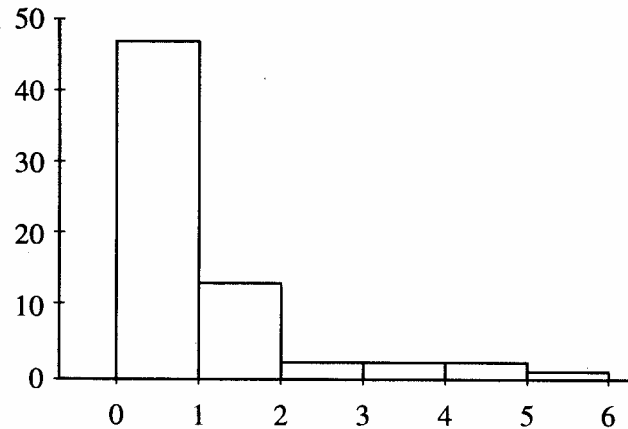
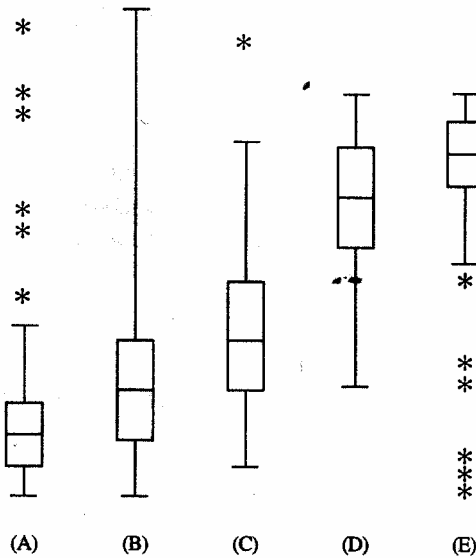


I. Multiple Choice.



1. The histogram above displays a set of measurements. Which of the boxplots below displays the same set of measurements?



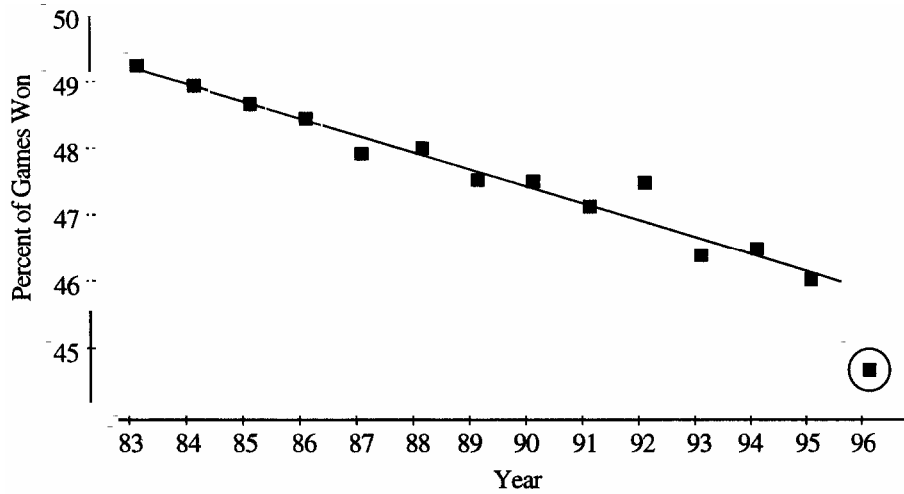
2. A random sample of size 10 was taken from a population. The sample has variance of zero. Which of the following statements must be true?

- I. The population also has a variance of zero.
- II. The sample mean is equal to the sample median.
- III. The ten data points in the sample are equal in numerical value.

- a) I only b) II only c) III only d) I and II e) II and III

3. A professor teaches two statistics classes. The morning class has 25 students and their average on the first test was 82. The evening class has 15 students and their average on the same test was 74. What is the average on this test if the professor combines the scores for both classes?

- a) 76 b) 78 c) 79 d) 80
 e) The average cannot be calculated since individual scores of each student are not available.



4. As shown above, a least-squares regression line has been fitted to the winning percentages for a local sports team in each of the years 1983 through 1995. The percentage for the 1996 season was then plotted (as circled above). Which of the following statements correctly describes how the value for the 1996 season will change the appearance of the least-squares regression line and the correlation coefficient if a new least-squares regression line is fitted to the 1983 through 1996 data?

