

Sample Final Exam

March 13

Your section: _____ Print your name: _____

Sign your name: _____

This is a closed book exam. However, you are allowed to bring two sheets (double-sided) of 8.5" × 11" paper with notes. This final exam consists of 12 problems; you have to answer one out of Problems 7 and 8 and any eight problems out of the remaining 10. The maximum you can score is 100.

Please select one of the following options for returning your graded exam.

- I am providing a self-addressed stamped envelope (55c postage).
Mail the exam.
- Deposit the exam in the box in front of room B-301 at Padelford Hall.
I will pick it there.
- Hold the exam. It will be available during the instructor's office hours
in Spring quarter, 2001.

Good luck!

Problem:	.. 1 2 3 4 5 6 ..
Points:	11	11	11	11	11	11
Credit:						

Problem:	.. 7 8 9 10 11 12 ..	Sum
Points:	12	12	11	11	11	11	100
Credit:							

Problem 1. (i) Some studies find an association between liver cancer and smoking. However alcohol consumption is a confounding variable. This means

(i) Alcohol causes liver cancer.

(ii) Drinking is associated with smoking, and alcohol causes liver cancer.

(ii) Sketch schematically (just show the shape by means of a curve) a histogram for income distribution in the U.S. What is the relation between mean income and median income ? (4+7=11 points)

Problem 2. (i) Construct a list of 5 numbers for which the mean and the median are the same.

(ii) Here is a list of numbers: 0.7, 1.6, 9.8, 3.2, 5.4, 0.8, 7.7, 6.3, 2.2, 4.1, 8.1, 6.5, 3.7, 0.6, 6.9, 9.9, 8.8, 3.1, 5.7, 9.1

(i) Without doing any arithmetic guess whether the average is around 1, 5 or 10. Justify briefly.

(ii) Without doing any arithmetic guess whether the SD is around 1, 3 or 6. Justify briefly.

(3+8 = 11 points)

Problem 3. Use the techniques developed in class to find the correlation coefficient for the following data set. (Do not use the correlation function on a statistical calculator.) Show all your work. (11 points)

x	y
3	11
5	6
7	4
5	9
8	1
2	3

Problem 4. (i) A law school finds the following relationship between LSAT scores and first-year scores (for students who finish the first year):

average LSAT score = 162, SD = 6

average 1'st year score = 68, SD = 10, $r = 0.60$

Of the students who scored 165 on the LSAT, about what percentage had first-year scores of over 75 ?

(ii) Find the correlation between X and Y where

x	y
3	7
5	13
7	19
9	25
11	31

(7 + 4 = 11 points)

Problem 5. (i) When is the R.M.S error for the regression line for Y on X, the same as the SD of Y ?

(ii) A statistician computes the slope of the regression line of Y on X as 2.5 and the slope of the regression line of X on Y as -.3. Do you think there is something wrong with his computations ? Justify your answer.

(iii) Another statistician now recomputes the slope of the regression line of Y on X as 2.5 and the slope of the regression line of X on Y as 0.5. Do you think he is computing the slopes correctly, or is he wrong ? (3 + 4 + 4 = 11 points)

Problem 6. A die is rolled four times. What is the chance that

- (i) not all the rolls show 3 or more spots ?
 - (ii) none of the rolls show 3 or more spots ?
 - (iii) the sum of the numbers on the rolls is 24 given that the first roll is a 5 ?
- (4 + 4 + 3 = 11 points)

Problem 7. A fair coin is tossed 400 times. You get to pick 21 numbers. If the number of tails turns out to be equal to one of your 21 numbers, you win 100 dollars. Which 21 numbers should you pick and what is approximately your chance of winning 100 dollars ? (12 points)

Problem 8. Sam and Russ each toss a die. If Sam gets 3 or more spots she records a 3, otherwise she records a 1. If Russ gets 3 or more spots he records a 1, otherwise he records a 3. They then multiply the numbers that they have recorded. They repeat the experiment a 100 times. The average of the 100 products will then be around (a) give or take (b) or so. Fill in (a) and (b) with appropriate numbers. (12 points)

Problem 9. A multiple-choice quiz has 50 questions. Each question has four possible answers, one of which is correct. Eight points are given for each correct answer, but 2 points are taken off for a wrong answer.

A student answers all 50 questions at random (by guessing).

- (a) The student's score is like the _____ of _____ draws made at random _____ replacement from the box



(specify an appropriate box)

For the first space, your options are *sum* and *average*. For the second space, select a number, and for the third space, choose from *with* and *without*.

- (b) If the passing score is 100, what is the student's chance of passing? Show all your work.

(11 points)

Problem 10. (a) If a sampling procedure is biased, can I rectify it by taking a larger sample ? Explain briefly.

(b) Is it always true that the absolute size of the sample and not its size relative to the population determines the accuracy of the sample percentage for the population percentage ? Discuss briefly. (4+7 = 11 points)

Problem 11. You want to estimate the proportion of Democrats in a town with 100000 eligible voters. You also want an accuracy measure for your estimate for the proportion. Cost constraints allow you to interview only 500 voters. Describe clearly how you would proceed to do so. (11 points)

Problem 12. One ticket is drawn at random from two boxes A and B below.

Box A: 1,2,3,4,5.

Box B:1,2,3,4,5,6.

Find the chance that the sum of the numbers is 7. Also find the chance that one of the numbers is strictly bigger than twice the other. (4+7=11 points)

