

STATISTICS

A CALCULATOR MAY BE USED IN BOTH SECTIONS OF THE EXAMINATION. CALCULATORS MAY NOT BE SHARED. REFERENCE MATERIALS MAY NOT BE USED.

Three hours are allotted for this examination: 1 hour and 30 minutes for Section I, which consists of multiple-choice questions, and 1 hour and 30 minutes for Section II, which consists of longer problems. In determining your grade, the two sections are given equal weight. Section I is printed in this examination booklet; Section II, in a separate booklet.

SECTION I

Time—1 hour and 30 minutes

Number of questions—35

Percent of total grade—50

Section I of this examination contains 35 multiple-choice questions. Please be careful to fill in only the ovals that are preceded by numbers 1 through 35 on your answer sheet.

General Instructions

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE INSTRUCTED TO DO SO.

INDICATE ALL YOUR ANSWERS TO QUESTIONS IN SECTION I ON THE SEPARATE ANSWER SHEET ENCLOSED. No credit will be given for anything written in this examination booklet, but you may use the booklet for notes or scratchwork. After you have decided which of the suggested answers is best, COMPLETELY fill in the corresponding oval on the answer sheet. Give only one answer to each question. If you change an answer, be sure that the previous mark is erased completely.

Example:

What is the arithmetic mean of the numbers 1, 3, and 6 ?

- (A) 1 (B) $\frac{7}{3}$ (C) 3
(D) $\frac{10}{3}$ (E) $\frac{7}{2}$

Sample Answer

(A) (B) (C) ☒ (E)

Many candidates wonder whether or not to guess the answers to questions about which they are not certain. In this section of the examination, as a correction for haphazard guessing, one-fourth of the number of questions you answer incorrectly will be subtracted from the number of questions you answer correctly. It is improbable, therefore, that mere guessing will improve your score significantly; it may even lower your score, and it does take time. If, however, you are not sure of the best answer but have some knowledge of the question and are able to eliminate one or more of the answer choices as wrong, your chance of answering correctly is improved, and it may be to your advantage to answer such a question.

Use your time effectively, working as rapidly as you can without losing accuracy. Do not spend too much time on questions that are too difficult. Go on to other questions and come back to the difficult ones later if you have time. It is not expected that everyone will be able to answer all the multiple-choice questions.

Formulas begin on page 12.

Questions begin on page 15.

Tables begin on page 34.

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

$$b_1 = r \frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}}}{\sqrt{\sum (x_i - \bar{x})^2}}$$

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

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

$$E(X) = \mu_x = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum (x_i - \mu_x)^2 p_i$$

If X has a binomial distribution with parameters n and p , then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_x = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If X has a normal distribution with mean μ and standard deviation σ , then:

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

(III) Inferential Statistics

Standardized test statistic: $\frac{\text{estimate} - \text{parameter}}{\text{standard deviation of estimate}}$

Confidence interval: $\text{estimate} \pm (\text{critical value}) \cdot (\text{standard deviation of estimate})$

Single-Sample

Statistic	Standard Deviation
Mean	$\frac{\sigma}{\sqrt{n}}$
Proportion	$\sqrt{\frac{p(1-p)}{n}}$

Two-Sample

Statistic	Standard Deviation
Difference of means (unequal variances)	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
Difference of means (equal variances)	$\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
Difference of proportions (unequal variances)	$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
Difference of proportions (equal variances)	$\sqrt{p(1-p)} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$$\text{Chi-square test statistic} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

STATISTICS

SECTION I

Time—1 hour and 30 minutes

Number of questions—35

Percent of total grade—50

Directions: Solve each of the following problems, using the available space for scratchwork. Decide which is the best of the choices given and fill in the corresponding oval on the answer sheet. No credit will be given for anything written in the test book. Do not spend too much time on any one problem.

1. *USA Today* reported that speed skater Bonnie Blair had “won the USA’s heart,” according to a *USA Today*/CNN/Gallup poll conducted on the final Thursday of the 1994 Winter Olympics. When asked who was the hero of the Olympics, 65 percent of the respondents chose Blair, who won five gold medals. The poll of 615 adults, done by telephone, had a margin of error of 4 percent. Which of the following statements best describes what is meant by the 4 percent margin of error?
- (A) About 4 percent of adults were expected to change their minds between the time of the poll and its publication in *USA Today*.
 - (B) About 4 percent of adults did not have telephones.
 - (C) About 4 percent of the 615 adults polled refused to answer.
 - (D) Not all of the 615 adults knew anything about the Olympics.
 - (E) The difference between the sample percentage and the population percentage is likely to be less than 4 percent.

GO ON TO THE NEXT PAGE

2. An automobile manufacturer claims that the average gas mileage of a new model is 35 miles per gallon (mpg). A consumer group is skeptical of this claim and thinks the manufacturer may be overstating the average gas mileage. If μ represents the true average gas mileage for this new model, which of the following gives the null and alternative hypotheses that the consumer group should test?

- (A) $H_0: \mu < 35$ mpg
 $H_a: \mu \geq 35$ mpg
 - (B) $H_0: \mu \leq 35$ mpg
 $H_a: \mu > 35$ mpg
 - (C) $H_0: \mu = 35$ mpg
 $H_a: \mu > 35$ mpg
 - (D) $H_0: \mu = 35$ mpg
 $H_a: \mu < 35$ mpg
 - (E) $H_0: \mu = 35$ mpg
 $H_a: \mu \neq 35$ mpg
-

3. A magazine has 1,620,000 subscribers, of whom 640,000 are women and 980,000 are men. Thirty percent of the women read the advertisements in the magazine and 50 percent of the men read the advertisements in the magazine. A random sample of 100 subscribers is selected. What is the expected number of subscribers in the sample who read the advertisements?

- (A) 30
- (B) 40
- (C) 42
- (D) 50
- (E) 80

4. A manufacturer makes lightbulbs and claims that their reliability is 98 percent. Reliability is defined to be the proportion of nondefective items that are produced over the long term. If the company's claim is correct, what is the expected number of nondefective lightbulbs in a random sample of 1,000 bulbs?
- (A) 20
(B) 200
(C) 960
(D) 980
(E) 1,000

-
5. When a virus is placed on a tobacco leaf, small lesions appear on the leaf. To compare the mean number of lesions produced by two different strains of virus, one strain is applied to half of each of 8 tobacco leaves, and the other strain is applied to the other half of each leaf. The strain that goes on the right half of the leaf is decided by a coin flip. The lesions that appear on each half are then counted. The data are given below.

<u>Leaf</u>	<u>Strain 1</u>	<u>Strain 2</u>
1	31	18
2	20	17
3	18	14
4	17	11
5	9	10
6	8	7
7	10	5
8	7	6

What is the number of degrees of freedom associated with the appropriate t -test for testing to see if there is a difference between the mean number of lesions per leaf produced by the two strains?

- (A) 7
(B) 8
(C) 11
(D) 14
(E) 16

6. Which of the following is a criterion for choosing a t -test rather than a z -test when making an inference about the mean of a population?

- (A) The standard deviation of the population is unknown.
- (B) The mean of the population is unknown.
- (C) The sample may not have been a simple random sample.
- (D) The population is not normally distributed.
- (E) The sample size is less than 100.

7. A certain county has 1,000 farms. Corn is grown on 100 of these farms but on none of the others. In order to estimate the total farm acreage of corn for the county, two plans are proposed.

- Plan I:
- (a) Sample 20 farms at random.
 - (b) Estimate the mean acreage of corn per farm in a confidence interval.
 - (c) Multiply both ends of the interval by 1,000 to get an interval estimate of the total.

- Plan II:
- (a) Identify the 100 corn-growing farms.
 - (b) Sample 20 corn-growing farms at random.
 - (c) Estimate the mean acreage of corn for corn-growing farms in a confidence interval.
 - (d) Multiply both ends of the interval by 100 to get an interval estimate of the total.

On the basis of the information given, which of the following is the better method for estimating the total farm acreage of corn for the county?

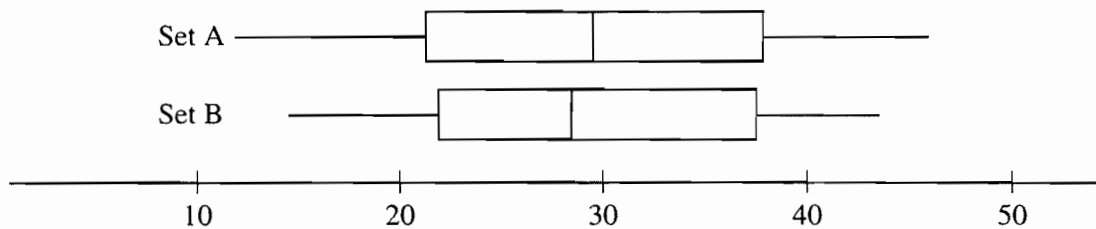
- (A) Choose plan I over plan II.
- (B) Choose plan II over plan I.
- (C) Choose either plan, since both are good and will produce equivalent results.
- (D) Choose neither plan, since neither estimates the total farm acreage of corn.
- (E) The plans cannot be evaluated from the information given.

8. Which of the following can be used to show a cause-and-effect relationship between two variables?

- (A) A census
- (B) A controlled experiment
- (C) An observational study
- (D) A sample survey
- (E) A cross-sectional survey

9. To check the effect of cold temperature on the elasticity of two brands of rubber bands, one box of Brand A and one box of Brand B rubber bands are tested. Ten bands from the Brand A box are placed in a freezer for two hours and ten bands from the Brand B box are kept at room temperature. The amount of stretch before breakage is measured on each rubber band, and the mean for the cold bands is compared to the mean for the others. Is this a good experimental design?

- (A) No, because the means are not proper statistics for comparison.
- (B) No, because more than two brands should be used.
- (C) No, because more temperatures should be used.
- (D) No, because temperature is confounded with brand.
- (E) Yes



10. The boxplots above summarize two data sets, A and B. Which of the following must be true?

- I. Set A contains more data than Set B.
- II. The box of Set A contains more data than the box of Set B.
- III. The data in Set A have a larger range than the data in Set B.

- (A) I only
- (B) III only
- (C) I and II only
- (D) II and III only
- (E) I, II, and III

11. The XYZ Office Supplies Company sells calculators in bulk at wholesale prices, as well as individually at retail prices. Next year's sales depend on market conditions, but executives use probability to find estimates of sales for the coming year. The following tables are estimates for next year's sales.

WHOLESALE SALES

Number Sold	2,000	5,000	10,000	20,000
Probability	0.1	0.3	0.4	0.2

RETAIL SALES

Number Sold	600	1,000	1,500
Probability	0.4	0.5	0.1

What profit does XYZ Office Supplies Company expect to make for the next year if the profit from each calculator sold is \$20 at wholesale and \$30 at retail?

- (A) \$10,590
- (B) \$220,700
- (C) \$264,750
- (D) \$833,100
- (E) \$1,002,500

GO ON TO THE NEXT PAGE

12. The heights of adult women are approximately normally distributed about a mean of 65 inches with a standard deviation of 2 inches. If Rachael is at the 99th percentile in height for adult women, then her height, in inches, is closest to
- (A) 60
(B) 62
(C) 68
(D) 70
(E) 74

-
13. Joe and Matthew plan to visit a bookstore. Based on their previous visits to this bookstore, the probability distributions of the number of books they will buy are given below.