- a) The 1996 point will make the least-squares regression line steeper and the correlation coefficient stronger.
 b) The 1996 point will make the least-squares regression line steeper and the correlation coefficient weaker.
 c) The 1996 point will make the least-squares regression line closer to horizontal and the correlation coefficient stronger.
 d) The 1996 point will make the least-squares regression line closer to horizontal and the correlation coefficient weaker.
 e) The 1996 point will not have any effect on the least-squares regression line since it follows the same downward trend.

8. A particular psychological test is used to measure academic motivation. The average test score for all female college students nationwide is 115. A large university estimates the mean test score for female students on its campus by testing a random sample of n female students and constructing a confidence interval based on their scores.

Which of the following statements about the confidence interval are true?

- I. The resulting interval will contain 115.
- II. The 95 percent confidence interval for $n = 100$ will generally be shorter than the 95 percent confidence interval for $n = 50$.
- III. For $n = 100$, the 95 percent confidence interval will be longer than the 90 percent confidence interval.

- a) I only
- b) II only
- c) III only
- d) II and III
- e) none of the above gives the complete set of true responses.

9. A survey was conducted at a movie theater to determine movie-goers' preference for different kinds of popcorn. The results of the survey showed that Brand A was preferred by 65 percent of the people with a margin of error of plus or minus 3 percent. What is meant by the statement "plus or minus 3 percent"?

- a) Three percent of the population that was surveyed will change their minds.
- b) Three percent of the time the results of such a survey are not accurate.
- c) Three percent of the population was surveyed.
- d) The true proportion of the population who preferred Brand A popcorn could be determined if 3 percent more of the population was surveyed.
- e) It would be unlikely to get the observed sample proportion of 65 percent unless the actual percentage of people in the population of movie-goers who prefer Brand A is between 62 percent and 68 percent.

10. When performing a test of significance for a null hypothesis, H_0 , against an alternative hypothesis, H_A , the p-value is

- a) the probability that H_0 is true.
- b) the probability that H_A is true.
- c) the probability that H_0 is false.
- d) the probability of observing a value of a test statistic at least as extreme as that observed in the sample if H_0 is true.
- e) the probability of observing a value of a test statistic at least as extreme as that observed in the sample if H_A is true.

11. 20 men and 20 women with high blood pressure were subjects in an experiment to determine the effectiveness of a new drug in lowering blood pressure. 10 of the 20 men and 10 of the 20 women were chosen at random to receive the new drug. The remaining 10 men and 10 women received a placebo. The change in blood pressure was measured for each subject. The design of this experiment is

- a) completely randomized with one factor, drug
- b) completely randomized with one factor, gender
- c) randomized block, blocked by drug and gender
- d) randomized block, blocked by drug
- e) randomized block, blocked by gender

12. A large elementary school has 15 classrooms, with 24 children in each classroom. A sample of 30 children is chosen by the following procedure:
Each of the 15 teachers selects 2 children from his or her classroom to be in the sample by numbering the children from 1 to 24, then using a random digit table to select two different random numbers between 01 and 24. The 2 children with those numbers are in the sample.

Did this procedure give a simple random sample of 30 children from the elementary school?

- a) No, because the teachers were not selected randomly.
- b) No, because not all possible groups of 30 children had the same chance of being chosen.
- c) No, because not all children had the same chance of being chosen.
- d) Yes, because each child had the same chance of being chosen.
- e) Yes, because the number were assigned randomly to the children.

13. The primary reason for using blocking when designing an experiment is to reduce

- a) the sensitivity of the experiment.
- b) variation
- c) the need for randomization
- d) bias
- e) confounding

14. The corn rootworm is a pest that can cause significant damage to corn resulting in a reduction in yield and thus in farm income. A farmer will examine a random sample of plants from a field in order to decide whether or not the number of corn rootworms in the whole field is at a dangerous level. If the farmer concludes that it is, the field will be treated. The farmer is testing the null hypothesis that the number of corn rootworms is not at a dangerous level against the alternative hypothesis that the number is at a dangerous level.

Which of the following is equal to the power of the test?

- a) The probability that the farmer will decide to treat the field.
- b) The probability that the farmer will decide not to treat the field.
- c) The probability that the farmer will fail to reject the null hypothesis.
- d) The probability that the farmer will reject the alternative hypothesis.
- e) The probability that the farmer will not get a statistically significant result.

