

## MATH 243 (PHILLIPS, SPRING 2008): SAMPLE MIDTERM 1 PROBLEMS

General comments:

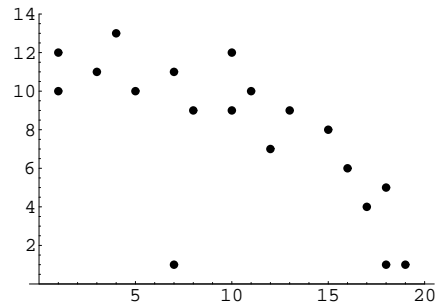
- Besides these problems, see Midterm 1 from the previous course, the list of problems from later exams from the previous course which deal with material we have already covered, the homework, the quizzes, the examples in the book, etc. (In the previous course, Midterm 1 was earlier in the quarter, so covered less material.)
- Many variations of these problems are possible. In particular, the questions could be asked in different ways, something which is multiple choice here could be short answer on the real exam, and the scatterplots and histograms could have very different features. For example, a problem might ask something about a stemplot, or for information about quartiles instead of the mean, etc. Thus, a problem could test the same ideas but superficially not look like anything here.
- There will be something on the exam which is not covered in the sample problems.
- The number of problems of a given type which appear here need not represent the weight given to a topic on the actual midterm. The choice of problem here is partly intended to reflect the possible variety of questions on the midterm.
- Note that normal probability problems will ask for reasonably well drawn graphs.
- Recall that you may bring one  $8.5 \times 11$  page of notes and a calculator; see the syllabus for details.
- Recall that **no** calculator help will be provided on the exam. If you need to, write instructions for your calculator on your  $8.5 \times 11$  page of notes.
- Recall that grading questions must be in writing and turned at the beginning of the lecture of Wednesday 7 May; see the syllabus for details.

### 1. MULTIPLE CHOICE PROBLEMS

1. Which of the following is a technique for improving a stemplot with too many stems and too few leaves?
  - a. Using separate plots, one for each stem.
  - b. Splitting stems.
  - c. Merging stems.
  - d. Rounding data.
  - e. Omitting some of the data.
  - f. Counting each data point three times.
2. A study of consumer behavior finds a positive correlation between sales of ice cream ( $x$ ) and sales of beer ( $y$ ). What is a plausible explanation for the observed correlation?
  - a. Bias, since there should be no correlation.
  - b. Eating ice cream makes people want to drink beer.
  - c. Drinking beer makes people want to eat ice cream.
  - d. The effect of a lurking variable such as outdoor temperature.
  - e. A mistake, since people under 21 are allowed to buy ice cream but not beer.

3. From the University of Oregon, 26 students are randomly selected. They are then asked the total cost of their textbooks for the term. What is the sample?
- The total cost of their textbooks for the term.
  - All University of Oregon students.
  - The 26 randomly selected students.
  - The textbooks bought by the 26 randomly selected students.
  - Impossible to tell from the information given.
  - None of the above.
4. A phone-in poll conducted by the athletic department reported that 88% of those who called in thought Ernie Kent was their number one choice for basketball coach. The number 88% is what?
- A sample.
  - A parameter.
  - A statistic.
  - A population.
  - Both (a) and (d).
  - Both (b) and (c).
  - None of the above.
5. Gary has 15 stocks in his portfolio. The mean value of the stocks is \$20,000, the median value is \$30,000, the standard deviation of the values is \$14,000, and the range of the values is \$55,000. The total value of Gary's portfolio is:
- \$450,000
  - \$300,000
  - \$210,000
  - \$825,000
  - Impossible to tell from the information given.
  - None of the above.
6. I decide to add 10 points to the Midterm 1 score of everybody in my section of Math 251. Then:
- The mean will be increased by 5, the median will be increased by 5, the standard deviation will stay the same, and the third quartile will be increased by 7.5.
  - The mean, median, standard deviation, and third quartile will all be increased by 10.
  - The mean will be increased by 5, the median will be increased by 5, the standard will be increased by 10, and the third quartile will be increased by 7.5.
  - The mean, median, and third quartile will all be increased by 10, and the standard deviation will stay the same.
  - The mean, median, and standard deviation will all be increased by 5, and the third quartile will be increased by 7.5.
  - Any of the above could happen.
  - None of the above is correct.
7. Research doctors test drugs by prescribing different amounts and observing the results on their patients. One question we could ask here is: "Does the amount of drug prescribed determine the length of the recovery time?" The explanatory variable is:
- The drug being tested.
  - The amount of drug prescribed.
  - The length of the recovery time.
  - The results on the patients.
  - The disease the patients had.
  - None of the above.

