach at least 40 million high income men.
      Goal 2: The advertisement should reach at least 60 million low income people.
      Goal 3: The advertisement should reach at least 20 million high income women.
      Each minute of advertisement costs $60,000 for soap operas and $100,000 for basketball games. The audience for a minute advertisement during Basketball games are 7 million high income men, 10 million low income persons and 5 million high income women. For one minute advertisement during Soap operas, the respective figures are 1 million high income men, 5 million low income persons and 4 million high income women. The available advertisement budget is $600,000. The shortage from Goal 1 carries a penalty of 20 cents per person short of targeted persons. The shortage from Goal 2 carries a penalty of 5 cents per person short of targeted amount. Formulate a suitable linear program that solves the Car manufacturer’s interest.

9. Stat 473. Liz chooses one of the numbers 1, 2 or 3 in the initial move. In the next move Tom is revealed the actual number chosen by Liz if the number chosen is 2. Otherwise, Tom knows that Liz chose 1 or 3, but does not know which one she chose. Now Tom chooses one of the numbers 4 or 5. The game ends when Liz chooses 2 in her move and Tom chooses 4 in his move. In the other cases knowing Tom’s choice, but forgetting her initial choice, in her earlier move Liz chooses 6 or 7. The payoff to Liz from Tom is the sum of the numbers they chose if odd, otherwise Liz pays Tom the sum of the numbers they chose.
1. Draw a suitable game tree with information sets for the players and payoffs to Liz at terminal vertices for the game.

2. Enumerate verbally the pure strategies for Tom.

3. If the payoff matrix is to be formed, what will be the size, namely the number of rows and the number of columns for the payoff matrix?

- **Stat 477.** Consider a five component system which functions if either component 1 or 2 with component 5 or component 1 or 2 simultaneously with component 3 and component 4 function. Find the structure function and the min path sets.
For any matrix $A$ (may be rectangular) show that $\text{rank}(A^T A) = \text{rank}(A)$.

Since $A^T A = (A^T)^T A$, 
\[ \text{rank}(A^T A) \leq \text{rank}(A). \] -- (1)

Also for a vector $x$, 
\[ A^T A x = 0 \]
\[ \Rightarrow x^T A^T A x = 0 \]
\[ \Rightarrow A x = 0. \]

So $x \in \text{NullSpace}(A^T A)$ implies $x \in \text{NullSpace}(A)$

So $\text{nullity}(A^T A) = \text{nullity}(A)$

Let $A$ be an $m \times n$ matrix.

Hence $n - \text{rank}(A^T A) \leq n - \text{rank}(A)$

ie $\text{rank}(A) \leq \text{rank}(A^T A)$ -- (2)

(1) and (2) gives:

\[ \text{rank}(A^T A) = \text{rank}(A). \]
Fall #2 (a)

1. \( \ln(1+x) \)

\[ \lim_{x \to 0} \frac{\ln(1+x)}{x} \]

\[ = \lim_{x \to 0} \frac{1}{1+x} \]

\[ = 1 \]

\[ = e \]

\[ = e \]

\[ = e \]

\[ \ln(1) \]

\[ = 0 \]

\[ \text{f is cont. at 0} \]

1. Using the definition \( f(1+x)^{\frac{1}{x}} \)

2. \( e^x \) is continuous everywhere.

3. l'Hôpital's rule

4. \( \frac{1}{1+x} \) is continuous at 0
Fall #2 (b)

(b) \( S^{10}(x) = \lim_{x \to 0} \frac{\frac{x}{10} - \frac{5}{10}}{x} = \lim_{x \to 0} \frac{x}{x} = 1 \)

4. \( \left( \frac{1}{x} \ln(1+x) \right)'(1+x) \frac{1}{x} = \lim_{x \to 0} \frac{1}{x} \ln(1+x) \cdot e \)

5. \( \lim_{x \to 0} \frac{x}{1+x} - \ln(1+x) \frac{1}{x^2} \)

6. \( \lim_{x \to 0} \frac{x - (1+x) \ln(1+x)}{x^2 (1+x)} \)

7. \( \lim_{x \to 0} \frac{-\ln(1+x)}{2x + 3x^2} \)

8. \( \lim_{x \to 0} \frac{1}{1+2+3x} \)

9. \( \lim_{x \to 0} \frac{1}{1+x(2+6x)} = \frac{1}{2+6x} \) \( \text{at } x = 0 \).