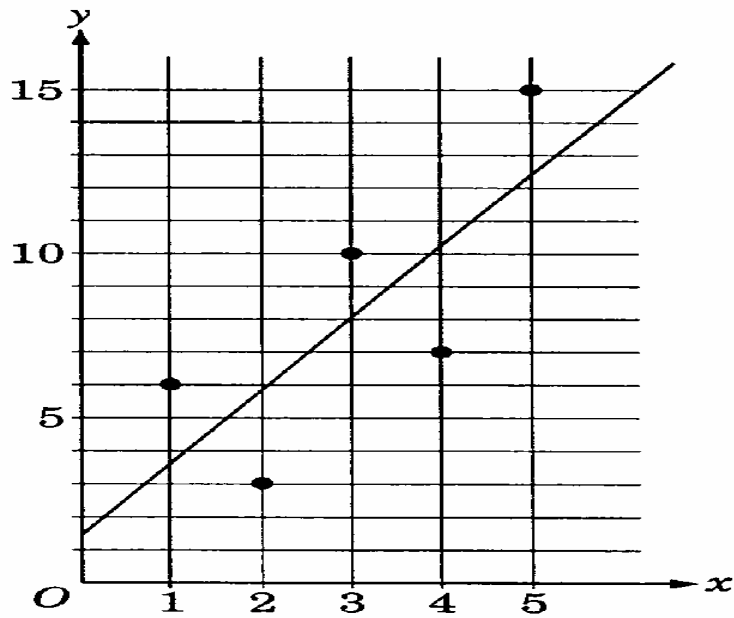
Answers to Multiple-Choice Questions

- | | | | | | | |
|------|------|-------|-------|-------|-------|-------|
| 1. A | 2. E | 3. C | 4. B | 5. A | 6. E | 7. A |
| 8. D | 9. E | 10. D | 11. E | 12. B | 13. B | 14. A |

II. Free Response Questions.

1)

The regression line $\hat{y} = 1.6 + 2.2x$ for the five points on the scatterplot above was computed using the method of least squares. Use the above scatterplot to demonstrate the meaning of the term "least squares."



2. Since most people are right-handed, a company that designs machinery has traditionally place the controls that demand the most hand strength so that they will be used by the right hand. The company decides that it should test its assumption that the right hand of right-handed adults tends to be stronger than the left hand. Nine right-handed adults are selected from employees of this company for the test. Hand strength is measured by using a calibrated hand gripper. It is reasonable to assume that the difference in people's left hand and right hand strength is approximately normally distributed. The people test their left hand first and then their right hand. The hand strengths, in kilograms, for each person and summary statistics are given below.

Person	1	2	3	4	5	6	7	8	9
Right hand	11.7	12.7	11.4	12.4	12.2	11.3	11.1	11.7	11.7
Left hand	11.6	11.2	10.6	11.2	10.9	10.9	10.7	10.3	10.5
Difference	0.1	1.5	0.8	1.2	1.3	0.4	0.4	1.4	1.2

Summary Statistics

	Right hand strength	Left hand strength	Right hand - left hand
Mean	11.8	10.9	0.9
Std. Deviation	0.5	0.4	0.5
Variance	0.28	0.16	0.26

Do these data support the conclusion that the right hand of right-handed adults tends to be stronger than the left hand? Give an appropriate statistical justification. Briefly discuss any concerns you have about the design of this study.

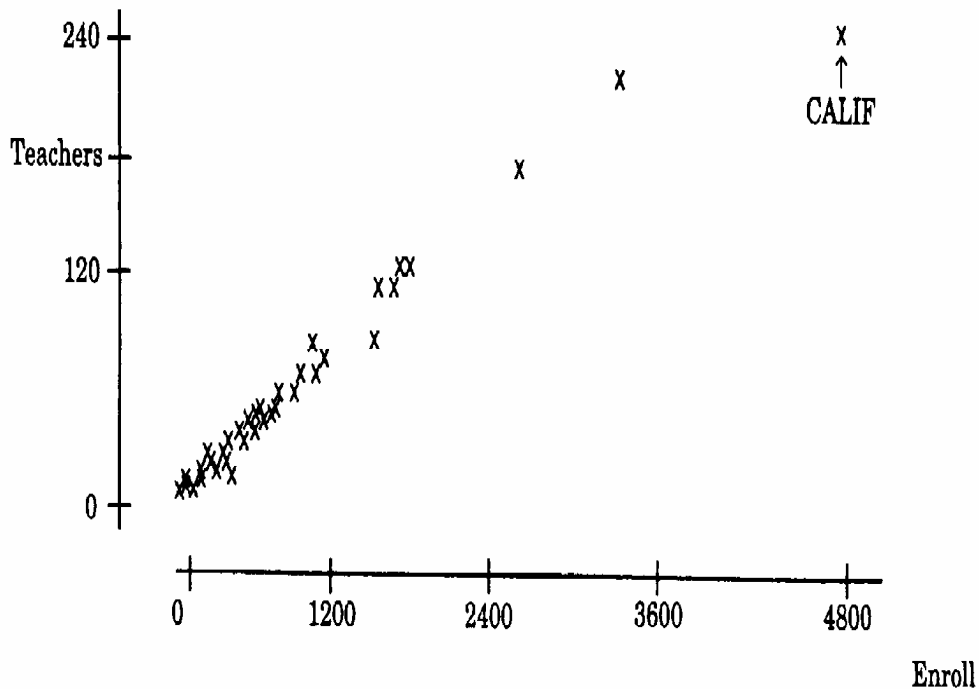
3. Cuckoos lay their eggs in the nest of other birds. Some biologists speculate that the size of the cuckoo's eggs might be different depending on whether the eggs are laid in warblers' nests or wrens' nests. To check this, biologists searched a wildlife refuge for warblers' and wrens' nests. Summary Statistics for the lengths (in mm) of cuckoo's eggs found in the nests of warblers' and wrens' are shown below:

