MONT 106N – Identifying Patterns
Information and Sample Questions for Final Exam

General Information

The final exam will be given on Thursday, December 17 from 2:30pm to 5:30pm in our regular classroom, Stein 307. The exam will be roughly twice as long as the midterm. If you are well-prepared and work steadily, I expect that you will complete it within two hours. However, if you need more time, you will have until 5:30pm to work on the exam.

The final will cover the material from Parts I - IV (Chapters 1 - 15) of Freedman, Pisani, Purves. This means in particular:

1. Issues about the design of studies – controlled experiments versus observational studies, questions like the ones from Problem Set 1 about whether conclusions are justified, possible confounding factors, and so forth. A question here might ask for a paragraph-length written response rather than just a calculation.
2. Histograms – be able to construct histograms for "small" data sets, explain the conceptual meaning of the histogram, etc.
3. Averages, medians, and standard deviations – be able to compute these for "small" data sets by hand. Also be prepared for conceptual questions about what these measures tell us about data.
4. Normal approximation to data – using the normal curve and the normal curve area table (questions similar to Problem Set 3), percentiles, etc.
5. Equations of lines in the plane.
6. Correlation: scatter plots, the correlation coefficient, its meaning; be able to compute the correlation coefficient in a relatively small case.
7. "Regression to the mean" and the "regression fallacy."
8. The SD and regression lines; making predictions with the regression line, the rms error for regression, distribution within vertical strips of the scatter plot.
9. Chances, conditional probabilities, independence, the addition and multiplication rules.
10. Binomial coefficients and the binomial formula for probabilities.

Some questions will involve computations, so you probably will want to bring a calculator. I will provide graph paper and a copy of the normal curve table for any questions that require those. Other questions may ask for a short answer or an explanation of a concept.

No use of cell phones, pagers, I-pods, or any other electronic devices beyond a calculator will be allowed during the exam – turn them off and stow them in your backpack.

Review Session

I would be happy to run an evening review session before the final during exam week if there is interest. We can discuss this in class on Monday, December 7.

Philosophical Comments and Suggestions on How to Prepare

- The reason we give final exams in almost all mathematics classes is to encourage students to "put whole courses together" in their minds. Also, preparing for the final should help to make the ideas "stick" so you will have the material at your disposal to use in the spring.
- It may not be necessary to say this, but here goes anyway: *You should take this exam seriously* – it is worth 25% of your course average and it can pull your course grade up *or down* depending on how you do.
- Get started reviewing early and do some work on this *every day* between now and the date of the final. Don't try to "cram" at the end.
- Reread your class notes in addition to the text, especially for topics where you lost points on the midterm. There are a lot of worked-out examples and discussions of all of the topics we have covered there.
- Look over the midterm exam with the solutions. Go over your corrected problem sets. If there were questions where you lost a lot of points, be sure you understand why what you did was not correct, and how to solve those questions.
- Be sure you actually do enough practice problems so that you have the facility to solve exam-type questions in a limited amount of time.
- Look over some of the problems from the text from each section in the chapters we have covered, and compare your answers with the ones given in the back of the text.
- Then take a couple of hours and try the practice exam problems below.

Practice Exam Questions

I. Short answer.

A) Suppose that a scatter diagram is football shaped. What is it about the rms error of the regression line that makes the regression line special?

B) George is playing roulette (where, in the US version, there are 18 red numbers, 18 greens, and 2 additional pockets, for a total of 38). For the last 50 spins of the wheel, he has bet \$1 on 17, his favorite number. Unfortunately, he has lost each time, losing \$50 in the process. Ever the optimist, George decides his luck must change on the next spin of the wheel, so he bets his remaining \$50 on 17. Georges friend Martha, who always brings George back to earth, says he is talking non-sense and that hes likely to lose again. Is George right, is Martha right, or are neither of them right? Explain your answer.

C) For men age 25-34 in the HANES2 sample, the regression equation for predicting height (y) from education level (x) is $y = (.25)x + 66.75$. Does going to college increase a man's height? Explain.

D) In the US in 1990, 20273 people were murdered, compared with 16848 in 1970 – nearly a 20% increase. True or false, and explain: The figures show that the US became a more violent society over the period 1970-1990.

E) A study is made of the age at entrance of college first-year students. Is the SD about 1 month, 1 year, or 5 years? Explain.

F) A coin is tossed 10 times. What is the chance of getting 7 heads and 3 tails?

G) True or False and explain: The median years of education completed for US men ages 30 - 50 would be greater than the average years of education completed for the same group.

H) A computer program is developed to predict the GPA of college first-year students from their high school GPA's. The program is tested on a class whose college GPA's are known. Assume both GPA's are reported on a standard $0-4$ scale. The RMS error for the computer's prediction is 3.12. Is something wrong? Yes or no, and explain.

II. A study of draftees into the Danish army showed a positive correlation between near-sightedness and intelligence.

A) True or false and explain: Draftees who were more near-sighted were also more intelligent on average.

B) True or false and explain: The data show that intelligence causes near-sightedness. (If you say this is false, give a specific example of a possible confounding factor that might account for the result.)

III. Exercise 10 on pages 53-54 from the text.

IV.

A) Draw the scatter plot and find the correlation coefficient for the dataset:

x:	4	5	7	8	8	10
y:	7	0	9	9	13	16

B) Find the equations of the SD- and regression lines for the data from part A, and add the SD- and regression lines to your scatter plot.

V. Exercise 14 on page 267 of the text.

VI. The ACT is an alternative test to the SAT. In 2007, in Massachusetts, the average score on the English portion of the ACT test was 23.5 with an SD of 5.6. The average score on the mathematics portion of the test was 23.6 with an SD of 5.2. (Each test is scored out of total of 36 points.) The correlation coefficient between the English scores and mathematics scores was $r = 0.6$.

A) A total of 350 students in Massachusetts scored 31 on the English portion of the ACT. Estimate the average mathematics score for these 350 students and the SD for these scores.

B) Estimate how many of these 350 students scored below 25 on the mathematics portion of the ACT.

VII. The aces and kings from a deck of cards are placed in a stack of eight cards by themselves (so there are two hearts, two spades, two diamonds and two clubs). Three draws with replacement are made from this stack of cards.

A) What are the chances that all three cards are spades?

B) What are the chances that none of the cards are spades?

C) What are the chances that there is a heart, club, or diamond among the three cards?

VIII. In a large class, the average score was 50 out of 100 and the SD was 20. The scores followed a normal distribution.

A) Two brothers took the final. One placed at the 70th percentile and the other placed at the 80th percentile. By how many points did their scores differ?

B) Two sisters took the final. One placed at the 80th percentile and the other placed at the 90th percentile. By how many points did their scores differ?

IX. One ticket is drawn at random from each of the two boxes: (i) 1 2 3 4 5 and (ii) 1 2 3 4 5 6 .

A) What is the chance that one of the numbers is 2 and the other is 5?

B) What is the chance that the sum of the numbers is 7?

C) What is the chance that one numbers is (strictly) bigger than twice the other?