Number of books Joe will buy	0	1	2
Probability	0.50	0.25	0.25

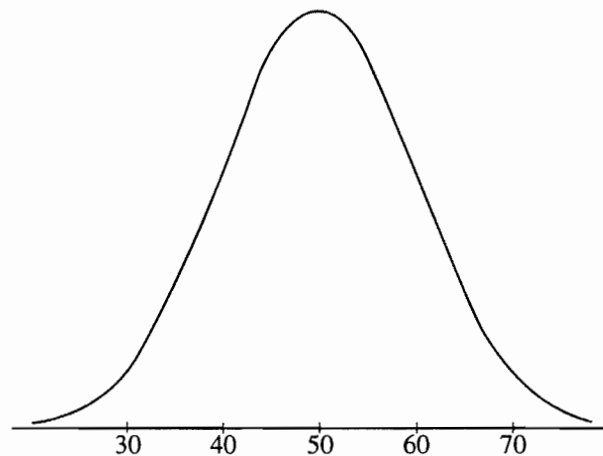
Number of books Matthew will buy	0	1	2
Probability	0.25	0.50	0.25

Assuming that Joe and Matthew make their decisions independently, what is the probability that they will purchase no books on this visit to the bookstore?

- (A) 0.0625
(B) 0.1250
(C) 0.1875
(D) 0.2500
(E) 0.7500

	Job	No Job	Total
Juniors	13	5	18
Seniors	13	26	39
Total	26	31	57

14. A survey of 57 students was conducted to determine whether or not they held jobs outside of school. The two-way table above shows the numbers of students by employment status (job, no job) and class (juniors, seniors). Which of the following best describes the relationship between employment status and class?
- (A) There appears to be no association, since the same number of juniors and seniors have jobs.
 (B) There appears to be no association, since close to half of the students have jobs.
 (C) There appears to be an association, since there are more seniors than juniors in the survey.
 (D) There appears to be an association, since the proportion of juniors having jobs is much larger than the proportion of seniors having jobs.
 (E) A measure of association cannot be determined from these data.



15. Which of the following is the best estimate of the standard deviation of the distribution shown in the figure above?
- (A) 5
 (B) 10
 (C) 30
 (D) 50
 (E) 60

GO ON TO THE NEXT PAGE

16. Ten students were randomly selected from a high school to take part in a program designed to raise their reading comprehension. Each student took a test before and after completing the program. The mean of the differences between the score after the program and the score before the program is 16. It was decided that all students in the school would take part in this program during the next school year. Let μ_A denote the mean score after the program and μ_B denote the mean score before the program for all students in the school. The 95 percent confidence interval estimate of the true mean difference for all students is (9, 23). Which of the following statements is a correct interpretation of this confidence interval?
- (A) $\mu_A > \mu_B$ with probability 0.95.
 - (B) $\mu_A < \mu_B$ with probability 0.95.
 - (C) μ_A is around 23 and μ_B is around 9.
 - (D) For any μ_A and μ_B with $(\mu_A - \mu_B) \geq 14$, the sample result is quite likely.
 - (E) For any μ_A and μ_B with $9 < (\mu_A - \mu_B) < 23$, the sample result is quite likely.
-
17. Gina's doctor told her that the standardized score (z-score) for her systolic blood pressure, as compared to the blood pressure of other women her age, is 1.50. Which of the following is the best interpretation of this standardized score?
- (A) Gina's systolic blood pressure is 150.
 - (B) Gina's systolic blood pressure is 1.50 standard deviations above the average systolic blood pressure of women her age.
 - (C) Gina's systolic blood pressure is 1.50 above the average systolic blood pressure of women her age.
 - (D) Gina's systolic blood pressure is 1.50 times the average systolic blood pressure for women her age.
 - (E) Only 1.5% of women Gina's age have a higher systolic blood pressure than she does.

18. The Physicians' Health Study, a large medical experiment involving 22,000 male physicians, attempted to determine whether aspirin could help prevent heart attacks. In this study, one group of about 11,000 physicians took an aspirin every other day, while a control group took a placebo. After several years, it was determined that the physicians in the group that took aspirin had significantly fewer heart attacks than the physicians in the control group. Which of the following statements explains why it would not be appropriate to say that everyone should take an aspirin every other day?
- I. The study included only physicians, and different results may occur in individuals in other occupations.
 - II. The study included only males and there may be different results for females.
 - III. Although taking aspirin may be helpful in preventing heart attacks, it may be harmful to some other aspects of health.
- (A) I only
(B) II only
(C) III only
(D) II and III only
(E) I, II, and III

Questions 19-20 refer to the following information.

Every Thursday, Matt and Dave's Video Venture has "roll-the-dice" day. A customer may choose to roll two fair dice and rent a second movie for an amount (in cents) equal to the numbers uppermost on the dice, with the larger number first. For example, if the customer rolls a two and a four, a second movie may be rented for \$0.42. If a two and a two are rolled, a second movie may be rented for \$0.22. Let X represent the amount paid for a second movie on roll-the-dice day. The expected value of X is \$0.47 and the standard deviation of X is \$0.15.

19. If a customer rolls the dice and rents a second movie every Thursday for 20 consecutive weeks, what is the total amount that the customer would expect to pay for these second movies?

(A) \$0.45
(B) \$0.47
(C) \$0.67
(D) \$3.00
(E) \$9.40

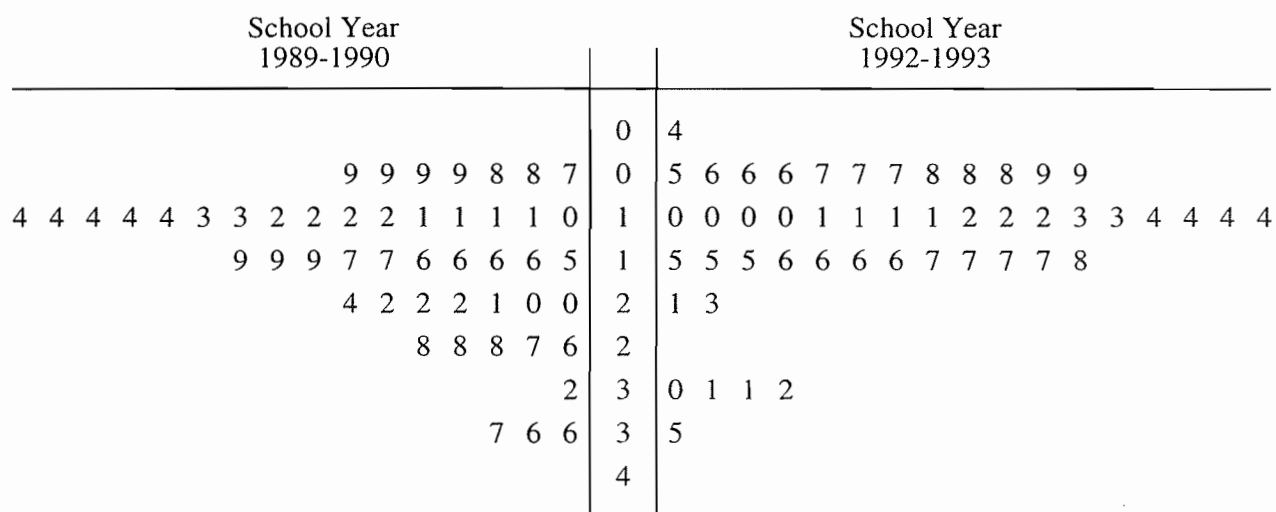
-
20. If a customer rolls the dice and rents a second movie every Thursday for 30 consecutive weeks, what is the approximate probability that the total amount paid for these second movies will exceed \$15.00 ?

(A) 0
(B) 0.09
(C) 0.14
(D) 0.86
(E) 0.91

21. A company wanted to determine the health care costs of its employees. A sample of 25 employees were interviewed and their medical expenses for the previous year were determined. Later the company discovered that the highest medical expense in the sample was mistakenly recorded as 10 times the actual amount. However, after correcting the error, the corrected amount was still greater than or equal to any other medical expense in the sample. Which of the following sample statistics must have remained the same after the correction was made?

(A) Mean
(B) Median
(C) Mode
(D) Range
(E) Variance

22. The back-to-back stem-and-leaf plot below gives the percentage of students who dropped out of school at each of the 49 high schools in a large metropolitan school district.



For 1992-1993, 1|2 represents 12%.

Which of the following statements is NOT justified by these data?

- (A) The drop-out rate decreased in each of the 49 high schools between the 1989-1990 and 1992-1993 school years.
(B) For the school years shown, most students in the 49 high schools did not drop out of high school.
(C) In general, drop-out rates decreased between the 1989-1990 and 1992-1993 school years.
(D) The median drop-out rate of the 49 high schools decreased between the 1989-1990 and 1992-1993 school years.
(E) The spread between the schools with the lowest drop-out rates and those with the highest drop-out rates did not change much between the 1989-1990 and 1992-1993 school years.

23. Circuit boards are assembled by selecting 4 computer chips at random from a large batch of chips. In this batch of chips, 90 percent of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 4 chips from this batch. What is the least probable value of X ?

(A) 0
(B) 1
(C) 2
(D) 3
(E) 4

-
24. A random sample of the costs of repair jobs at a large muffler repair shop produces a mean of \$127.95 and a standard deviation of \$24.03. If the size of this sample is 40, which of the following is an approximate 90 percent confidence interval for the average cost of a repair at this repair shop?

(A) $\$127.95 \pm \4.87
(B) $\$127.95 \pm \6.25
(C) $\$127.95 \pm \7.45
(D) $\$127.95 \pm \30.81
(E) $\$127.95 \pm \39.53

25. At a college the scores on the chemistry final exam are approximately normally distributed, with a mean of 75 and a standard deviation of 12. The scores on the calculus final are also approximately normally distributed, with a mean of 80 and a standard deviation of 8. A student scored 81 on the chemistry final and 84 on the calculus final. Relative to the students in each respective class, in which subject did this student do better?

- (A) The student did better in chemistry.
- (B) The student did better in calculus.
- (C) The student did equally well in each course.
- (D) There is no basis for comparison, since the subjects are different from each other and are in different departments.
- (E) There is not enough information for comparison, because the number of students in each class is not known.

26. A fair coin is flipped 10 times and the number of heads is counted. This procedure of 10 coin flips is repeated 100 times and the results are placed in a frequency table. Which of the frequency tables below is most likely to contain the results from these 100 trials?

(A)

Number of Heads	Frequency
0	19
1	12
2	9
3	6
4	2
5	1
6	3
7	5
8	8
9	14
10	21

(B)

Number of Heads	Frequency
0	9
1	9
2	9
3	9
4	9
5	10
6	9
7	9
8	9
9	9
10	9

(C)

Number of Heads	Frequency
0	0
1	0
2	6
3	9
4	22
5	24
6	18
7	12
8	7
9	2
10	0

(D)

Number of Heads	Frequency
0	7
1	10
2	6
3	11
4	8
5	10
6	9
7	12
8	7
9	11
10	9

(E)

Number of Heads	Frequency
0	0
1	0
2	0
3	2
4	24
5	51
6	22
7	1
8	0
9	0
10	0

GO ON TO THE NEXT PAGE 

27. The student government at a high school wants to conduct a survey of student opinion. It wants to begin with a simple random sample of 60 students. Which of the following survey methods will produce a simple random sample?
- (A) Survey the first 60 students to arrive at school in the morning.
 - (B) Survey every 10th student entering the school library until 60 students are surveyed.
 - (C) Use random numbers to choose 15 each of first-year, second-year, third-year, and fourth-year students.
 - (D) Number the cafeteria seats. Use a table of random numbers to choose seats and interview the students until 60 have been interviewed.
 - (E) Number the students in the official school roster. Use a table of random numbers to choose 60 students from this roster for the survey.
-

28. There is a linear relationship between the number of chirps made by the striped ground cricket and the air temperature. A least squares fit of some data collected by a biologist gives the model

$$\hat{y} = 25.2 + 3.3x \quad 9 < x < 25,$$

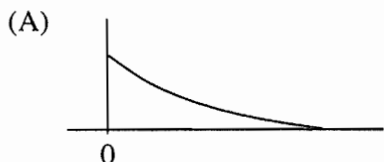
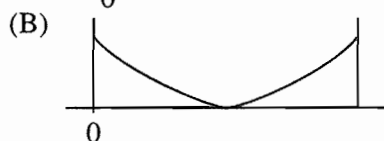
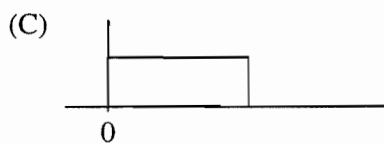
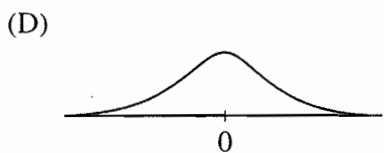
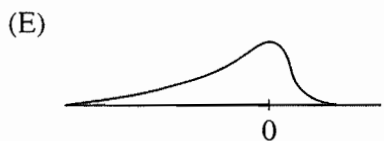
where x is the number of chirps per minute and $\hat{y}$ is the estimated temperature in degrees Fahrenheit. What is the estimated increase in temperature that corresponds to an increase of 5 chirps per minute?