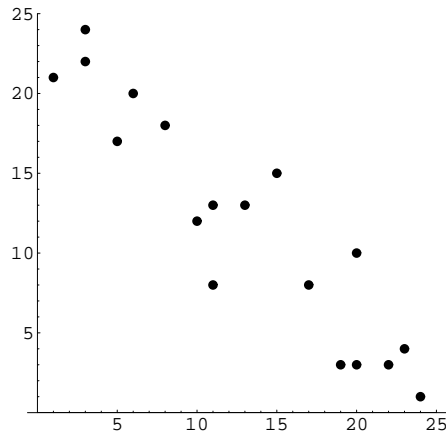
8. Consider the following scatterplot:



The association of the variables plotted is:

- Clearly positive with no outliers.
- Clearly positive with one or more outliers.
- Neither clearly positive nor clearly negative.
- Clearly negative with no outliers.
- Clearly negative with one or more outliers.
- Not defined.
- None of the above.

9. Consider the following scatterplot:

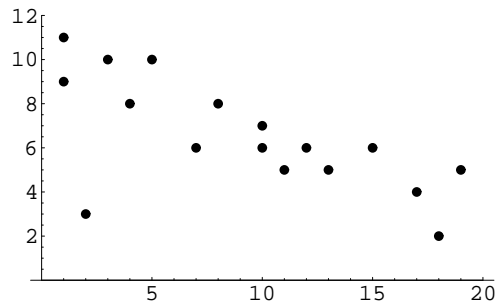


- The variables plotted are strongly negatively associated.
- The variables plotted are weakly negatively associated.
- The variables plotted are not significantly associated.
- The variables plotted are weakly positively associated.
- The variables plotted are strongly positively associated.
- The association of the variables plotted is not defined.

10. Which of the following are useful ways to illustrate one-variable data?

- A histogram.
- A scatterplot.
- A residual plot.
- A stemplot.
- Both (a) and (d).
- Both (b) and (c).
- Both (a) and (c).
- None of the above.

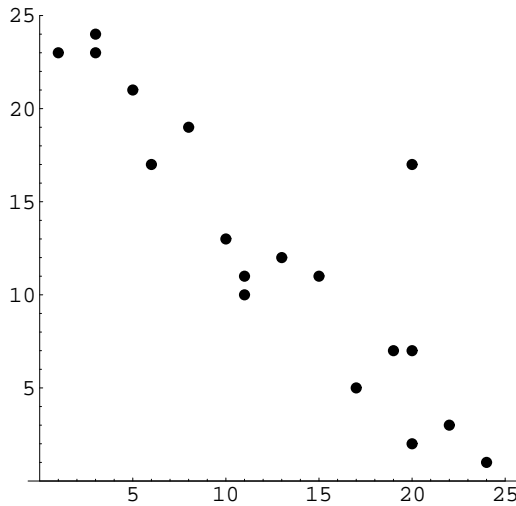
11. Consider the following scatterplot:



The association of the variables plotted is:

- a. Clearly positive with no outliers.
- b. Clearly positive with one or more outliers.
- c. Neither clearly positive nor clearly negative.
- d. Clearly negative with no outliers.
- e. Clearly negative with one or more outliers.

12. Consider the following scatterplot:



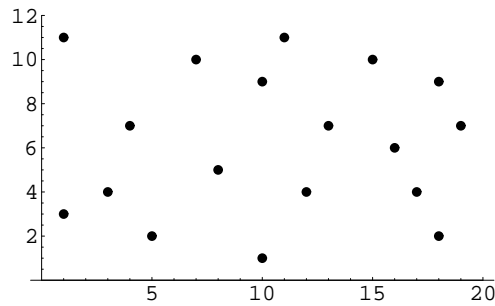
The association of the variables plotted is:

- a. Clearly positive and roughly linear.
- b. Clearly positive and clearly nonlinear.
- c. Neither clearly positive nor clearly negative.
- d. Clearly negative and roughly linear.
- e. Clearly negative and clearly nonlinear.

13. A data set has mean 100 and standard deviation 20. The median:

- a. Must be 100.
- b. Must be 120.
- c. Must be 80.
- d. Must be 20.
- e. Cannot be determined from the information given.
- f. None of the above.

