

Practice Examination 2

SECTION I

Questions 1–40

Spend 90 minutes on this part of the exam.

Directions: The questions or incomplete statements that follow are each followed by five suggested answers or completions. Choose the response that best answers the question or completes the statement.

- Suppose that the regression line for a set of data, $y = mx + 3$, passes through the point $(2, 7)$. If \bar{x} and \bar{y} are the sample means of the x - and y -values, respectively, then $\bar{y} =$
 - \bar{x} .
 - $\bar{x} - 2$.
 - $\bar{x} + 3$.
 - $2\bar{x} + 3$.
 - $3.5\bar{x} + 3$.
- A study is made to determine whether more hours of academic studying leads to higher point scoring by basketball players. In surveying 50 basketball players, it is noted that the 25 who claim to study the most hours have a higher point average than the 25 who study less. Based on this study, the coach begins requiring the players to spend more time studying. Which of the following is a correct statement?
 - While this study may have its faults, it still does prove causation.
 - There could well be a confounding variable responsible for the seeming relationship.
 - While this is a controlled experiment, the conclusion of the coach is not justified.
 - To get the athletes to study more, it would be more meaningful to have them put in more practice time on the court to boost their point averages, as higher point averages seem to be associated with more study time.
 - No proper conclusion is possible without somehow introducing *blinding*.
- The longevity of people living in a certain locality has a standard deviation of 14 years. What is the mean longevity if 30% of the people live longer than 75 years? Assume a normal distribution for life spans.
 - 61.00
 - 67.65
 - 74.48
 - 82.35
 - The mean cannot be computed from the information given.

GO ON TO THE NEXT PAGE >>

4. Which of the following is a correct statement about correlation?
- If the slope of the regression line is exactly 1, then the correlation is exactly 1.
 - If the correlation is 0, then the slope of the regression line is undefined.
 - Switching which variable is called x and which is called y changes the sign of the correlation.
 - The correlation r is equal to the slope of the regression line when z -scores for the y -variable are plotted against z -scores for the x -variable.
 - Changes in the measurement units of the variables may change the correlation.
5. Which of the following are affected by outliers?
- Mean
 - Median
 - Standard deviation
 - Range
 - Interquartile range
- I, III, and V
 - II and IV
 - I and V
 - III and IV
 - I, III, and IV
6. An engineer wishes to determine the quantity of heat being generated by a particular electronic component. If she knows that the standard deviation is 2.4, how many of these components should she consider to be 99% sure of knowing the mean quantity to within ± 0.6 ?
- 27
 - 87
 - 107
 - 212
 - 425
7. A company that produces facial tissues continually monitors tissue strength. If the mean strength from sample data drops below a specified level, the production process is halted and the machinery inspected. Which of the following would result from a Type I error?
- Halting the production process when sufficient customer complaints are received.
 - Halting the production process when the tissue strength is below specifications.
 - Halting the production process when the tissue strength is within specifications.
 - Allowing the production process to continue when the tissue strength is below specifications.
 - Allowing the production process to continue when the tissue strength is within specifications.
8. Suppose that for a certain Caribbean island in any 3-year period the probability of a major hurricane is .25, the probability of water damage is .44, and the probability of both a hurricane and water damage is .22. What is the probability of water damage given that there is a hurricane?
- $.25 + .44 - .22$
 - $\frac{.22}{.44}$
 - $.25 + .44$
 - $\frac{.22}{.25}$
 - $.25 + .44 + .22$

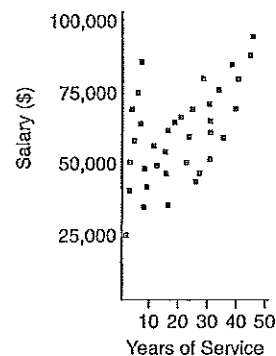
9. Two possible wordings for a questionnaire on a proposed school budget increase are as follows:

- I. This school district has one of the highest per student expenditure rates in the state. This has resulted in low failure rates, high standardized test scores, and most students going on to good colleges and universities. Do you support the proposed school budget increase?
- II. This school district has one of the highest per student expenditure rates in the state. This has resulted in high property taxes, with many people on fixed incomes having to give up their homes because they cannot pay the school tax. Do you support the proposed school budget increase?

One of these questions showed that 58% of the population favor the proposed school budget increase, while the other question showed that only 13% of the population support the proposed increase. Which produced which result and why?