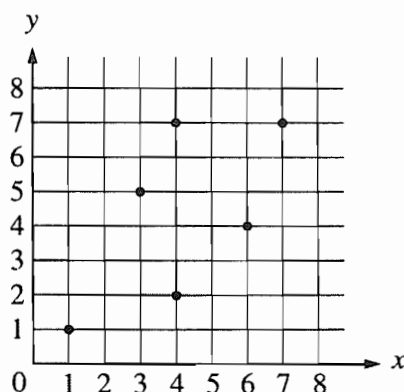
- (A) 3.3° F
- (B) 16.5° F
- (C) 25.2° F
- (D) 28.5° F
- (E) 41.7° F

29. In a test of the null hypothesis $H_0: \mu = 10$ against the alternative hypothesis $H_a: \mu > 10$, a sample from a normal population produces a mean of 13.4. The z -score for the sample is 2.12 and the p -value is 0.017. Based on these statistics, which of the following conclusions could be drawn?

- (A) There is reason to conclude that $\mu > 10$.
- (B) Due to random fluctuation, 48.3 percent of the time a sample produces a mean larger than 10.
- (C) 1.7 percent of the time, rejecting the alternative hypothesis is in error.
- (D) 1.7 percent of the time, the mean is above 10.
- (E) 98.3 percent of the time, the mean is below 10.

30. For which of the following distributions is the mean greater than the median?





31. The equation of the least squares regression line for the points on the scatterplot above is $\hat{y} = 1.3 + 0.73x$. What is the residual for the point (4, 7) ?

(A) 2.78
 (B) 3.00
 (C) 4.00
 (D) 4.22
 (E) 7.00

32. The distribution of the weights of loaves of bread from a certain bakery follows approximately a normal distribution. Based on a very large sample, it was found that 10 percent of the loaves weighed less than 15.34 ounces, and 20 percent of the loaves weighed more than 16.31 ounces. What are the mean and standard deviation of the distribution of the weights of the loaves of bread?

(A) $\mu = 15.82$, $\sigma = 0.48$
 (B) $\mu = 15.82$, $\sigma = 0.69$
 (C) $\mu = 15.87$, $\sigma = 0.50$
 (D) $\mu = 15.93$, $\sigma = 0.46$
 (E) $\mu = 16.00$, $\sigma = 0.50$

33. A 95 percent confidence interval of the form $\hat{p} \pm E$ will be used to obtain an estimate for an unknown population proportion p . If $\hat{p}$ is the sample proportion and E is the margin of error, which of the following is the smallest sample size that will guarantee a margin of error of at most 0.08 ?

(A) 25
(B) 100
(C) 175
(D) 250
(E) 625

-
34. The process of producing pain-reliever tablets yields tablets with varying amounts of the active ingredient. It is claimed that the average amount of active ingredient per tablet is at least 200 milligrams. The Consumer Watchdog Bureau tests a random sample of 70 tablets. The mean content of the active ingredient for this sample is 194.3 milligrams, while the standard deviation is 21 milligrams. What is the approximate p -value for the appropriate test?

(A) 0.012
(B) 0.024
(C) 0.050
(D) 0.100
(E) 0.488

35. A survey was conducted to determine what percentage of college seniors would have chosen to attend a different college if they had known then what they know now. In a random sample of 100 seniors, 34 percent indicated that they would have attended a different college. A 90 percent confidence interval for the percentage of all seniors who would have attended a different college is
- (A) 24.7% to 43.3%
 - (B) 25.8% to 42.2%
 - (C) 26.2% to 41.8%
 - (D) 30.6% to 37.4%
 - (E) 31.2% to 36.8%
-

END OF SECTION I

IF YOU FINISH BEFORE TIME IS CALLED, YOU MAY
CHECK YOUR WORK ON SECTION I.
DO NOT GO ON TO SECTION II UNTIL YOU ARE TOLD TO DO SO.

Table entry for z is the probability lying below z .

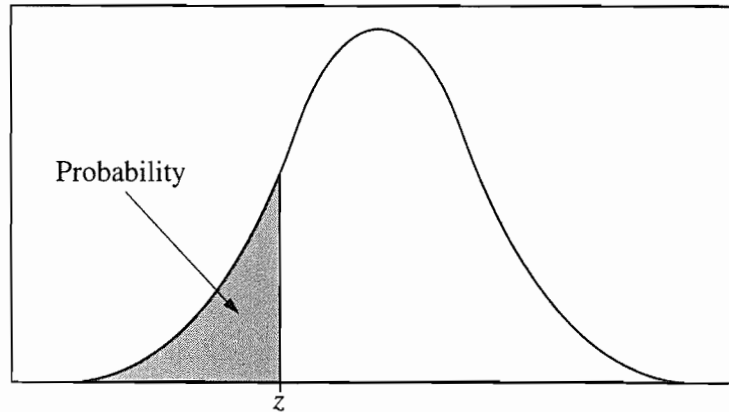


Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

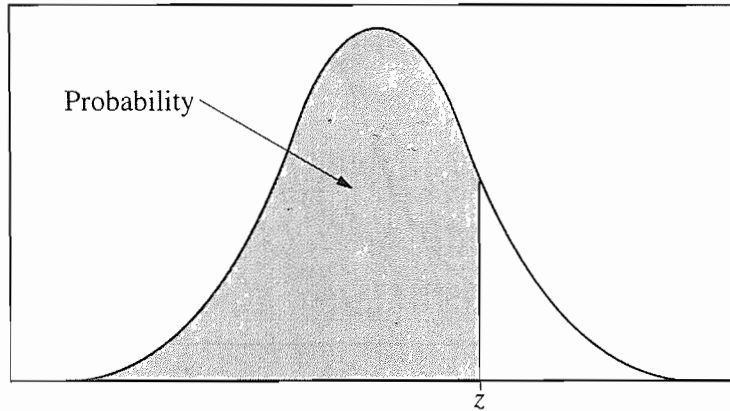


Table entry for z is the probability lying below z .

Table A (Continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^* .

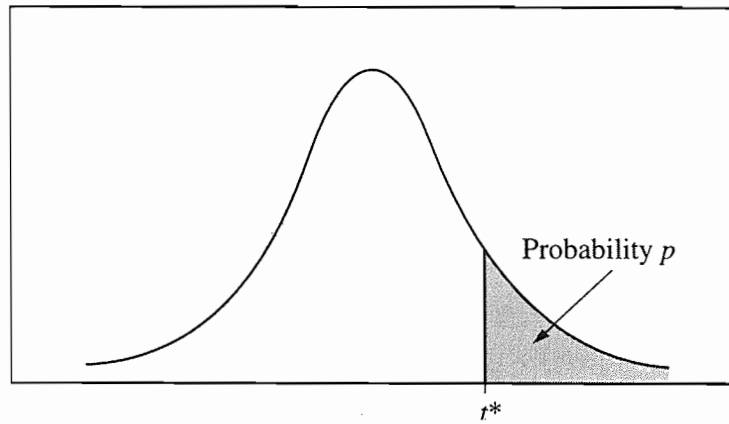


Table B t distribution critical values

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
Confidence level C												

Table entry for p is the point (χ^2) with probability p lying above it.

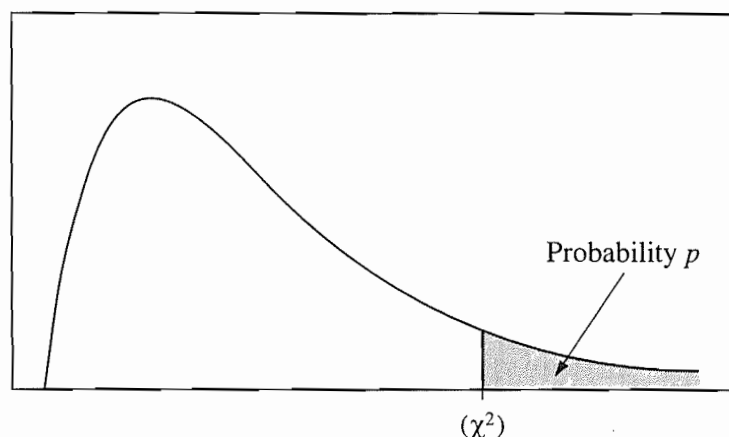


Table C χ^2 critical values

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

STATISTICS

SECTION II

Time — 1 hour and 30 minutes

Number of problems — 6

Percent of total grade — 50

GENERAL INSTRUCTIONS

There are two parts to this section of the examination. Part A consists of five equally weighted problems that represent 75 percent of the total weight of this section. Spend about 65 minutes on this part of the exam. Part B consists of one longer problem that represents 25 percent of the total weight of this section. Spend about 25 minutes on this part of the exam. Since it is not expected that everyone will be able to complete all parts of all problems, you may wish to look over all the problems before you begin to work. The questions are printed in the booklet and on the green insert: it may be easier for you to look over all the problems on the insert. Statistical tables and useful formulas are printed in the green insert. When you are told to begin, open your booklet, carefully tear out the green insert, and start work.

- You should write all work for each part of each problem in the space provided for that part in this booklet. Be sure to write clearly and legibly. If you make an error, you may save time by crossing it out rather than trying to erase it. Erased or crossed-out work will not be graded. No credit will be given for any work shown on the green insert.
- Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as the accuracy of your final answers. Correct answers without supporting work may not receive credit.

STATISTICS

SECTION II

Part A

Questions 1-5

Spend about 65 minutes on this part of the exam.

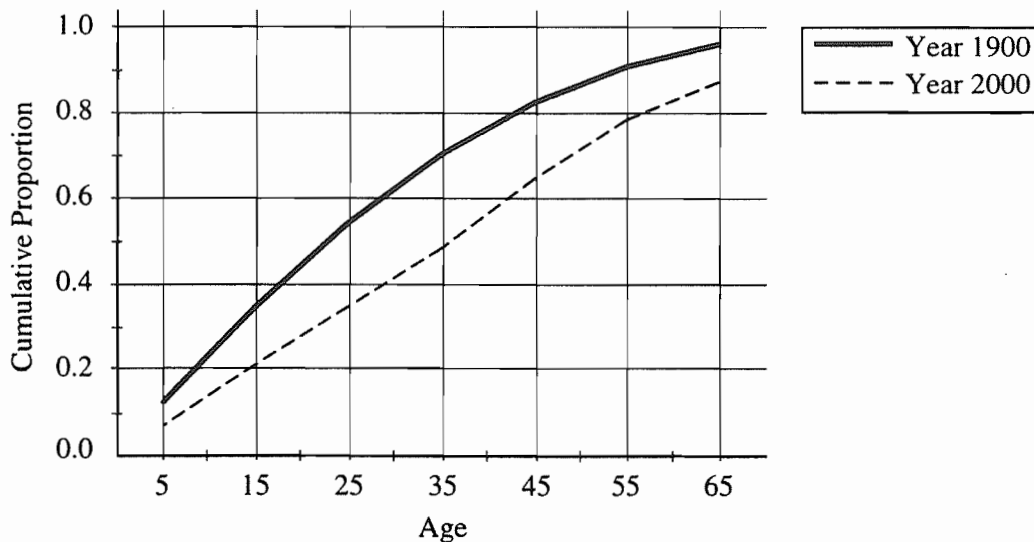
Percent of Section II grade—75

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanation.

