Math 105 Final Exam Review

The final exam will take place on Thursday, December 11, 2003, from 9:00–11:30 am in Warren Hall B45.

Slopes and equations of lines (1.1). A line with slope m and y-intercept b has equation $y = mx + b$. A line with slope m passing through the point (x_1, y_1) has equation $y - y_1 = m(x - x_1)$. (Both forms are equivalent, except one might be easier to use in a particular problem.)

Linear functions and appications (1.2). Finding the intersection of two lines. Applications in economics such as supply and demand curves, or cost analysis. The equilibrium price occurs at the intersection point of the supply line $p = S(q) = mq + b$ and the demand line $p = D(q) = nq + b$. The marginal cost is the slope of the linear cost function $C(x) = mx + b$.

Least squares line (1.3). The least squares line $Y = mx + b$ that gives the best fit to the data points (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ has slope m and y-intercept b that satisfy the equations

$$\begin{aligned} nb + (\textstyle\sum x)m &= \textstyle\sum y \\ (\textstyle\sum x)b + (\textstyle\sum x^2)m &= \textstyle\sum xy. \end{aligned}$$

The coefficient of correlation given by

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \cdot \sqrt{n(\sum y^2) - (\sum y)^2}}$$

measures the strength of the fit of the least squares line. If $|r|$ is near one, then there is a strong linear fit. It $|r|$ is near 0, then there is not a strong linear fit. The sign of r indicates the sign of the slope of the least squares line. If $r > 0$, then the least squares line has positive slope, but if $r < 0$, then the least squares line has negative slope.

Solutions of linear equations by the echelon method (2.1) or Gaussian elimination (2.2). On the exam you can use the method you prefer. You should expect to have to find the equilibrium vector of a two or three state Markov chain by solving linear equations.

Multiplication of matrices (2.4). The consideration of Markov chains in 10.1 has given us a new reason to be interested in this. The kth power of the transition matrix P gives the k step transition probabilities.

Inverses of matrices (2.5). Use inverses to solve the matrix equation $Ax = b$, or to find the fundamental matrix $F = (I - Q)^{-1}$ of an absorbing Markov chain.

Basic concepts of probability and set theory (7.1, 7.2, 7.3, 7.4). $n(A \cup B) = n(A) + n(B) - n(A \cap B)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, $P(A') = 1 - P(A)$.

Conditional probability (7.5). The definition of the conditional probability of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

provided $P(B) \neq 0$. We have the multiplication rule $P(A \cap B) = P(A|B) \cdot P(B)$. Two events A and B are independent if $P(A|B) = P(A)$, or equivalently, if $P(A \cap B) = P(A) \cdot P(B)$.

Bayes' Theorem (7.6). Since $P(E \cap F) = P(E|F)P(F)$, we have

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

Note that $P(E \cap F) = P(F \cap E)$ so that

$$P(E|F)P(F) = P(F|E)P(E) \quad \text{or} \quad P(F|E) = \frac{P(E|F)P(F)}{P(E)}.$$

Since we can write $E = (E \cap F) \cup (E \cap F')$, as a disjoint union, we use the definition of conditional probability to find $P(E) = P(E \cap F) + P(E \cap F') = P(E|F)P(F) + P(E|F')P(F')$. We now substitute this into the above equation for $P(E)$ to derive Bayes' Theorem:

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F')P(F')}.$$

The more general form is as follows. Suppose that F_1, F_2, $\ldots$, F_n are disjoint events whose union is S, the sample space. Then,

$$P(F_i|E) = \frac{P(F_i)P(E|F_i)}{P(F_1)P(E|F_1) + P(F_2)P(E|F_2) + \cdots + P(F_n)P(E|F_n)}$$

Multiplication principle (8.1). If there are m_1 ways to make choice 1, and m_2 ways to make choice 2, $\ldots$, and m_n ways to make choice n, then there are $m_1 \cdot m_2 \cdots m_n$ ways to make all of the choices. (Note that multiplication is commutative: $x \cdot y = y \cdot x$. One way to think about counting the total number of ways of making all the choices is via a tree diagram. Draw a tree diagram with m_1 branches for the first choice, and then for each of these first branches there are m_2 second branches. The total number of choices is the product $m_1 \cdot m_2$, etc.)