- (A) The first showed 58% and the second 13% because of the lack of randomization as evidenced by the wording of the questions.
- (B) The first showed 13% and the second 58% because of a placebo effect due to the wording of the questions.
- (C) The first showed 58% and the second 13% because of the lack of a control group.
- (D) The first showed 13% and the second 58% because of response bias due to the wording of the questions.
- (E) The first showed 58% and the second 13% because of response bias due to the wording of the questions.

10. A union spokesperson is trying to encourage a college faculty to join the union. She would like to argue that faculty salaries are not truly based on years of service as most faculty believe. She gathers data and notes the following scatterplot of salary versus years of service.



Which of the following most correctly interprets the overall scatterplot?

- (A) The faculty member with the fewest years of service makes the lowest salary, and the faculty member with the most service makes the highest salary.
- (B) A faculty member with more service than another has the greater salary than the other.
- (C) There is a strong positive correlation with little deviation.
- (D) There is no clear relationship between salary and years of service.
- (E) While there is a strong positive correlation, there is a distinct deviation from the overall pattern for faculty with fewer than ten years of service.

11. Two random samples of students are chosen, one from those taking an AP Statistics class and one from those not. The following back-to-back stemplots compare the GPAs.

AP Statistics		No AP Statistics
	1	89
97653	2	015688
98775332110	3	133344777888
1100	4	

Which of the following is true about the ranges and standard deviations?

- (A) The first set has both a greater range and a greater standard deviation.
 (B) The first set has a greater range, while the second has a greater standard deviation.
 (C) The first set has a greater standard deviation, while the second has a greater range.
 (D) The second set has both a greater range and a greater standard deviation.
 (E) The two sets have equal ranges and equal standard deviations.
12. In a group of 10 scores, the largest score is increased by 40 points. What will happen to the mean?
- (A) It will remain the same.
 (B) It will increase by 4 points.
 (C) It will increase by 10 points.
 (D) It will increase by 40 points.
 (E) There is not sufficient information to answer this question.

13. Suppose X and Y are random variables with $\mu_x = 32$, $\sigma_x = 5$, $\mu_y = 44$, and $\sigma_y = 12$. Given that X and Y are independent, what are the mean and standard deviation of the random variable $X + Y$?

- (A) $\mu_{x+y} = 76$, $\sigma_{x+y} = 8.5$
 (B) $\mu_{x+y} = 76$, $\sigma_{x+y} = 13$
 (C) $\mu_{x+y} = 76$, $\sigma_{x+y} = 17$
 (D) $\mu_{x+y} = 38$, $\sigma_{x+y} = 17$
 (E) There is insufficient information to answer this question.
14. Suppose you toss a fair die three times and it comes up an even number each time. Which of the following is a true statement?
- (A) By the law of large numbers, the next toss is more likely to be an odd number than another even number.
 (B) Based on the properties of conditional probability the next toss is more likely to be an even number given that three in a row have been even.
 (C) Dice actually do have memories, and thus the number that comes up on the next toss will be influenced by the previous tosses.
 (D) The law of large numbers tells how many tosses will be necessary before the percentages of evens and odds are again in balance.
 (E) The probability that the next toss will again be even is .5.

15. A pharmaceutical company is interested in the association between advertising expenditures and sales for various over-the-counter products. A sales associate collects data on nine products, looking at sales (in \$1000) versus advertising expenditures (in \$1000). The results of the regression analysis are shown below.

Dependent variable: Sales

Source	df	Sum of Squares	Mean Square	F-ratio
Regression	1	9576.1	9576.1	1118.45
Residual	7	59.9	8.6	

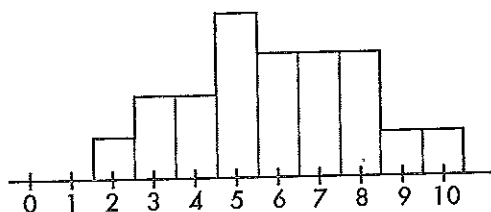
Variable	Coefficient	SE Coef	t-ratio	P
Constant	123.800	1.798	68.84	0.000
Advertising	12.633	0.378	33.44	0.000