AGE DATA

Age	1900	2000
5	0.121	0.066
15	0.344	0.209
25	0.540	0.344
35	0.700	0.480
45	0.822	0.643
55	0.906	0.781
65	0.959	0.870

- The table of data above provides the cumulative proportions for the United States population at selected ages for the years 1900 and 2000 (projected). For example, 0.344 or 34.4 percent of the population was at or under age 15 in 1900, while only 0.209 or 20.9 percent will be at or under age 15 in the year 2000. The graph below shows the cumulative proportions plotted against age for the years 1900 and 2000 (projected). The data and graph are to be used to compare the age distribution for the year 1900 with the projected age distribution for the year 2000.

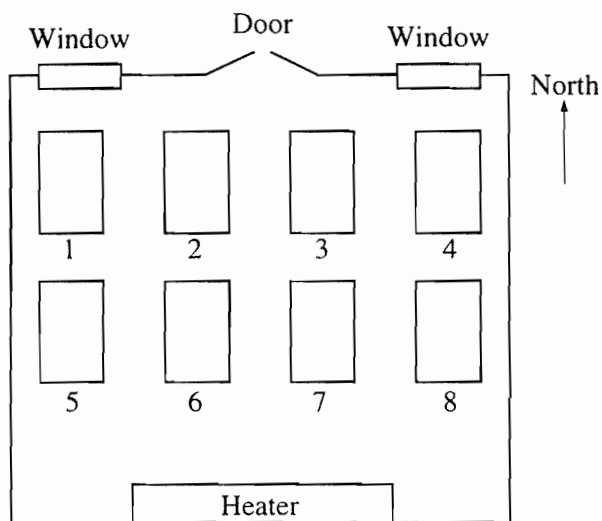


(a) Approximate the median age for each distribution.

(b) Approximate the interquartile range for each distribution.

(c) Using the results from parts (a) and (b), write a sentence or two for a history textbook comparing the age distributions for the years 1900 and 2000.

2. A new type of fish food has become available for salmon raised on fish farms. Your task is to design an experiment to compare the weight gain of salmon raised over a six-month period on the new and the old types of food. The salmon you will use for this experiment have already been randomly placed in eight large tanks in a room that has a considerable temperature gradient. Specifically, tanks on the north side of the room tend to be much colder than those on the south side. The arrangement of tanks is shown on the diagram below.



Describe a design for this experiment that takes account of the temperature gradient.

GO ON TO THE NEXT PAGE

3. A laboratory test for the detection of a certain disease gives a positive result 5 percent of the time for people who do not have the disease. The test gives a negative result 0.3 percent of the time for people who have the disease. Large-scale studies have shown that the disease occurs in about 2 percent of the population.
- (a) What is the probability that a person selected at random would test positive for this disease? Show your work.
- (b) What is the probability that a person selected at random who tests positive for the disease does not have the disease? Show your work.

**GO ON TO THE NEXT PAGE**

4. A random sample of 415 potential voters was interviewed 3 weeks before the start of a state-wide campaign for governor; 223 of the 415 said they favored the new candidate over the incumbent. However, the new candidate made several unfortunate remarks one week before the election. Subsequently, a new random sample of 630 potential voters showed that 317 voters favored the new candidate.

Do these data support the conclusion that there was a decrease in voter support for the new candidate after the unfortunate remarks were made? Give appropriate statistical evidence to support your answer.

5. A company bakes computer chips in two ovens, oven A and oven B. The chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips coming from these ovens for each hour of production throughout a day is shown below.

Percentage of Defective Chips

Hour	Oven A	Oven B
1	45	36
2	32	37
3	34	33
4	31	34
5	35	33
6	37	32
7	31	33
8	30	30
9	27	24

The percentage of defective chips produced each hour by oven A has a mean of 33.56 and a standard deviation of 5.20. The percentage of defective chips produced each hour by oven B has a mean of 32.44 and a standard deviation of 3.78. The hourly differences in percentages for oven A minus oven B have a mean of 1.11 and a standard deviation of 4.28.

Does there appear to be a difference between oven A and oven B with respect to the mean percentages of defective chips produced? Give appropriate statistical evidence to support your answer.

STATISTICS

SECTION II

Part B

Question 6

Spend about 25 minutes on this part of the exam.

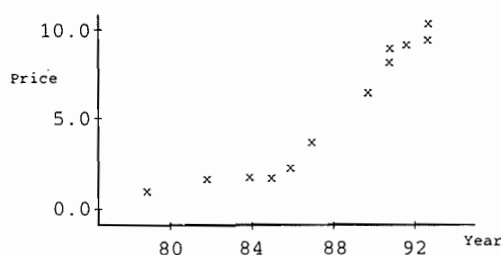
Percent of Section II grade—25

6. You are planning to sell a used 1988 automobile and want to establish an asking price that is competitive with that of other cars of the same make and model that are on the market. A review of newspaper advertisements for used cars yields the following data for 12 different cars of this make and model. You want to fit a least squares regression model to these data for use as a model in establishing the asking price for your car.

Production Year	1990	1991	1992	1987	1993	1991	1993	1985	1984	1982	1986	1979
Asking Price (in thousands of dollars)	6.0	7.7	8.8	3.4	9.8	8.4	8.9	1.5	1.6	1.4	2.0	1.0

The computer printouts for three different linear regression models are shown below. Model 1 fits the asking price as a function of the production year, Model 2 fits the natural logarithm of the asking price as a function of the production year, and Model 3 fits the square root of the asking price as a function of the production year. Each printout also includes a plot of the residuals from the linear model *versus* the fitted values, as well as additional descriptive data produced from the least squares procedure.

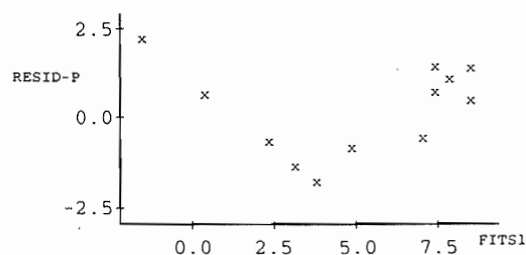
Model 1



The regression equation is
 $\text{Price} = -58.1 + 0.719 \text{ Year}.$

Predictor	Coef	Stdev	t-ratio	p
Constant	-58.050	7.205	-8.06	0.000
Year	0.71900	0.08200	8.77	0.000

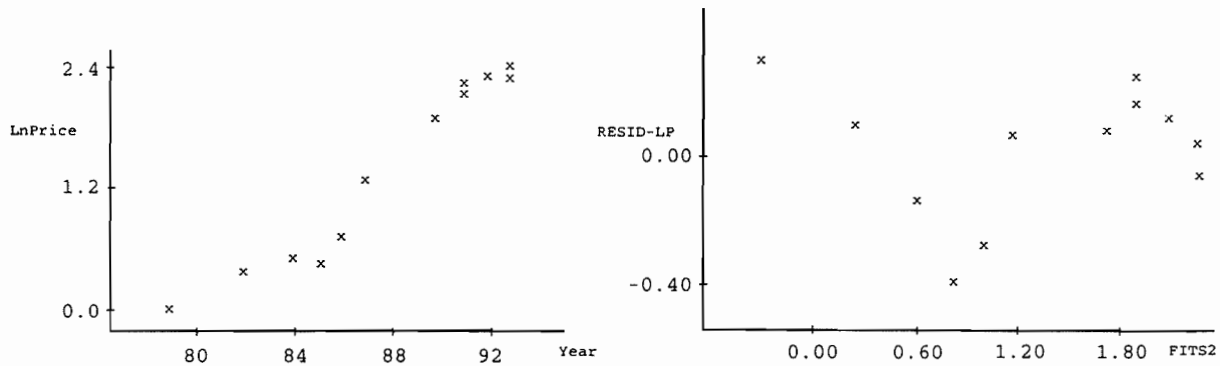
s = 1.255 R-sq = 88.5%



Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	121.10	121.10	76.88	0.000
Error	10	15.75	1.58		
Total	11	136.85			

Model 2



The regression equation is
 $\text{LnPrice} = -14.9 + 0.185 \text{ Year}$.

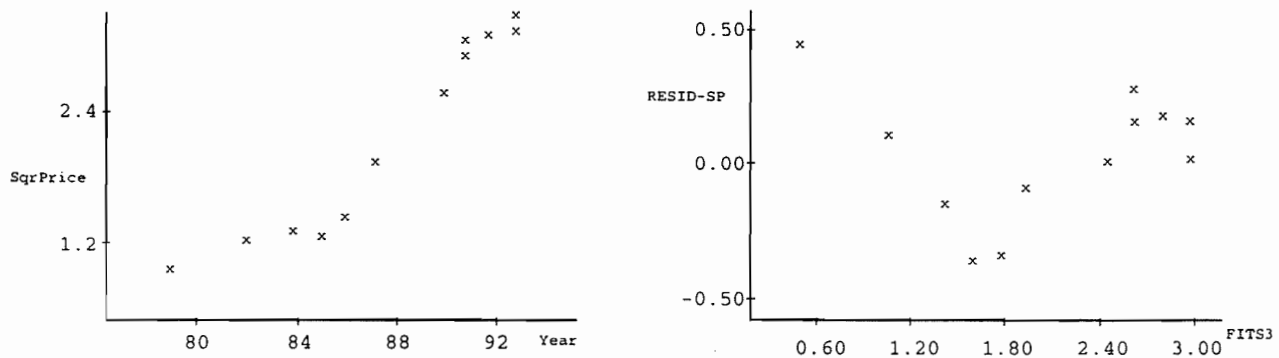
Predictor	Coef	Stdev	t-ratio	p
Constant	-14.924	1.223	-12.21	0.000
Year	0.18502	0.01392	13.30	0.000

s = 0.2130 R-sq = 94.6%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	8.0190	8.0190	176.77	0.000
Error	10	0.4536	0.0454		
Total	11	8.4726			

Model 3



The regression equation is
 $\text{SqrPrice} = -13.3 + 0.176 \text{ Year}$.

Predictor	Coef	Stdev	t-ratio	p
Constant	-13.313	1.447	-9.20	0.000
Year	0.17559	0.01647	10.66	0.000

s = 0.2520 R-sq = 91.9%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	7.2221	7.2221	113.72	0.000
Error	10	0.6351	0.0635		
Total	11	7.8572			

(a) Use Model 1 to establish an asking price for your 1988 automobile.

(b) Use Model 2 to establish an asking price for your 1988 automobile.

(c) Use Model 3 to establish an asking price for your 1988 automobile.

(d) Describe any shortcomings you see in these three models.

(e) Use some or all of the given data to find a better method for establishing an asking price for your 1988 automobile. Explain why your method is better.

If you need more room for your work, use the space below.

END OF EXAMINATION

AP Statistics

1997

APPENDIX - ANSWERS

Matching multiple-choice answers and free-response scoring follow.

Original 1997 format has 35 MCQs and 6 FRQs.

Chapter III

Answers to the 1997 AP Statistics Examination

- Section I: Multiple Choice
 - Blank Answer Sheet
- Section II: Free Response
 - Student Preparation for the Exam
 - Free-Response Questions, Scoring Guidelines, and Sample Student Responses
 - Section II, Part A
 - Section II, Part B

Section I: Multiple Choice

Listed below are the correct answers to the multiple-choice questions and the percentage of AP candidates who answered each question correctly. A copy of the blank answer sheet appears on the following pages for reference.