Permutations (8.1). The number of ways n people can stand in a line is $n!$. The number of ways of selecting r things out of n total when order is important is

$$P(n, r) = \frac{n!}{(n-r)!}.$$

Suppose that we have m_1 objects of type 1, m_2 objects of type 2, ..., and m_n objects of type n (e.g., the letters in STATISTICS). The number of ways these $m = m_1 + m_2 + \cdots + m_n$ things can be arranged is

$$\frac{m!}{m_1! \cdot m_2! \cdots m_n!}.$$

Combinations (8.2). The number of ways of selecting r things out of n total when order does NOT matter is

$$\binom{n}{r} = C(n,r) = \frac{n!}{r!(n-r)!}.$$

Probability applications of counting (8.3). Remember that the probability of an event E is $P(E) = n(E)/n(S)$. The difficult part is computing $n(E)$. This is where combinations and permutations may be helpful.

Ex. Suppose a bag contains 10 tiles, three labelled T, three labelled S, two labelled I, one labelled C, and one labelled A. If the bag is well-shaken, and tiles are selected one at a time then placed down in the order selected adjacent to the previous letter (like in Scrabble), what is the probability that they spell the word STATISTICS?

Binomial distribution (8.4). If independent events have probability p of success (and therefore probability $1-p$ of failure), then the probability of exactly k successes among n trials is

$$\binom{n}{k} p^k (1-p)^{n-k}.$$

The mean of a binomial random variable is np, and the standard deviation is $\sqrt{np(1-p)}$.

Distribution and mean of a random variable (8.5). Let X be a random variable representing the possible outcomes of some experiment. The distribution of X is the listing of all the probabilities $P(X = x)$ where x are the possible values for X. This information is often displayed in a table or histogram. The mean of X, or the expected value of X, written $E(X)$, is the weighted average of the possible outcomes:

$$E(X) = x_1 \cdot P(X = x_1) + \cdots + x_n \cdot P(X = x_n).$$

Ex. Suppose two dice are rolled. Let X be the sum of the two upturned faces. The distribution of X is

$X = x$	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Summary statistics (9.1,9.2). For a data set $\{x_1, x_2, \ldots, x_n\}$, the sample mean is

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n},$$

the sample variance is

$$s^2 = \frac{\sum x_i^2 - n(\overline{x})^2}{n-1}$$

and the sample standard deviation is $s = +\sqrt{s^2}$.

Normal distribution (9.3). If X is a normal random variable with mean μ and standard deviation σ, then the random variable

$$Z = \frac{X - \mu}{\sigma}$$

has a normal distribution with mean 0 and standard deviation 1. Using this "Z-transformation" we can compute "z-scores" corresponding to observations of X. This is useful because we can evaluate probabilities from *any* normal distribution by making the Z-transformation and looking up the appropriate z-scores in the table of normal values.

Note that we can determine "inverse z-scores" as well. If Z is normal with mean 0 and SD 1, then $X = \sigma X + \mu$ is normal with mean μ and SD σ.

Normal approximation to the binomial (9.4). Let Y be the number of successes in n independent trials of a binomial experiment (a.k.a., binomial trials, Bernoulli process) with each trial having success probability p. The mean number of successes is $E(Y) = np$ and the standard deviation of the number of successes is $SD(Y) = \sqrt{np(1-p)}$. The normal approximation to the binomial says that

$$\frac{Y - np}{\sqrt{np(1-p)}}$$

has a distribution that is close to a normal with mean 0 and SD 1.

Ex. A binomial experiment is repeated 80 times, with each trial having success probability 0.4. Let Y be the number of successes. We compute $E(Y) = 80 \times 0.4 = 32$, $SD(Y) = \sqrt{80 \times 0.4 \times 0.6} \approx 4.38$. Compute $P(Y = 35)$. This can be computed *exactly* as

$$P(Y = 35) = \binom{80}{35}(0.4)^{35}(0.6)^{55}.$$

However, this number cannot be evaluated by a calculuator (such as the TI-83). Instead we can *approximate* it via the normal approximation. That is, $P(Y = 35) \approx P(34.5 \leq X \leq 35.5)$ where X is normal mean 32 and SD 4.38. Now make the z-transformation to conclude

$$P(Y = 35) \approx P(34.5 \le X \le 35.5) = P\left(\frac{34.5 - 32}{4.38} \le \frac{X - 32}{4.38} \le \frac{35.5 - 32}{4.38}\right)$$
$$\approx P(0.57 \le Z \le 0.80) = 0.7881 - 0.7157 = 0.0724.$$

To compute $P(29 \le Y \le 35)$, the approximation is $P(28.5 \le X \le 35.5)$, but to compute $P(29 < Y < 35)$, the approximation is $P(29.5 \le X \le 34.5)$. WHY?