10. \( \frac{1}{(1+x)(2+6x)} = \frac{1}{2+6x} \) follow from \( e \)'s derivatives rule.

Now using 10, in time we can work backward to obtain all the equalities.
\[ \begin{align*}
F(x) &= \int_{-\infty}^{x} f(y) \, dy \\
&= \begin{cases} 
0 & \text{if } x \leq 0 \\
\frac{2}{\pi} \int_{0}^{x} \cos(y^2) \, dy & \text{if } 0 < x \leq 1 \\
1 & \text{if } x > 1
\end{cases}
\end{align*} \]

\[ \begin{align*}
\int_{-\infty}^{\infty} F(x) \, dx &= \int_{-\infty}^{0} f(y) \, dy + \int_{0}^{1} F(x) \, dx + \int_{1}^{\infty} f(y) \, dy \\
&= \frac{1}{2} + \frac{1}{\sqrt{\pi}} + \frac{1}{2} \\
&= 1
\end{align*} \]
Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with parameter $\theta$, where $\theta > 0$. Suppose we are interested in the function $h(\theta) = P(X = 0) = e^{-\theta}$.

(1) Find the maximum likelihood estimator for $h(\theta)$.
(2) Find the minimum variance unbiased estimator for $h(\theta)$.

[You have to justify your answers. If you are using a theorem, you have to mention it.]

\[
(1) \text{MLE of } \theta: \quad L(\theta) = e^{-n\theta} \frac{\theta^{\sum X_i}}{\prod x_i!} \quad \text{ml} \theta = \bar{X}.
\]

$h(\theta)$ is a 1-1 function. So $h(\overline{X}) = \text{MLE of } h(\theta) = h(\overline{X}) = e^{-\overline{X}}$.

(2) Let $Y = \begin{cases} 1 & \text{if } X_i = 0 \\ 0 & \text{if } X_i > 0 \end{cases}$

$E(Y) = P(X_i = 0) = e^{-\theta} = h(\theta)$.

$\bar{X} = \frac{\sum X_i}{n}$ is complete sufficient for $\theta$.

So MUNE is $E(Y | \sum Z = z) = P(X_i = 0 | Z = z) = \frac{P(X_i = 0, \sum \frac{X_i}{n} = z)}{P(\sum \frac{X_i}{n} = z)}$

Since $\sum X_i \sim \text{Poisson}$,

$$= \frac{e^{-\theta} \cdot e^{-\frac{(n-1)\theta}{z}} \left[\frac{z}{(n-1)\theta}\right]^{\frac{n}{z} - 1}}{e^{-n\theta} \left[\frac{n0}{z}\right]^{\frac{n}{z} - 1}}$$