Section I Answer Key and Percent Answering Correctly

Item No.	Correct Answer	Percent Correct	Item No.	Correct Answer	Percent Correct	Item No.	Correct Answer	Percent Correct
1	E	88%	13	B	73%	25	C	65%
2	D	81%	14	D	73%	26	C	55%
3	C	75%	15	B	83%	27	E	82%
4	D	93%	16	E	69%	28	B	55%
5	A	73%	17	B	74%	29	A	60%
6	A	68%	18	E	63%	30	A	46%
7	B	61%	19	E	91%	31	A	69%
8	B	80%	20	C	36%	32	D	42%
9	D	80%	21	B	69%	33	C	30%
10	B	87%	22	A	56%	34	A	53%
11	B	74%	23	A	68%	35	C	65%
12	D	76%	24	B	70%			

Section II: Free Response

Student Preparation for the Exam

The free-response questions, for the most part, intentionally gave little direction as to how the student was to approach the problem and the rubrics were designed to give credit for any correct method. However, for inference problems the rubrics required students to explain the method used along with the assumptions for the method, and to look carefully at the data to see if they thought the assumptions were satisfied. The analysis had to be followed by a clearly stated correct conclusion, written in the context of the original question.

Many students had difficulty determining what method to use, often trying two or three approaches or mixing approaches in ways that were unclear to the faculty consultants at the Reading. In addition, students had difficulty explaining assumptions and rarely looked at the data to see if assumptions were satisfied. Conclusions were often poorly written both in terms of the statistical concepts and the grammar. Students must learn to view statistical inference as a process that involves five steps:

- 1) understanding the problem to be solved,
- 2) finding an appropriate method,
- 3) checking the assumptions for the method against all the available data,
- 4) carrying out the analysis, and
- 5) writing a clear and correct conclusion in the context of the original problem.

Student preparation on the basic principles of blocking and randomization for designed studies was also quite weak. For question 2, on the design of an experiment, many students appeared to be confused about what constituted the treatments and what constituted the experimental units. Among those who seemed to grasp this idea, many did not understand that blocking is the best way to reduce variation among the experimental units or got confused about how to block appropriately. In addition, many did not understand that randomization is necessary to reduce bias and to allow probability-based methods of analysis to be used after the data are collected.

In summary, the examination contained a fairly easy multiple-choice section and a challenging free-response section. The questions in the latter appeared to be fair, in light of the Course Description, but many students were not adequately prepared to explain and defend methods used and to write clear conclusions reflecting both good statistics and good writing.

Free-Response Questions, Scoring Guidelines, and Sample Student Responses

The answers presented here are actual student responses to the six free-response questions on the 1997 AP Statistics Examination. The students gave permission to have their work reproduced at the time they took the exam. These responses were read and scored by the leaders and faculty consultants assigned to each particular question and were used as sample responses for the training of faculty consultants during the AP Reading in June 1997. The actual scores that these students received, as well as a brief explanation of why, are indicated.

Question 1 Scoring Guidelines

4 Complete Response

Correctly uses the graph or (informally) interpolates in the table to approximate the medians and quartiles. The numerical values must make practical sense within the context of the problem.

Approximates medians for 1900 to be about 23 or 24 and for 2000 about 35 or 36.

Approximates the interquartile ranges for 1900 to be about $40 - 12 = 28$ and for 2000 about $54 - 18 = 36$.

May not calculate the IQR but gives a clear indication in the writing that the distance between the quartiles is a measure of spread.

States clearly that the U.S. age distribution will shift toward larger values for 2000 as compared to 1900, with greater variability in 2000. The statement is consistent with the numerical results from parts (a) and (b) and is written in such a way that it would be understood by the general population.

3 Substantial Response

Approximates medians and interquartile ranges nearly correctly from the graph or the table. Need not calculate the IQR if there is some indication that the quartiles can measure variability.

States clearly that the U.S. age distribution will shift toward larger values for 2000 as compared to 1900, but may not specifically mention the increase in variability.

2 Developing Response

Gives a poor approximation for the medians or interquartile ranges, but seems to have some understanding of what these measures represent.

Or, provides at least a weak but correct statement on the nature of the changing populations.

1 Minimal Response

Indicates some understanding of at least one of the three parts of the question.

Sample 1

(a) Approximate the median age for each distribution.

1900

$15 < \text{median age} < 25$

median age ≈ 22.5 yrs

2000

$35 < \text{median age} < 45$

median age ≈ 36 yrs

Sample 1 (cont.)

(b) Approximate the interquartile range for each distribution.

1900

2000

.25 .75

$10-11 \leq x \leq 39$

$$39 - 11 = \boxed{28 \text{ yrs}}^{\text{size}}$$

$11 \Rightarrow 39$
values

.25 .75

$18 \leq x \leq 53$ ← values

$$53 - 18 = \boxed{35 \text{ yrs}}^{\text{size}}$$

(c) Using the results from parts (a) and (b), write a sentence or two for a history textbook comparing the age distributions for the years 1900 and 2000.

At the turn of the century in 1900, the age of people was much lower than it will be when we hit the year 2000. In 1900 the median age was only 22 ± 23 years old while in the year 2000 it will be about 36. People are living longer now. The middle 50% of people in 1900 were between 11 and 39 years old while in 2000 the middle 50% will be between 18 and 53 years old.

Commentary:

The medians and interquartile ranges are clearly specified for both the 1900 and 2000 data. The student describes the shift toward an older population in 2000 by stating how the median increases and how the spread of the middle 50% also increases. Score = 4

Sample 2

- (a) Approximate the median age for each distribution.

The median age for each distribution would appear at the age where 50% of the population was at or below that age. In the case of 1900, the approx. median age was close to 25 yrs. of age, with 54.4% of the population being at or below the age of 25. For the 2000 distribution the approx. median age was slightly above 35 yrs., with 48% of the population at or below the age of 35.

- (b) Approximate the interquartile range for each distribution.

The interquartile range is equivalent to the range of the ages in the middle 50% of each population distribution. This it can be found by subtracting the 25% age (approx.) from the 75% age (approx.)

estimated from graph... / 1900: approx. 25% - 12 yrs. → 40 yrs - 12 yrs = 28 yrs.
approx. 75% - 40 yrs.

2000: approx. 25% - 20 yrs. → 55 yrs - 20 yrs = 35 yrs.
approx. 75% - 35 yrs.

The approx. IQR for 1900 is 28 yrs., while for 2000, it is 35 yrs.

- (c) Using the results from parts (a) and (b), write a sentence or two for a history textbook comparing the age distributions for the years 1900 and 2000.

In the year 2000, the population, as a whole will be older than it was in the year 1900, as can be based on the medians. The spread of the population ages, will also be fairly different as is demonstrated by the IQRs. In the year 2000, the ages of the population will be spread more widely relative to the population in 1900.

Commentary:

The thinking of this student is clearly stated in words; the reasoning on both the medians and the interquartile ranges is correct. The conclusion concisely states how both the center and spread of the ages increase between 1900 and 2000. Score = 4

Sample 3

- (a) Approximate the median age for each distribution.

The median age for the population of 1900 lies between the ages of 15 and 25, where 50% of the population is above & below the median. This age is contained in $(15, 25)$, roughly 23.

The median age for the 2000 population lies between the ages of 35 and 45, roughly 37.

- (b) Approximate the interquartile range for each distribution.

The IQR is $Q_3 - Q_1$. For 1900, $Q_3 \approx 37$ and $Q_1 \approx 10$. Therefore, the IQR ≈ 27 years

For 2000, $Q_3 \approx 53$ and $Q_1 \approx 17$.
Therefore, the IQR ≈ 36 years

- (c) Using the results from parts (a) and (b), write a sentence or two for a history textbook comparing the age distributions for the years 1900 and 2000.

During the year 1900, the majority of the U.S. population ranged from the ages of 10 to 37, while in the year 2000, the population ranged from 17 to 53 years old. In general, the age of the "average" American citizen has increased dramatically.

Commentary:

Correct reasoning and good approximations are provided for both the medians and the interquartile ranges. Although the conclusion states that ages are increasing, the use of interquartile ranges to express this idea is a bit confused. Both the "majority" and "population ranged" terms are not quite correct. Score = 3

Question 2 Scoring Guidelines

4 Complete Response

Shows four blocks of two tanks each, with the pairs of tanks being in nearly identical conditions with regard to temperature. The blocks are (1,4), (2,3), (5,8), and (6,7).

States clearly that this is the potentially most effective arrangement of blocks since these pairs should be the most homogeneous.

Explains a correct method for randomly assigning treatments to blocks.

3 Substantial Response

Shows two blocks of four tanks each which take into account the north-south temperature gradient. The blocks are (1,2,3,4) and (5,6,7,8).

Or, shows four blocks of two tanks each which take into account the north-south temperature gradient. These blocks are (1,2), (3,4), (5,6), and (7,8).

Or, shows a correct design for the case in which both foods could be used in each tank.

Explains a method of randomly assigning the two treatments to tanks that shows some understanding of randomization.

2 Developing Response

Shows a design that correctly takes temperature into account, but ignores the randomization.

Or, explains a method of randomly assigning the two treatments to tanks and indicates that temperature is to be taken into account in the analysis.

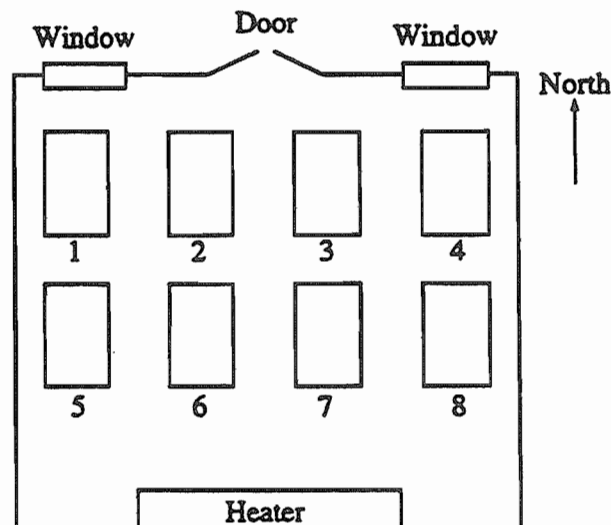
1 Minimal Response

Shows some knowledge of what is meant by blocks, but is not clear on any rationale for choosing blocks or any method of assigning treatments.

Or, states that the two treatments must be randomly assigned to tanks but ignores the temperature gradient.

Sample 1

2. A new type of fish food has become available for salmon raised on fish farms. Your task is to design an experiment to compare the weight gain of salmon raised over a six-month period on the new and the old types of food. The salmon you will use for this experiment have already been randomly placed in eight large tanks in a room that has a considerable temperature gradient. Specifically, tanks on the north side of the room tend to be much colder than those on the south side. The arrangement of tanks is shown on the diagram below.



Sample 1 (cont.)

Describe a design for this experiment that takes account of the temperature gradient.

Because of the room's symmetry, it is possible to match each tank with another that should have very similar, if not equal temperatures. The matchings are 1, 4 2, 3 5, 8 and 6, 7. From each pair, choose one tank to be fed with the new food and one tank to be fed with the old food. This choice should be done randomly (possibly by the toss of a coin).

At both the beginning and end of the six week period, find mean weight of the fish on each type of food.