Markov chains (10.1, 10.2, 10.3). Also see review notes for Chapter 2 above. To help visualize a Markov chain, draw a transition diagram. (I visualize a frog in a pond jumping between lilypads.)

The transition probabilities $p_{i,j}$ can be recorded in the transition matrix P. The (i, j)th entry in P records the probability $p_{i,j}$ and represents the probability of a one step transition from state i to state j.

The matrix powers P^k give the k step transition probabilities. The (i, j)th entry in the matrix P^k gives the probability of being in state j after k steps, starting in state i.

To determine the long run behaviour of the Markov chain, investigate the matrix power P^n for large n.

There are two special kinds of Markov chains. Regular Markov chains are ones whose transition matrix powers P^k have all non-zero entries for some k.

Ex.

$$\begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.5 & 0.5 \\ 0.6 & 0.4 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.5 & 0 & 0.5 \\ 0.6 & 0 & 0.4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0.75 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

are regular (WHY?), but

$$\begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

is not regular (WHY?). (Observe that there is not much difference between these four matrices. In order to determine regularity, you must do some work! Don't try to guess based on zeroes in the original matrix.)

Regular Markov chains have the property that there exists an equilibrium probability distribution V which may be found by solving $VP = V$. Use techniques from Chapter 2 to solve this. Furthermore, for regular Markov chains, in the long run P^n converges to a matrix, each of whose rows have the same entries as V. This means that in the long run, where you end up is independent of where you start.

An absorbing Markov chain is one that has at least one absorbing state, and is such that no matter which state you start in, you can end up in some absorbing state. For an absorbing Markov chain, no matter where you start you will be absorbed in a finite number of moves. However, where you are absorbed is highly dependent on where you start.

To determine the long run behaviour, write P as the block matrix

$$P = \left[\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array}\right].$$

Compute the associated fundamental matrix $F = (I - Q)^{-1}$.

Compute the matrix product FR.

In the long run P^n converges to the matrix

$$\left[\begin{array}{c|c} I & 0 \\ \hline FR & 0 \end{array}\right].$$

Selected Review Problems

Chapter 1. Page 44 # 15, 19, 27, 33, 41, 47.

Chapter 2. Page 117 # 3, 5, 7, 9, 21, 23, 25, 27, 31, 35, 43, 45, 48.

Chapter 7. Page 373 # 39, 41, 43, 59, 61, 63, 67, 69. Page 378 # 1, 2, 3.

Chapter 8. Page 432 # 1, 3, 5, 7, 9, 13, 15, 17, 19, 21, 23, 25, 27, 33, 35, 39, 53, 55.

Chapter 9. Page 484 # 17 (also find mean, median, mode), 23, 25, 29, 31, 33, 35, 41, 43.

Chapter 10. Page 520 # 7, 9, 11, 13, 17, 29, 30, 31, 34, 37, 43, 46—51.

Selected Answers

(Odd numbered answers may be found in the back of the textbook. There is a copy of the student solutions manual on reserve in the Mathematics Library on the fourth floor of Malott Hall.)

Chapter 2. Page 117 # 48. The solution of the system is $(2, 3, -1)$, and

$$A^{-1} = \begin{bmatrix} 5/22 & 7/22 & 1/22 \\ 7/22 & 1/22 & -3/22 \\ 3/22 & -9/22 & 5/22 \end{bmatrix}.$$

Chapter 7. Page 378 # 2. 0.16/0.295

Chapter 9. Page 484 # 17. mean$= 52$, mode$= 29$, median$= (43 + 51)/2 = 47$

Chapter 10. Page 520 # 30. $V = [2/3 \ \ 1/3]$ so the long range market share for Dogkins is 2/3.

Chapter 10. Page 520 # 34. $V = [47/114 \ \ 32/114 \ \ 35/114]$

Chapter 10. Page 522 #46. Just check this. It is straightforward.

Chapter 10. Page 522 #48. Assuming the states in the rearranged block matrix for P are 1, 6, 2, 3, 4, 5,

$$Q = \begin{bmatrix} 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/2 \end{bmatrix}.$$

Chapter 10. Page 522 #50. We want the entry in row 3, column 3 of F which is 8/3.