R-Sq = 99.4% R-Sq (adj) = 99.3%

s = 2.926 with 9-2 = 7 degrees of freedom

- Which of the following gives a 90% confidence interval for the slope of the regression line?
- (A) $12.633 \pm 1.415(0.378)$
 (B) $12.633 \pm 1.895(0.378)$
 (C) $123.800 \pm 1.414(1.798)$
 (D) $123.800 \pm 1.895(1.798)$
 (E) $123.800 \pm 1.645(1.798/\sqrt{9})$
16. Suppose you wish to compare the AP Statistics exam results for the male and female students taking AP Statistics at your high school. Which is the most appropriate technique for gathering the needed data?
- (A) Census
 (B) Sample survey
 (C) Experiment
 (D) Observational study
 (E) None of these is appropriate.
17. Jonathan obtained a score of 80 on a statistics exam, placing him at the 90th percentile. Suppose five points are added to everyone's score. Jonathan's new score will be at the
- (A) 80th percentile.
 (B) 85th percentile.
 (C) 90th percentile.
 (D) 95th percentile.
 (E) There is not sufficient information to answer this question.
18. To study the effect of music on piecework output at a clothing manufacturer, two experimental treatments are planned: day-long classical music for one group versus day-long light rock music for another. Which one of the following groups would serve best as a control for this study?
- (A) A third group for which no music is played
 (B) A third group that randomly hears either classical or light rock music each day
 (C) A third group that hears day-long R & B music
 (D) A third group that hears classical music every morning and light rock every afternoon
 (E) A third group in which each worker has earphones and chooses his or her own favorite music
19. Suppose $H_0: p = .6$, and the power of the test for $H_a: p = .7$ is .8. Which of the following is a valid conclusion?
- (A) The probability of committing a Type I error is .1.
 (B) If H_a is true, the probability of failing to reject H_0 is .2.
 (C) The probability of committing a Type II error is .3.
 (D) All of the above are valid conclusions.
 (E) None of the above are valid conclusions.

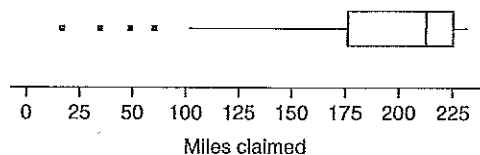
20. Following is a histogram of the numbers of ties owned by bank executives.



Which of the following is a correct statement?

- (A) The median number of ties is five.
 (B) More than four executives own over eight ties each.
 (C) An executive is equally likely to own fewer than five ties or more than seven ties.
 (D) One tie is a reasonable estimate for the standard deviation.
 (E) Removing all the executives with three, nine, and ten ties may change the median.
21. Which of the following is a binomial random variable?
- (A) The number of tosses before a "5" appears when tossing a fair die.
 (B) The number of points a hockey team receives in 10 games, where two points are awarded for wins, one point for ties, and no points for losses.
 (C) The number of hearts out of five cards randomly drawn from a deck of 52 cards, without replacement.
 (D) The number of motorists not wearing seat belts in a random sample of five drivers.
 (E) None of the above.
22. Company I manufactures bomb fuses that burn an average of 50 minutes with a standard deviation of 10 minutes, while company II advertises fuses that burn an average of 55 minutes with a standard deviation of 5 minutes. Which company's fuse is more likely to last at least 1 hour? Assume normal distributions of fuse times.
- (A) Company I's, because of its greater standard deviation
 (B) Company II's, because of its greater mean
 (C) For both companies, the probability that a fuse will last at least 1 hour is 15.9%
 (D) For both companies, the probability that a fuse will last at least 1 hour is 84.1%
 (E) The problem cannot be solved from the information given.
23. Which of the following is *not* important in the design of experiments?
- (A) Control of confounding variables
 (B) Randomization in assigning subjects to different treatments
 (C) Use of a lurking variable to control the placebo effect
 (D) Replication of the experiment using sufficient numbers of subjects
 (E) All of the above are important in the design of experiments.

24. The travel miles claimed in weekly expense reports of the sales personnel at a corporation are summarized in the following boxplot.



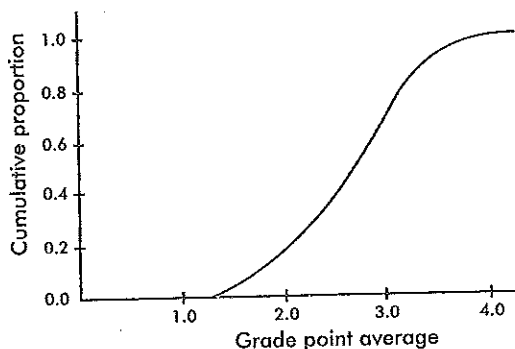
Which of the following is the most reasonable conclusion?