$$= \left(\frac{n-1}{n}\right)^{\frac{n}{z}}, \text{ where } z = \sum \frac{X_i}{z}.$$
Solution: Let $X,Y$ denote the life of the two bulbs. Since $X,Y$ are independent random variables with exponential distribution, the random variable $Z = \min(X,Y)$ is also exponentially distributed. Its expectation is simply $\frac{1}{\mu}$ where $\mu = \lambda_1 + \lambda_2$ where $\frac{1}{\lambda_1}$ is the mean of $X$ and $\frac{1}{\lambda_2}$ is the mean of $Y$. In our case $\frac{1}{\lambda_1} = 1000$, and $\frac{1}{\lambda_2} = 800$. Thus $E(Z) = \frac{1}{\frac{1}{1000} + \frac{1}{800}}$. Now $X + Y = \min(X,Y) + \max(X,Y)$. Thus $E(X + Y) = E(X) + E(Y) = 1000 + 800 = E(\min(X,Y) + \max(X,Y)) = E(\min(X,Y)) + E(\max(X,Y))$. Thus $E(\max(X,Y)) = 1800 - \frac{1}{\frac{1}{1000} + \frac{1}{800}} = 1355.56$ hrs. This is the expected time the room will be lighted.
Problem 6: (Stat 461): Let $T$ be the time until next rain of 8 A.M. arrives. $T$ is uniform in $(0, 1)$. Conditional on $T$, the number of passengers is Poisson at rate $10$ per hour $\Rightarrow \mathcal{X}$ Poisson with mean $10T$. \[ E(\mathcal{X}) = E(E(\mathcal{X}|T)) = E(10T) = 10E(T) = 5 \] (since $E(T) = \frac{1}{2}$).

Problem 7: Let $N_j(t)$, $N_k(t)$, $N_\lambda(t)$, be the number of Japanese, American, and German accidents among $N(t)$ accidents in $(0, t)$ hours.

The following theorem is applicable.

If $\{N(t)\}$ is a Poisson process having rate $\lambda$ and suppose each event is classified into types $N_j(t)$, $N_k(t)$, $N_\lambda(t)$ with probabilities $0.3$, $0.5$, $0.2$, $0.1$, then $N_j(t)$, $N_k(t)$, $N_\lambda(t)$ are independent Poisson with parameters $0.3t$, $0.5t$, $0.2t$, $0.1t$. Here $\lambda = 2$.

(i) We want $P\left[ N_j(t) = 1, N_k(t) = 1 \right]$ by 12 noon.

Since time is measured in days, we want $N(t) = N_j(t)$ to be Poisson with mean $2.5$ (i.e., $N_j(t)$ is Poisson with mean $2.5t$). Thus by independence the required prob is

\[
\frac{e^{-0.5} \cdot 0.5^1}{1!} \cdot \frac{e^{-0.3} \cdot 0.3^1}{1!} = \frac{0.5}{1!} \cdot \frac{0.3}{1!}
\]

Prob 8: Example 10, page 191 (Winston, Venkatraman, Class Test Book)
tained 150,000 variables and 12,000 constraints and was solved in one hour of computer time using Karmarkar's method. Using the simplex method, an LP with similar structure containing 36,000 variables and 10,000 constraints required four hours of computer time. Delta Airlines has used Karmarkar's method to develop monthly schedules for 7,000 pilots and more than 400 aircraft. When the project is completed, Delta expects to have saved millions of dollars.

### 4.16 Multiattribute Decision Making in the Absence of Uncertainty: Goal Programming

In some situations, a decision maker may face multiple objectives, and there may be no point in an LP's feasible region satisfying all objectives. In such a case, how can the decision maker choose a satisfactory decision? **Goal programming** is one technique that can be used in such situations. The following example illustrates the main ideas of goal programming.

#### Example 10: Burnit Goal Programming

The Leon Burnit Advertising Agency is trying to determine a TV advertising schedule for Priceler Auto Company. Priceler has three goals:

- **Goal 1**: Its ads should be seen by at least 40 million high-income men (HIM).
- **Goal 2**: Its ads should be seen by at least 60 million low-income people (LIP).
- **Goal 3**: Its ads should be seen by at least 35 million high-income women (HIW).

Leon Burnit can purchase two types of ads: those shown during football games and those shown during soap operas. At most, $600,000 can be spent on ads. The advertising costs and potential audiences of a one-minute ad of each type are shown in Table 52. Leon Burnit must determine how many football ads and soap opera ads to purchase for Priceler.