14. Consider the following scatterplot:



A least squares regression line is likely:

- To be useful to describe the data.
- Not to be useful, because the association isn't linear.
- Not to be useful, because there are too many outliers.
- Not to be useful, because there is little association between the variables.
- Not to be useful, because the regression line is not defined.

15. A researcher is studying the gas mileage of cars for the model year 2003. He finds the gas mileage (in miles per gallon) for each of a number of cars, and statistically analyzes the data he gets. The units of the standard deviation of the gas mileages are:

- Miles per gallon.
- Gallons per mile.
- Miles.
- Gallons.
- None: the standard deviation has no units.
- Cannot be determined from the information given.

16. A researcher finds the age (in years) and height (in meters) of each member of a sample of Douglas fir trees, and statistically analyzes the data he gets. With age being the explanatory variable, units of the correlation of age and height are:

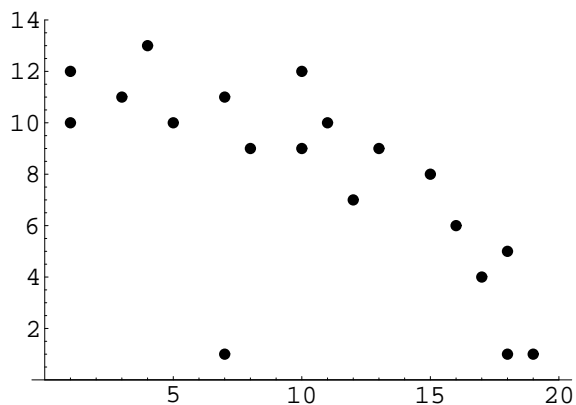
- Meters per year.
- Years per meter.
- Years.
- Meters.
- None: the correlation has no units.
- Cannot be determined from the information given.

17. The masses (in kilograms) of crumple-horned snorkacks are normally distributed with mean 70 and standard deviation 8. The masses (in kilograms) of spiral-horned snorkacks are normally distributed with mean 66 and standard deviation 4. The masses (in kilograms) of purple-spotted snorkacks are normally distributed with mean 45 and standard deviation 5. The Zoo of Magical Creatures has a 56 kilogram crumple-horned snorkack, a 58 kilogram spiral-horned snorkack, and a 50 kilogram purple-spotted snorkack. Which of these is the most underweight for its species?

- The crumple-horned snorkack.
- The spiral-horned snorkack.
- The purple-spotted snorkack.
- The crumple-horned snorkack and the spiral-horned snorkack are equally underweight.
- Cannot be determined from the information given.
- None of the above.

## 2. SHORT ANSWER PROBLEMS

1. Consider the following scatterplot:



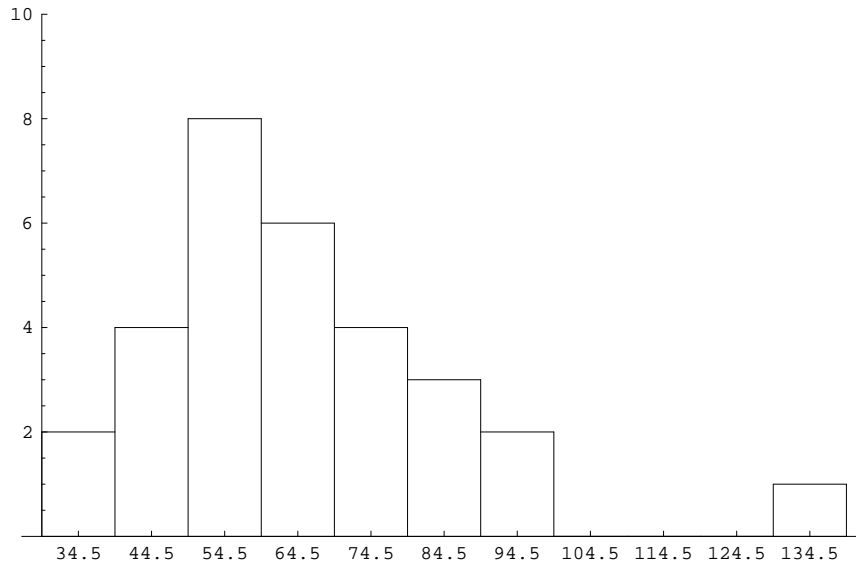
Give the approximate coordinates of all outliers. If there are none, write "NONE".