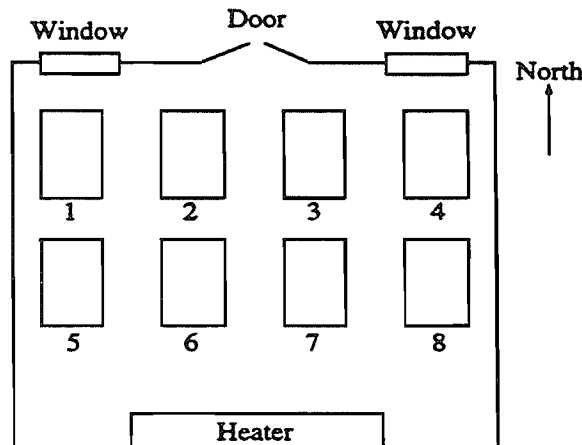
Then calculate the difference in weight gain and determine if it is significant.

Commentary:

The idea of pairing similar tanks to form homogeneous blocks is clearly stated, as is the notion of randomly assigning treatments to the tanks within a block. It is not necessary to discuss the analysis of the resulting data in any detail. *Score = 4*

Sample 2

2. A new type of fish food has become available for salmon raised on fish farms. Your task is to design an experiment to compare the weight gain of salmon raised over a six-month period on the new and the old types of food. The salmon you will use for this experiment have already been randomly placed in eight large tanks in a room that has a considerable temperature gradient. Specifically, tanks on the north side of the room tend to be much colder than those on the south side. The arrangement of tanks is shown on the diagram below.



Describe a design for this experiment that takes account of the temperature gradient.

To account for the temp. gradient, the salmon should be divided into the following groups:

1 + 4
5 + 8
2 + 3
6 + 7

Each fish in these group would be in the relatively same temp. environment as the other fish in the group, based on room location.

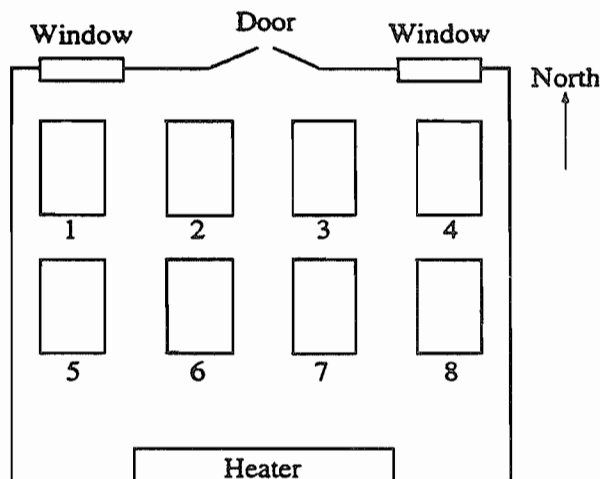
You would now randomly assign 1 fish in each group the new food, and give the other fish the old food. After 6 months, compare the weight gain of each fish in each group. Any differences in the weight gain would have to be attributed to the new food, since temperature differences have been neutralized.

Commentary:

In this response, the block arrangement for tanks is chosen correctly and randomization is used properly. The conclusion is a clear statement on the purpose of blocking and randomization, showing that the student understands the basic principle. Score = 4

Sample 3

2. A new type of fish food has become available for salmon raised on fish farms. Your task is to design an experiment to compare the weight gain of salmon raised over a six-month period on the new and the old types of food. The salmon you will use for this experiment have already been randomly placed in eight large tanks in a room that has a considerable temperature gradient. Specifically, tanks on the north side of the room tend to be much colder than those on the south side. The arrangement of tanks is shown on the diagram below.



Describe a design for this experiment that takes account of the temperature gradient.

The room would be divided into two blocks. Two block would be of tanks 1, 2, 3, and 4 on the north side of the room. The other would be of tanks 5, 6, 7 and 8 on the south side of the room. Starting with tanks 1 and 4, each tank will be randomly assigned either new or old food. Only 2 tanks in each block can receive the same type of food. Therefore, after two of the same types are assigned in a block the other one or two tanks will receive the remaining unused treatment. The fish will be fed once a day and checked for weight gain once every 2 weeks. At the end of the six-month period, they will again be checked for weight gain and the results will be assessed.

Commentary:

Next to the four blocks of two tanks seen on previous responses, this is the second best blocking arrangement. The student correctly explains how randomization will be used. Score = 3.

Question 3 Scoring Guidelines

4 Complete Response

Either by formula or by constructing a table or tree diagram, demonstrates understanding of the three pieces of information given in the problem.

$$P(\text{positive} \mid \text{no disease}) = .05$$

$$P(\text{negative} \mid \text{disease}) = .003$$

$$P(\text{disease}) = .02$$

- a. Correctly reworks that conditional information to answer the unconditional probability as

$$\begin{aligned} P(\text{positive}) &= P(\text{disease}) * P(\text{positive} \mid \text{disease}) + \\ &\quad P(\text{no disease}) * P(\text{positive} \mid \text{no disease}) \\ &= (.02)(.997) + (.98)(.05) = .0689 \end{aligned}$$

- b. Correctly reworks that conditional information to answer the conditional probability as

$$\begin{aligned} P(\text{no disease} \mid \text{positive}) &= \\ &\quad \frac{P(\text{positive} \mid \text{no disease})P(\text{no disease})}{P(\text{positive})} \\ &= \frac{(.05)(.98)}{.0689} = .7112 \end{aligned}$$

May have minor errors in arithmetic or decimal conversions, but the reasoning is correct throughout. The following tree diagram or table may be used to show the structure of the problem.

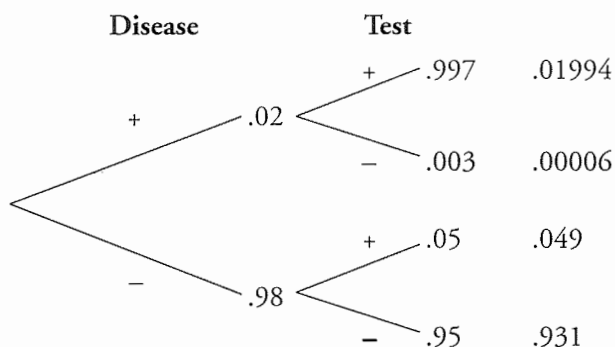


Table of Expected Cell Sizes for a Population of 100

	Disease +	Disease -	Totals
Test +	1.994	4.9	6.894
Test -	.006	93.1	93.106
Totals	2	98	100

$$P(\text{positive test}) = 6.894 / 100 = .0689$$

$$P(\text{no disease} \mid \text{positive test}) = 4.9 / 6.894 = .7108$$

3 Substantial Response

Either by formula or by constructing a table or tree diagram, demonstrates understanding of the three pieces of information given in the problem. Reasons through both parts correctly but may have misread the data in the problem or have errors in arithmetic (e.g., correctly reasons through part a and uses the result in the denominator for the solution to part b but uses incorrect product in the numerator).

2 Developing Response

Reasons through one of parts a or b correctly, but not the other. May have errors in arithmetic.

1 Minimal Response

Shows some understanding of what is being asked for, but does not correctly reason through either part of the problem [e.g., understands the conditional information given or attempts to find $P(\text{no disease} \mid \text{positive})$ or $P(\text{positive})$].

Sample 1

3. A laboratory test for the detection of a certain disease gives a positive result 5 percent of the time for people who do not have the disease. The test gives a negative result 0.3 percent of the time for people who have the disease. Large-scale studies have shown that the disease occurs in about 2 percent of the population.

- (a) What is the probability that a person selected at random would test positive for this disease? Show your work.



$$P(\text{HAVE DISEASE AND POSITIVE RESULT}) \text{ OR } P(\text{DON'T HAVE DISEASE AND POS. RESULT})$$

$$(.02)(.997) + (.98)(.05) = .06894$$

.06894 IS THE PROBABILITY OF TESTING POSITIVE

- (b) What is the probability that a person selected at random who tests positive for the disease does not have the disease? Show your work.

$$P(\text{DOES NOT HAVE DISEASE} \mid \text{POSITIVE RESULT}) = \frac{P(\text{DOESN'T HAVE DISEASE} \cap \text{POSITIVE RESULT})}{P(\text{POSITIVE RESULT})}$$

$$= \frac{(.98)(.05)}{.06894}$$

$$= .71076$$

.71076 IS THE PROBABILITY A PERSON TESTS POSITIVE AND DOES NOT HAVE A DISEASE.

Commentary:

The tree diagram shows that the student understands the structure of the problem. The additive and multiplicative rules of probability are used correctly in (a) and conditional probability is assessed correctly in (b). Score = 4

- The person can either have it & test positive or not have it and test positive!

$$P(\text{positive}) = 0.02 \times 0.997 + 0.98 \times 0.05 = 0.069$$

- This is $\frac{P(\text{no disease}) \cdot P(\text{show pos.})}{P(\text{disease}) \cdot P(\text{show pos.}) + P(\text{no disease}) \cdot P(\text{show pos.})}$

This means that 71.1% of people who test positive for the disease do not have it.. quite a high percentage, but not really avoidable

The student writes out the structure of the events under consideration, thereby showing an understanding of the structure of the problem. The rules of probability are used correctly throughout. *Score = 4*

Sample 3

3. A laboratory test for the detection of a certain disease gives a positive result 5 percent of the time for people who do not have the disease. The test gives a negative result 0.3 percent of the time for people who have the disease. Large-scale studies have shown that the disease occurs in about 2 percent of the population.

- (a) What is the probability that a person selected at random would test positive for this disease? Show your work.

Out of 100 people, 2 on average have the disease, who will give positive results 99.7% of the time. 98 people who don't have the disease (on average) will test positive 5% of the time.

Out of 100 people, the probability of testing positive is

$$\frac{2 \cdot 0.997 + 98 \cdot 0.05}{100} = 0.069$$

- (b) What is the probability that a person selected at random who tests positive for the disease does not have the disease? Show your work.

$$p(\text{not have the disease} \mid \text{test positive for disease}) = \frac{p(\text{not have disease and test pos.})}{p(\text{test positive})}$$

$$= \frac{0.05}{0.069} = 0.725$$

Commentary:

This student correctly assesses what will happen to a typical group of 100 people in (a), but forgets to multiply the numerator in (b) by the probability of not having the disease. Score = 3

Question 4 Scoring Guidelines

4 Complete Response

Recognizes that the problem involves a significance test on the difference between two proportions. Recognizes that the large sample approximation procedure is justified. Sets up correct hypotheses, including a one-sided alternative, with explanations of notation; i.e., enough information is provided to identify symbols with groups.

Identifies and correctly evaluates an appropriate test statistic. The test statistic may be standard normal with either the pooled or unpooled version of the standard error. The test statistic is approximately equivalent for both, $z = 1.08$. The chi-square value is 1.17; the 95% large sample confidence interval for the difference between the two proportions is $(-0.028, 0.096)$.

May make trivial arithmetic errors.

Finds the p-value (or critical value) that is consistent with α, H_a , etc.

States what the significance test allows one to conclude with regard to the original question, using the obtained p-value correctly. Uses statements in the conclusion that are conceptually correct.

3 Substantial Response

Recognizes that the problem involves a significance test on the difference between two proportions, but makes a major error such as:

- fails to address the necessary assumptions
- draws incorrect inference
- does not draw a conclusion or fails to give the conclusion in context

OR

makes two or three less significant errors, such as:

1. Hypotheses errors
2. Notation errors such as not identifying p_1 & p_2 in some manner, not defining non-standard notation, using the same notation for parameters and statistics, etc.
3. Gives $z = 1.08$, but does not identify the test statistic either by name or by formula.

4. Uses a p-value, critical value, confidence interval, etc., that is inconsistent with α, H_a , etc. (includes failing to adjust p-value for chi-square).
5. Uses two confidence intervals and looks for overlap.
6. Uses language such as “accept H_0 ,” “reject H_a ,” “the probability that H_0 is true is .14,” etc. in the conclusion statement.