#### Solution

Let

\[
\begin{align*}
    x_1 &= \text{number of minutes of ads shown during football games} \\
    x_2 &= \text{number of minutes of ads shown during soap operas}
\end{align*}
\]

Then any feasible solution to the following LP would meet Priceler's goals:

\[
\begin{align*}
    \min \ (\text{or max}) \ z &= 0x_1 + 0x_2 \quad \text{(or any other objective function)} \\
    \text{s.t.} \quad 7x_1 + 3x_2 &\geq 40 \quad \text{(HIM constraint)} \\
    10x_1 + 5x_2 &\leq 60 \quad \text{(LIP constraint)} \\
    5x_1 + 4x_2 &\leq 35 \quad \text{(HIW constraint)} \\
    100x_1 + 60x_2 &\leq 600 \quad \text{(Budget constraint)} \\
    x_1, x_2 &\geq 0
\end{align*}
\]

From Figure 14, we find that no point that satisfies the budget constraint meets all three of Priceler's goals. Thus, (21) has no feasible solution. It is impossible to meet all of Priceler's goals, so Burnit might ask Priceler to identify, for each goal, a cost (per-unit short of meeting each goal) that is incurred for failing to meet the goal. Suppose Priceler determines that
Each million exposures by which Prizeler falls short of the HIM goal costs Prizeler a $200,000 penalty because of lost sales.

Each million exposures by which Prizeler falls short of the LIP goal costs Prizeler a $100,000 penalty because of lost sales.

Each million exposures by which Prizeler falls short of the HIW goal costs Prizeler a $50,000 penalty because of lost sales.

Burnit can now formulate an LP that minimizes the cost incurred in deviating from Prizeler’s three goals. The trick is to transform each inequality constraint in (21) that represents one of Prizeler’s goals into an equality constraint. Because we don’t know whether the cost-minimizing solution will undersatisfy or oversatisfy a given goal, we need to define the following variables:

\( s^+ \) = amount by which we numerically exceed the 7th goal

\( s^- \) = amount by which we are numerically under the 7th goal

The \( s^+ \) and \( s^- \) are referred to as \textit{deviational variables}. For the Prizeler problem, we assume that each \( s^+ \) and \( s^- \) is measured in millions of exposures. Using the deviational variables, we can rewrite the first three constraints in (21) as

\[
7x_1 + 3x_2 + x_3 - s_3^+ = 40 \quad \text{(HIM constraint)}
\]

\[
10x_1 + 5x_2 + x_3 - s_3^- = 60 \quad \text{(LIP constraint)}
\]

\[
5x_1 + 4x_2 + x_3 - s_3^+ = 35 \quad \text{(HIW constraint)}
\]
For example, suppose that \( x_1 = 5 \) and \( x_2 = 2 \). This advertising schedule yields \( 7(5) + 3(2) = 41 \) million HIM exposures. This exceeds the HIM goal by \( 41 - 40 = 1 \) million exposures, so \( s_1 = 0 \) and \( s_1^* = 1 \). Also, this schedule yields \( 10(5) + 5(2) = 60 \) million LIP exposures. This exactly meets the LIP requirement, and \( s_2 = s_2^* = 0 \). Finally, this schedule yields \( 5(5) + 4(2) = 33 \) million HW exposures. We are numerically under the HW goal by \( 35 - 33 = 2 \) million exposures, so \( s_3 = 2 \) and \( s_3^* = 0 \).

Suppose Priceler wants to minimize the total penalty from the lost sales. In terms of the deviational variables, the total penalty from lost sales (in thousands of dollars) caused by deviation from the three goals is \( 200s_1^2 + 100s_2^2 + 50s_3^2 \). The objective function coefficient for the variable associated with goal \( i \) is called the weight for goal \( i \). The most important goal has the largest weight, and so on. Thus, in the Priceler example, goal 1 (HIM) is most important, goal 2 (LIP) is second most important, and goal 3 (HW) is least important.

