

STATISTICS
SECTION I

Time—1 hour and 30 minutes
Number of questions—35
Percent of total grade—50

1997 exam

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.

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

15

SECTION II

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.

Leaf	Strain 1	Strain 2
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

GO ON TO THE NEXT PAGE

17

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

GO ON TO THE NEXT PAGE

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.

GO ON TO THE NEXT PAGE

18

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

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

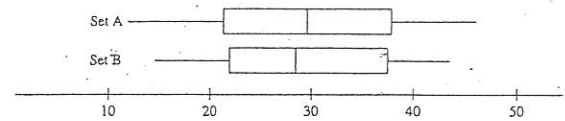
13. Joe and Matthew plan to visit a bookstore. Based on their previous visits to this bookstore, the probability distributions of the number of books they will buy are given below.

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

GO ON TO THE NEXT PAGE



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.

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

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?

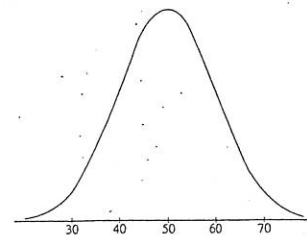
- (A) \$10,590
- (B) \$220,700
- (C) \$264,750
- (D) \$833,100
- (E) \$1,002,500

GO ON TO THE NEXT PAGE

	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

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) μ_B with probability 0.95.
- (B) μ_B with probability 0.95.
- (C) μ_B with probability 0.95.
- (D) μ_B with probability 0.95.
- (E) For any μ_A and μ_B with $(\mu_A - \mu_B) \geq 14$, the sample result is quite likely.

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.

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

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

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.

School Year 1989-1990	School Year 1992-1993
	0 4
9 9 9 9 8 8 7	0 5 6 6 6 7 7 7 8 8 8 9 9
4 4 4 4 3 3 2 2 2 2 1 1 1 0	1 0 0 0 0 1 1 1 1 2 2 2 3 4 4 4 4
9 9 9 7 7 6 6 6 6 5	1 5 5 5 6 6 6 6 7 7 7 8
4 2 2 2 1 0 0	2 1 3
8 8 8 7 6	2
	2 3 0 1 1 2
	3 5
7 6 6	4

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.

3. 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

GO ON TO THE NEXT PAGE

17. 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.

18. 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

GO ON TO THE NEXT PAGE

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)

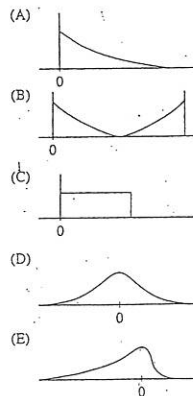
Number of Heads	Frequency
0	0
1	0
2	0
3	2
4	24
5	51
6	22
7	1
8	0
9	0
10	0

GO ON TO THE NEXT PAGE

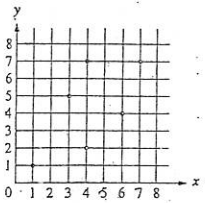
29. In a test of the null hypothesis $H_0: \mu = 10$ against the alternative hypothesis $H_a: \mu > 10$, a sample from a normal population produces a mean of 13.4. The z -score for the sample is 2.12 and the p -value is 0.017. Based on these statistics, which of the following conclusions could be drawn?

- (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?



GO ON TO THE NEXT PAGE



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) $\mu = 16.00$, $\sigma = 0.30$

GO ON TO THE NEXT PAGE

31

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.

33

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

GO ON TO THE NEXT PAGE

32

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%			