Eggs from warblers' nests: $n_1 = 29$, $\bar{x}_1 = 22.20$, $s_1 = 0.65$

Eggs from wrens' nests: $n_2 = 35$, $\bar{x}_2 = 21.12$, $s_2 = 0.75$

Use a 99 percent confidence interval to determine if these data support the biologists' speculation that the size of the eggs differs depending on whether they are laid in warblers' nests or wrens' nests.

4. The following scatterplot shows the number of teachers (in thousands) for each of the states plus the District of Columbia plotted against the number of students (in thousands) enrolled in grades K-12 for those states and the District of Columbia. The computer printout under the plot shows the results of fitting a straight line to the data by the method of least squares.



The regression equation is
 $\text{Teachers} = 4.49 + 0.0534 \text{ Enroll}$

Predictor	Coef	Stdev	t-ratio	p	
Constant	4.486	2.025	2.22	0.031	
Enroll	0.053401		0.001692	31.57	0.000
s = 10.67	R-sq = 95.3%				

- Find the correlation coefficient for the relationship between the number of teachers and the student enrollment based on these data.
- Provide an interpretation of the slope of this line in the context of these data.
- How will the size of the correlation coefficient change if California is removed from the data set and a new correlation coefficient is calculated? (The data point for California is labeled on the scatterplot.) Explain your answer.
- How will the size of the slope of the least squares regression line change if California is removed from the data set and a new least square regression line is calculated? Explain your answer.

III. Investigative Tasks

1. A company has 11 mathematicians on its staff and only 3 are women. The president of the company is concerned about the small number of women mathematicians. The president learns that about 40 percent of the mathematicians in the United States are women, and asks you to investigate whether or not the number of women mathematicians in the company is consistent with the national pool. The president knows very little about statistics.

You decide to answer the president's question using simulation. You will use the Table of Random Digits to simulate the situation of hiring 11 randomly selected mathematicians (and counting the number of women hired).

- a) Describe how you will use the table to perform one trial of your simulation. Include a description of what each of the digits 0,1,2,3,4,5,6,7,8, and 9 will represent in your simulation.
- b) Perform your simulation 20 times. Start at the left most digit in the first row of the table and move across. Make your procedure clear so that someone can follow what you did. You must do this by marking directly on or above the table.
- c) Place the results of your 20 simulations in a frequency table that show the number of women mathematicians hired.
- d) Based on your frequency table, estimate the probability that there are 3 or fewer women in a randomly selected group of 11 mathematicians.
- e) Based on your simulation, write a report to the president of the company.

2. The table below provides information about the scores on the 1997 and 1998 AP Statistics examination. The data for the 1997 were obtained from all students who took the exam. The data for 1998 were obtained from a random sample of 200 students. You will use the data from this sample to describe how you expect the performance of all students who took the exam in 1998 to differ from the performance of all students who took the exam in 1997.

	1997			1998	
	Examination Grade	Number of students	%	Number of students	%
Extremely well qualified	5	1193	15.6	21	10.5
Well qualified	4	1692	22.1	58	29
Qualified	3	1872	24.5	46	23
Possibly qualified	2	1502	19.6	45	22.5
No recommendation	1	1387	18.1	30	15
Total number of students	7646			200	
Mean grade	2.97			2.975	
Standard Deviation	1.33			1.242	

- Do these data provide evidence of a change in the distribution of grades from 1997 to 1998? Give an appropriate statistical justification to support your conclusion.
- Is there evidence that there is a change in the mean grade from 1997 to 1998? Give appropriate statistical evidence to support your conclusion.
- Using the evidence from this sample, describe how you expect the performance of all student who took the exam in 1998 to differ from the performance of all students who took the exam in 1997.