Burnt can minimize the penalty from Priceler’s lost sales by solving the following LP:

\[
\begin{align*}
\text{min } z &= 200s_1^2 + 100s_2^2 + 50s_3^2 \\
\text{s.t. } &7x_1 + 3x_2 + s_1 - s_1^* = 40 \quad (\text{HIM constraint}) \\
&10x_1 + 5x_2 + s_2 - s_2^* = 60 \quad (\text{LIP constraint}) \\
&5x_1 + 4x_2 + s_3 - s_3^* = 35 \quad (\text{HW constraint}) \\
&100x_1 + 60x_2 \approx 600 \quad (\text{Budget constraint})
\end{align*}
\]

All variables nonnegative

The optimal solution to this LP is \( z = 250, x_1 = 6, x_2 = 0, s_1 = 2, s_2 = 0, s_3 = 0, s_1^* = 0, s_2^* = 0, s_3^* = 5 \). This meets goal 1 and goal 2 (the goals with the highest costs, or weights, for each unit of deviation from the goal) but fails to meet the least important goal (goal 3).

**Remarks**

If failure to meet goal \( i \) occurs when the attained value of an attribute is numerically smaller than the desired value of goal \( i \), then a term involving \( s_i^* \) will appear in the objective function. If failure to meet goal \( i \) occurs when the attained value of an attribute is numerically larger than the desired value of goal \( i \), then a term involving \( s_i \) will appear in the objective function. Also, if we want to meet a goal exactly and a penalty is assessed for being over or under a goal, then terms involving both \( s_i^* \) and \( s_i \) will occur in the objective function.

Suppose we modify the Priceler example by deciding that the budget restriction of \$600,000 is a goal. If we decide that a \$1 penalty is assessed for each dollar by which this goal is unmet, then the appropriate goal programming formulation would be

\[
\begin{align*}
\text{min } z &= 200s_1^2 + 100s_2^2 + 50s_3^2 + s_4^2 \\
\text{s.t. } &7x_1 + 3x_2 + s_1 - s_1^* = 40 \quad (\text{HIM constraint}) \\
&10x_1 + 5x_2 + s_2 - s_2^* = 60 \quad (\text{LIP constraint}) \\
&5x_1 + 4x_2 + s_3 - s_3^* = 35 \quad (\text{HW constraint}) \\
&100x_1 + 60x_2 + s_4 - s_4^* = 600 \quad (\text{Budget constraint})
\end{align*}
\]

All variables nonnegative

In contrast to our previous optimal solution, the optimal solution to this LP is \( z = 33\frac{1}{2}, x_1 = 4\frac{1}{2}, x_2 = 2\frac{1}{2}, s_1 = \frac{1}{2}, s_2 = 0, s_3 = 0, s_4 = 33\frac{1}{2}, s_1^* = 0, s_2^* = 0, s_3^* = 0, s_4^* = 0 \). Thus, when we define the budget restriction to be a goal, the optimal solution is to meet all three advertising goals by going \$33\frac{1}{2} thousand over budget.

4.16 Multiattribute Decision Making in the Absence of Uncertainty: Goal Programming
Prob 9:

Pure strategies for Tom

1. If Tom chooses 1, 4 all the time.
   Liz choose 2.
2. If 1, 3 choose 4 else 5.
3. If Liz chose 1, 3 choose 5 else 4.
4. Choose 5 all the time.

Liz must decide what to choose in the chance move & what to do case she has 5 move again.

There are 3 possibilities for first move & 2 for next move of Liz all × 2 = 6 possible pure strategies for Liz. They are (1, 6), (1, 7), (2, 6), (2, 7), (3, 6), (3, 7).

Pay off has 4 rows & 6 columns for Tom.
Problem 10: We can write the five-component system with its structure function as

\[ \phi(x) = \max(x, x_2) \max(x_3 x_4, x_5) \]

\[ = (x + x_2 - x_2)(x_3 x_4 + x_5 - x_3 x_5 x_5) \]

There are 4 min path sets, namely \{1, 3, 4\}, \{2, 3, 4\}, \{1, 5\}, \{2, 5\}.

The following representation can also be useful for visualizing the structure function.