2. In order to determine the effect of attending review sessions on midterm scores in Math 243, I ask in class for volunteers. For each volunteer, I record how many of the review sessions that person went to and the score on Midterm 1.

- What is the response variable?
- What is the explanatory variable?
- What is the sample?
- What kind of sample did I choose? Circle one.
  - simple random sample
  - explanatory and response sample
  - convenience sample
  - voluntary response sample
  - stratified random sample
  - variable sample

3. I decide to survey my students. Since 22% are seniors, 20% are juniors, 25% are sophomores, 30% are freshman, and 3% are community education students, I decide to take a sample of 22 seniors, 20 juniors, 25 sophomores, 30 freshman, and 3 CEP students. Is this an example of a simple random sample? Why or why not?

4. A small town library kept a record of the number of patrons using the library in the evening during the past 30 days. Below is a histogram of the data. The midpoint of each class is given.



- Give a brief description of the overall shape of this distribution.
- Is the mean larger, smaller, or about the same as the median? Why?
- What was the approximate mean number of evening patrons?
- About what were the highest and lowest number of evening patrons?

5. In a study of the heights of mothers and adult daughters, the correlation was found to be  $r = 0.70$ . The mean and standard deviation of the mothers' heights were 64.0 inches and 2.5 inches, respectively. The mean and standard deviation of the daughters' heights were 66 inches and 3 inches, respectively.

- Which is the reasonable choice for response variable?
- Show that the least-squares regression line is given by  $\hat{y} = 12.24 + 0.84x$ .
- Use the least-squares regression line to predict the height of an adult daughter of a woman who is 61 inches tall.
- Use the least-squares regression line to predict the height of an adult daughter of a woman who is 70 inches tall.
- Carol is 68 inches tall and her adult daughter Joanne is only 64 inches tall. What is the residual for this observation?
- Approximately what percentage of the variation in daughters' heights is explained by the linear relationship? Circle one.
 

99%	95%	70%	50%	30%	0%
-----	-----	-----	-----	-----	----
- Mention at least one lurking variable that you would consider if you wanted to better understand the relationship between mothers' heights and their daughters' heights.

### 3. LONGER ANSWER PROBLEMS

1. You want to test the effectiveness of a new cold medication in reducing nasal congestion. You have 50 test subjects available, each of whom is suffering from nasal congestion as a result of a cold. Use a diagram to outline in detail the design of a randomized comparative double blind experiment. Include information about the treatment groups and the response variable.

2. In a study of car prices and their safety (suitably quantified), it was found that the correlation  $r$  between car prices and safety is  $r \approx 0.13$ . What can you conclude from this? Does this suggest further study of this data set, for example, the calculation of the least squares regression line, etc.? Why?

3. The following data has been collected on five individuals for some explanatory variable  $x$  and some response variable  $y$ .

Individual	A	B	C	D	E
$x$	1	9	22	14	5
$y$	19	17	3	16	18

a. Draw a scatterplot on the axes provided. Be sure to label your axes, and make an appropriate choice of which variable to put on the horizontal axis. [Grid omitted from the sample questions.]

b. Find the equation of the least squares regression line and plot it on the graph above.

c. What is the correlation between  $x$  and  $y$ ? What percentage of the variation in  $y$  is explained by variation in  $x$ ?

d. Use the least squares regression line to predict  $y$  when  $x = 17$ .

e. Explain clearly what an influential point for the regression line is, and identify such a point in this data set.

4. Make a stemplot of the following data:

154 109 137 115 152 140 154 178 101  
103 126 126 137 165 165 129 200 148

5. Make a boxplot of the following data:

154 109 137 115 152 140 154 178 101  
103 126 126 137 165 165 129 200 148

6. Construct a histogram from the following classes of data.

Class	[0, 5)	[5, 10)	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)
Frequency	5	8	14	34	4	9	2

7. Tensile strength is the ability of a material to resist rupture when pressure is applied under specified conditions to one of its sides. In the manufacture of a certain woven polypropylene, the process is operating properly when the standard deviation of the tensile strength is  $\sigma = 4$  pounds per square inch (psi). Measurements of the tensile strength on a random sample of 40 rolls of woven polypropylene produced a mean of 87.3 psi. Find a 98% confidence interval for  $\mu$ , the mean tensile strength of the material. Give the margin of error. Illustrate your answer with a graph, labelling appropriate areas.