2 Developing Response

Performs a statistical inference on the two sample proportions but makes two major errors or an equivalent combination of errors.

1 Minimal Response

Gives a statistical inferential argument that incorporates some measure of error or some notion of chance.

Does little more than carry out the mechanics of the test; e.g., uses the calculator to obtain the test statistic.

Possible Complete Response:

$$H_0: p_1 = p_2 \quad H_a: p_1 > p_2$$

where p_1 denotes the proportion of voters favoring the new candidate before the incident and p_2 denotes the proportion of voters favoring the new candidate after the incident.

$$\hat{p}_1 = \frac{223}{415} = .537 \quad \hat{p}_2 = \frac{317}{630} = .503$$

Both samples are large enough that the normal distribution approximation is appropriate.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.537 - .503}{\sqrt{.517(1 - .517)\left(\frac{1}{415} + \frac{1}{630}\right)}} = 1.08$$

$$z = 1.08, \text{ approximately.}$$

$$p\text{-value} = 1 - .8599 = .14$$

While the proportion of voters in the sample supporting the candidate after the unfortunate remarks were made is less than the proportion of voters supporting the candidate before the remarks were made, the difference is not significant and can reasonably be attributed to chance.

Sample 1

4. A random sample of 415 potential voters was interviewed 3 weeks before the start of a state-wide campaign for governor; 223 of the 415 said they favored the new candidate over the incumbent. However, the new candidate made several unfortunate remarks one week before the election. Subsequently, a new random sample of 630 potential voters showed that 317 voters favored the new candidate.

Do these data support the conclusion that there was a decrease in voter support for the new candidate after the unfortunate remarks were made? Give appropriate statistical evidence to support your answer.

The null hypothesis is that the proportion of voters who supported the new candidate before the campaign, $\hat{p}_1$, equals the proportion of voters who supported the new candidate a week before the election, $\hat{p}_2$. The alternative hypothesis is that the proportion of voters who supported the new candidate before the campaign, $\hat{p}_1$, is greater than the proportion of voters who supported the new candidate a week before the election. Since I have two sample proportions, each observation within each sample are independent, and $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ are all greater than or equal to five, I will use a two-proportion Z-test. I found the standard error of the proportion to be
$$SE_p = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = .03155738$$
 since the samples are large and they have unequal variances, I computed the critical z-value to be $\frac{\hat{p}_1 - \hat{p}_2}{SE_p}$ which equaled 1.0817.

The probability that z is greater than 1.0817 is .1398. There is not enough significant evidence to reject the null hypothesis at the 5% level. The new candidate's unfortunate remarks did not significantly affect his support from voters.

Commentary:

The student writes out in great detail exactly what the reasoning process is. The reasoning is very clear and easy to follow. It covers the nature of the problem, the assumptions, the mechanics of the test, and the conclusion, which is stated in the context of the problem. *Score = 4*

Sample 2

4. A random sample of 415 potential voters was interviewed 3 weeks before the start of a state-wide campaign for governor; 223 of the 415 said they favored the new candidate over the incumbent. However, the new candidate made several unfortunate remarks one week before the election. Subsequently, a new random sample of 630 potential voters showed that 317 voters favored the new candidate. $n_1 = 415$ $x_1 = 223$ $n_2 = 630$ $x_2 = 317$

Do these data support the conclusion that there was a decrease in voter support for the new candidate after the unfortunate remarks were made? Give appropriate statistical evidence to support your answer.

$$H_0: p_1 = p_2$$

$p_1 \rightarrow$ proportion of voter support 3 weeks before election

$$H_a: p_1 > p_2$$

$p_2 \rightarrow$ proportion of voter support 1 week before election

a proportion Z test:

$$Z = 1.082$$

$$p\text{-val} = .1397$$

$$\begin{array}{l} x_1 = 223 \\ n_1 = 415 \\ \hat{p}_1 = .5373 \end{array} \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} 3 \text{ weeks} \\ \text{before} \\ \text{election} \end{array}$$

$$\begin{array}{l} x_2 = 317 \\ n_2 = 630 \\ \hat{p}_2 = .5032 \end{array} \left. \begin{array}{l} \\ \\ \end{array} \right\} \begin{array}{l} 1 \text{ week} \\ \text{before} \\ \text{election} \end{array}$$

$$\hat{p} = .5167$$

pooled

Conclusion: do not reject H_0 because $p\text{-val} = .1397$ is not statistically significant at $\alpha = .10$; while there is some evidence against H_0 , there is not enough evidence to conclude that there was a decrease in voter support for the new candidate after the unfortunate remarks were made

test to see that it's safe to use a prop Z test:

$$n_1 \hat{p}_1 > 5 \rightarrow 415(.5373) = 222.98 > 5$$

$$n_1(1 - \hat{p}_1) > 5 \rightarrow 415(.4627) = 192.02 > 5$$

$$n_2 \hat{p}_2 > 5 \rightarrow 630(.5032) = 317.02 > 5$$

$$n_2(1 - \hat{p}_2) > 5 \rightarrow 295.97 > 5$$

most likely that state has $> 10 \times$ sample size

Commentary:

The response clearly covers the type of test being done, the assumptions (with a small error in notation), the test statistic and p-value, and the correct conclusion. The conclusion provides a statement in the context of the problem. Score = 4

Sample 3

4. A random sample of 415 potential voters was interviewed 3 weeks before the start of a state-wide campaign for governor; 223 of the 415 said they favored the new candidate over the incumbent. However, the new candidate made several unfortunate remarks one week before the election. Subsequently, a new random sample of 630 potential voters showed that 317 voters favored the new candidate.

Do these data support the conclusion that there was a decrease in voter support for the new candidate after the unfortunate remarks were made? Give appropriate statistical evidence to support your answer.

H_0 : There was no change in voter support ~~after~~ after the remarks $p_1 = p_2 = 0$

H_a : There was a decrease in voter support after the remarks $p_1 = p_2 > 0$

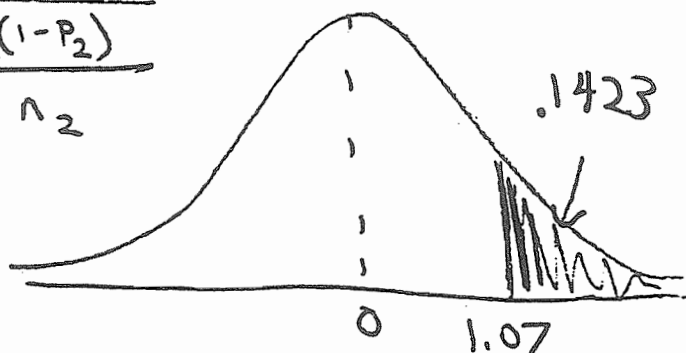
$$\hat{p}_1 = .537\% = \frac{223}{415} \quad \hat{p}_2 = .503\% = \frac{317}{630}$$

$$Z^* = \frac{(p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

$$Z^* = 1.07$$

$$p^* = .1423$$

$$\alpha = .05$$



Based on lack of sufficient evidence, fail to reject H_0 that there was no change in voter support after the remarks the new candidate made.

Commentary:

The student carries out the correct test and makes a proper conclusion, but some of the symbols are not defined and no mention is made of assumptions. Score = 3

Question 5 Scoring Guidelines

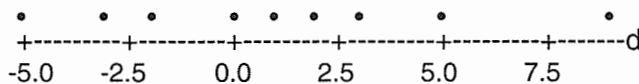
4 Complete Response

- A. Selects an appropriate test of significance.

Observes that data are paired by hour of day and indicates a paired t-test is appropriate OR observes that the data are paired by hour of day and justifies choosing a two-sample t-test (or confidence interval for the difference of two proportions) OR worries about assumptions for a t-test and so chooses a sign test OR performs a substantial data analysis.

- B. States and addresses assumptions necessary for t-test.

The dot plot below shows no evidence that the differences aren't normally distributed.



(For a two-sample t-test, mentions that outliers indicate a potential problem.)

- C. Carries out a test of significance.

States the hypotheses correctly, with two-sided alternative, defining any nonstandard symbols.

$$H_0 : \mu_A = \mu_B \text{ or } \mu_d = 0$$

$$H_a : \mu_A \neq \mu_B \text{ or } \mu_d \neq 0$$

Calculates an appropriate test statistic, possibly with trivial arithmetic error.

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{1.11}{4.28 / \sqrt{9}} = 0.78$$

(For two-sample t-test, $t = 0.519$ and $p = 0.61$, pooled or unpooled. For a sign test, $++++-0+$ so $P = 0.727$.)

Correctly approximates the P-value from a table or calculator.

With 8 degrees of freedom, the two-sided P-value is 0.46 (OR compares $t = 0.78$ with critical value of 2.306).

OR constructs and interprets a 95% confidence interval for the mean difference:

$$\bar{d} \pm t^* \frac{s_d}{\sqrt{n}} = 1.11 \pm (2.306) \left(\frac{4.28}{\sqrt{9}} \right) = 1.11 \pm 3.29 \text{ or } (-2.18, 4.40)$$

OR a 95% confidence interval for the difference of two means: $(-3.47, 5.69)$

- D. Correctly states a conclusion in the context of the situation.

The difference in the mean percentages in the samples of defective chips produced by ovens A and B isn't statistically significant. If there was no difference in the percentages, a mean absolute difference of 1.11 or greater would happen 46% of the time with random samples of size 9.

Or, for the confidence interval: Since the confidence interval includes zero, there is no evidence that one oven differs from the other in terms of mean percentage of defective chips.

3 Substantial Response

Performs a test of significance but

Doesn't mention the assumptions for the test used

OR

States an incorrect conclusion

OR

Makes two or three less significant omissions or errors in the test of significance, such as

- doesn't make it clear by words or formula which test they are using
- chooses a two-sample t-test or a sign test without considering the paired t-test
- fails to list hypotheses
- does a one-tailed test or forgets to double the P-value
- says must use a t-test because n is small
- uses z-test, not t-test
- mentions normality assumption but doesn't check it graphically or appeal to robustness
- confuses parameters and statistics
- doesn't define nonstandard symbols

Question 5 Scoring Guidelines (cont.)

- forgets $\frac{1}{\sqrt{n}}$ in formula for t
- uses wrong critical value or degrees of freedom
- "accepts" the null hypothesis or uses similar careless language in conclusion
- gives correct, but minimal, conclusion ("do not reject H_0 ") or a generic conclusion that doesn't relate to the ovens

1 Minimal Response

Recognizes that the problem involves an inferential statistical argument.

Does little more than carry out the mechanics of a test of significance, such as those done by a calculator.

Note: Since we are using holistic grading, a deficiency in one part can be compensated for by outstanding work in another.

2 Developing Response

Performs test of significance except makes both major errors OR an equivalent combination of errors.

Sample 1

5. A company bakes computer chips in two ovens, oven A and oven B. The chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips coming from these ovens for each hour of production throughout a day is shown below.

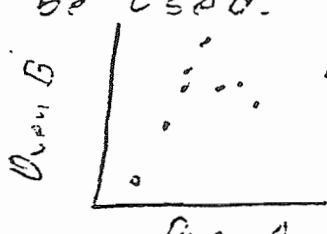
Percentage of Defective Chips

Hour	Oven A	Oven B
1	45	36
2	32	37
3	34	33
4	31	34
5	35	33
6	37	32
7	31	33
8	30	30
9	27	24

The percentage of defective chips produced each hour by oven A has a mean of 33.56 and a standard deviation of 5.20. The percentage of defective chips produced each hour by oven B has a mean of 32.44 and a standard deviation of 3.78. The hourly differences in percentages for oven A minus oven B have a mean of 1.11 and a standard deviation of 4.28.

Does there appear to be a difference between oven A and oven B with respect to the mean percentages of defective chips produced? Give appropriate statistical evidence to support your answer.

To test for a difference between oven A and oven B defective chips, a matched pairs design will be used. The reason is a scatter plot of Δ shows a good positive association of defective chips. The probability



Sample 1 (cont.)

of obtaining such a high linearity is below .05 according to a linear regression test. If dependence exists, a two sample t-test cannot be done. Thus, the null hypothesis is that the ^{mean of paired} differences of the percent defective chips is 0. The alternative hypothesis is that the mean difference is not equal to 0. A standard alpha of .05 will be used.

The general equation is a one sample t test upon the difference. or $t = \frac{\bar{x} - 0}{s/\sqrt{n}}$

The degrees of freedom is 9-1 or 8.

Substituting $t = \frac{7.77 - 0}{4.78/\sqrt{10}}$ or $t = .7780$

The probability of obtaining such a t statistic or $\geq P(t \geq .7780)$ give 8 degrees of freedom is .4588. This probability value is greater than alpha. The null hypothesis, no difference exists between Oven A and Oven B for the proportion of defective chips. must be accepted.

Commentary:

The response provides a detailed explanation of why the data are regarded as paired (although the regression part is not essential and even plots the data. The paired t-test is carried out and a correct conclusion is stated in terms of a comparison between the two ovens. (It would be good to plot the differences to show that there are no outliers.) Score = 4

Sample 2

5. A company bakes computer chips in two ovens, oven A and oven B. The chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips coming from these ovens for each hour of production throughout a day is shown below.

Percentage of Defective Chips

Hour	Oven A	Oven B
1	45	36
2	32	37
3	34	33
4	31	34
5	35	33
6	37	32
7	31	33
8	30	30
9	27	24

The percentage of defective chips produced each hour by oven A has a mean of 33.56 and a standard deviation of 5.20. The percentage of defective chips produced each hour by oven B has a mean of 32.44 and a standard deviation of 3.78. The hourly differences in percentages for oven A minus oven B have a mean of 1.11 and a standard deviation of 4.28.

Does there appear to be a difference between oven A and oven B with respect to the mean percentages of defective chips produced? Give appropriate statistical evidence to support your answer.

The data is given to us in matched pairs form, matched by hour. To analyze matched pairs data, I will use a one-sample t-test. On the differences, I do this to account for the non-independence that matched pairs has.

$$H_0: \mu_{diff} = 0 \quad H_a: \mu_{diff} \neq 0$$

$$\bar{x}_{diff} = 1.11$$

$$s = 4.28$$

$$n = 9$$

$$\alpha = .05$$

$$df = 8$$

$$t = \frac{\bar{x} - H_0}{\frac{s}{\sqrt{n}}} = \frac{1.11 - 0}{\frac{4.28}{\sqrt{9}}} = .778$$

$$\text{critical } t_{\alpha=0.05, df=8} = 2.306$$

$$2.3067.778$$

Because the critical t is less than the t statistic, I cannot reject H_0 in favor of H_A . I have no evidence to suggest a difference in the percentage of baked chips leaving the oven.

Commentary:

The student correctly recognizes this as a matched pairs problem and analyzes the data accordingly. The conclusion is correctly stated in context. Although assumptions are not mentioned, the student did look at the differences, as indicated by the numbers added to the chart. *Score = 4*

Sample 3

5. A company bakes computer chips in two ovens, oven A and oven B. The chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips coming from these ovens for each hour of production throughout a day is shown below.

Percentage of Defective Chips

Hour	Oven A	Oven B
1	45	36
2	32	37
3	34	33
4	31	34
5	35	33
6	37	32
7	31	33
8	30	30
9	27	24

The percentage of defective chips produced each hour by oven A has a mean of 33.56 and a standard deviation of 5.20. The percentage of defective chips produced each hour by oven B has a mean of 32.44 and a standard deviation of 3.78. The hourly differences in percentages for oven A minus oven B have a mean of 1.11 and a standard deviation of 4.28.

Does there appear to be a difference between oven A and oven B with respect to the mean percentages of defective chips produced? Give appropriate statistical evidence to support your answer.

See student response on facing page.

Sample 3 (cont.)

$H_0: \mu_A = \mu_B$ $\mu_A =$ percentage of defective chips per oven A
 $H_a: \mu_A \neq \mu_B$ $\mu_B =$ percentage of defective chips per oven B

$$t = \frac{33.56 - 32.44}{\sqrt{\frac{5.20^2}{9} + \frac{3.78^2}{9}}} = .5226548$$

$df = 14.6092 \leftarrow$ calculated from TI-82

$p = .609015$ (from TI-82 with $df = 14.6092$) $\therefore$
 $\alpha = .61538$ (with $df = 8$)

There is not enough evidence to reject H_0 in favor of H_a there for there is no difference between the percentage of defective chips between ovens A and B

*Note data for oven A is skewed right and since sample size is less than 15 it may effect the t-procedures. We should use a larger sample size.

Commentary:

Here, the data are analyzed by a two-sample t-test on the difference between two means. Given the assumption of two independent samples, the analysis and conclusion are correct. The student correctly notes that there is some skewness in the data. Score = 3

Question 6 (Investigative Task) Scoring Guidelines

4 Complete Response

- I. Correctly estimates the asking price in dollars for at least two of the three models, including back transforming the predicted value for at least one of models b or c.
- a. $\text{Price} = -58.1 + 0.719(88) = 5.172$
asking price is \$5,172
- b. $\text{Ln}(\text{price}) = -14.9 + 0.185(88) = 1.380$
 $\exp(1.38) = 3.9749$
asking price = \$3,975
- c. $\text{Sqrt}(\text{price}) = -13.3 + 0.176(88) = 2.188$
 $(2.188)^2 = 4.7873$
asking price = \$4,787
- II. Describes the major shortcoming to be the non-linear pattern in the scatterplot or residual plot for all three models.
- III. Suggests a new model that successfully deals with the non-linearity of the data. The prime contenders for this model are:
- Fit a simple linear model after the first two or three years are dropped.
($R^2 = .978$ with little pattern in the residuals after dropping 79, 82 and 84.)
- Fit separate linear models to the earlier and later years.
($R^2 = .985$ for the years 86 to 93, with little pattern in the residuals.)
- Fit a model that attempts to model the curvature in the data. For example, fit a quadratic model to all of the data.
($R^2 = .974$ with little pattern in the residuals for a quadratic model.)
- IV. Justifies why this model is better. For example, comments that there is no pattern in the residuals as seen from looking at the fitted model on the scatterplot or from looking at the residual plot.

3 Substantial Response

Fails to do one of the following satisfactorily:

- Part I (for example, by failing to back transform the prediction)
- describe the major shortcomings of the three models
- justify why the new model is better (for example, by failing to comment on the residual plot)

2 Developing Response

Fails to do one of the following satisfactorily:

- Part I and describe the major shortcomings of the three models
- Part I and justify why the new model is better
- suggest a new model and justify it
- describe the major shortcomings of the three models and justify why the new model is better.

1 Minimal Response

Gives correct responses to one of the items I, II, or III above. **Note:** Students cannot get IV correct without first specifying a new model.

Sample 1

(a) Use Model 1 to establish an asking price for your 1988 automobile.

$$\text{Price} = -58.1 + 0.719(88) \approx \boxed{\$5,170}$$

(b) Use Model 2 to establish an asking price for your 1988 automobile.

$$\ln \text{Price} = -14.9 + 0.185(88) = 1.38$$

$$\text{Price} = e^{1.38} \approx \boxed{\$3,970}$$

(c) Use Model 3 to establish an asking price for your 1988 automobile.

$$\sqrt{\text{Price}} = -13.3 + 0.176(88) = 2.188$$

$$\text{Price} = 2.188^2 \approx \boxed{\$4,790}$$

Sample 1 (cont.)

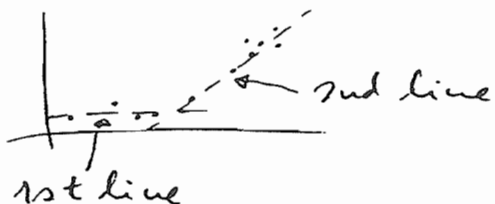
(d) Describe any shortcomings you see in these three models.

Although the correlation coefficients are all quite large, the residual plots all show a distinct pattern, which means that neither fit is the correct one for the data.

(e) Use some or all of the given data to find a better method for establishing an asking price for your 1988 automobile. Explain why your method is better.

From the plot of Model 1 it is clear that a better fit for the data would be two lines instead of one. The first one would be the least-squares regression line on the years prior to 1986 and would be valid only for those years, while the other one would be the regression line on the years of 1986 and later, and would be valid for those years.

In this case, any pattern in the residual plot would disappear, and the asking price of the 1988 automobile could be determined from the second line of the fit.



Commentary:

Parts (a), (b), and (c) show correct calculations and (d) contains a comment on the pattern in the residual plots. Part (e) contains a clear description of how the data could be split into two groups, each showing a linear trend but with different slopes. Score = 4

Sample 2

- (a) Use Model 1 to establish an asking price for your 1988 automobile.

$$\begin{aligned}\text{price} &= -58.1 + .719(\text{year}) \\ &= -58.1 + .719 \cdot 88 \\ &= 5.172 (\text{thousand dollars}) \\ \therefore \text{price} &= \boxed{\$5172}\end{aligned}$$

- (b) Use Model 2 to establish an asking price for your 1988 automobile.

$$\begin{aligned}\ln \text{price} &= -14.9 + .185 \cdot \text{year} \\ &= -14.9 + .185 \cdot 88 \\ \ln p &= 1.38 \\ p &= e^{1.38} = 3.975 (\text{thousand \$}) \\ \therefore \text{price} &= \boxed{\$3975}\end{aligned}$$

- (c) Use Model 3 to establish an asking price for your 1988 automobile.

$$\begin{aligned}\sqrt{\text{price}} &= -13.3 + .176 \cdot \text{year} \\ &= -13.3 + .176 \cdot 88 \\ &= 2.188 \\ \text{price} &= 2.188^2 = 4.787 (\text{thousand \$}) \\ \therefore \text{price} &= \boxed{\$4787}\end{aligned}$$

Sample 2 (cont.)

(d) Describe any shortcomings you see in these three models.

The residuals, which should be random in the best regression, exhibit curved patterns on all three graphs. Also, r^2 , which should be at least .97 or .98, is .946 or lower on every graph.

(ln price vs. ln year does not correct this. I tried.)

(e) Use some or all of the given data to find a better method for establishing an asking price for your 1988 automobile. Explain why your method is better.

Eliminate from our data pool all cars more than three years away from our 1988 car: before 1985 or after 1991. Then, we get a data set with the following graph (sketched from calculator):

price | and a least squares regression of $[\text{price} = -92 + 1.096(\text{year})]$.



The plot of the residuals is



which is fairly random. Substituting in our year, we get a price of $\text{price} = -92 + 1.096(88) = 4.448 = \4448 which seems to correspond to nearby years fairly well.

Commentary:

The numerical calculations in the first three parts are correct. The pattern in the residuals is recognized. The solution to improve the fit is to concentrate on the years around 1988, and this is a reasonable way to improve the fit. Score = 4

Sample 3

(a) Use Model 1 to establish an asking price for your 1988 automobile.

$$f(88) = \underline{5.172 \text{ thousand dollars}}$$

$$-58.1 + 0.719 \cdot 88$$

(b) Use Model 2 to establish an asking price for your 1988 automobile.

$$f(88) = 1.38$$

$$e^{1.38} = \underline{3.9749 \text{ thousand dollars}}$$

$$-14.9 + 0.185 \cdot 88 =$$

(c) Use Model 3 to establish an asking price for your 1988 automobile.

$$-13.3 + 0.176 \cdot 88 = 2.188 \quad 2.188^2 = \underline{4.7873 \text{ thousand dollars}}$$