- (A) The mean and median numbers of travel miles are roughly equal.
 (B) The mean number of travel miles is greater than the median number.
 (C) Most of the claimed numbers of travel miles are in the $[0, 200]$ interval.
 (D) Most of the claimed numbers of travel miles are in the $[200, 240]$ interval.
 (E) The left and right whiskers contain the same number of values from the set of personnel travel mile claims.
25. Which of the following statements about residuals is true?
 (A) Influential scores have large residuals.
 (B) If the linear model is good, the number of positive residuals will be the same as the number of negative residuals.
 (C) The mean of the residuals is always zero.
 (D) If the correlation is 0, there will be a distinct pattern in the residual plot.
 (E) If the correlation is 1, there will not be a distinct pattern in the residual plot.
26. Four pairs of data are used in determining a regression line $y = 3x + 4$. If the four values of the independent variable are 32, 24, 29, and 27, respectively, what is the mean of the four values of the dependent variable?
 (A) 68
 (B) 84
 (C) 88
 (D) 100
 (E) The mean cannot be determined from the given information.
27. According to one poll, 12% of the public favor legalizing all drugs. In a simple random sample of six people, what is the probability that at least one person favors legalization?
 (A) $6(.12)(.88)^5$
 (B) $(.88)^6$
 (C) $1 - (.88)^6$
 (D) $1 - 6(.12)(.88)^5$
 (E) $6(.12)(.88)^5 + (.88)^6$
28. Sampling error occurs
 (A) when interviewers make mistakes resulting in bias.
 (B) because a sample statistic is used to estimate a population parameter.
 (C) when interviewers use judgment instead of random choice in picking the sample.
 (D) when samples are too small.
 (E) in all of the above cases.

29. A telephone executive instructs an associate to contact 104 customers using their service to obtain their opinions in regard to an idea for a new pricing package. The associate notes the number of customers whose names begin with *A* and uses a random number table to pick four of these names. She then proceeds to use the same procedure for each letter of the alphabet and combines the $4 \times 26 = 104$ results into a group to be contacted. Which of the following is a correct conclusion?

- (A) Her procedure makes use of chance.
 (B) Her procedure results in a simple random sample.
 (C) Each customer has an equal probability of being included in the survey.
 (D) Her procedure introduces bias through *sampling error*.
 (E) With this small a sample size, it is better to let the surveyor pick representative customers to be surveyed based on as many features such as gender, political preference, income level, race, age, and so on, as are in the company's data banks.

30. The graph below shows cumulative proportions plotted against GPAs for high school seniors.



What is the approximate interquartile range?

- (A) 0.85
 (B) 2.25
 (C) 2.7
 (D) 2.75
 (E) 3.1

31. PCB contamination of a river by a manufacturer is being measured by amounts of the pollutant found in fish. A company scientist claims that the fish contain only 5 parts per million, but an investigator believes the figure is higher. The investigator catches six fish that show the following amounts of PCB (in parts per million): 6.8, 5.6, 5.2, 4.7, 6.3, and 5.4. In performing a hypothesis test with $H_0: \mu = 5$ and $H_a: \mu > 5$, what is the test statistic?

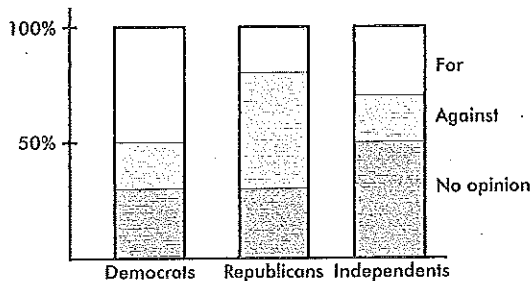
- (A) $t = \frac{5.67 - 5}{0.763}$
 (B) $t = \frac{5.67 - 5}{\sqrt{0.763/5}}$
 (C) $t = \frac{5.67 - 5}{\sqrt{0.763/6}}$
 (D) $t = \frac{5.67 - 5}{0.763/\sqrt{5}}$
 (E) $t = \frac{5.67 - 5}{0.763/\sqrt{6}}$

32. The distribution of weights of 16-ounce bags of a particular brand of potato chips is approximately normal with a standard deviation of 0.28 ounce. How does the weight of a bag at the 40th percentile compare with the mean weight?

- (A) 0.40 ounce above the mean
 (B) 0.25 ounce above the mean
 (C) 0.07 ounce above the mean
 (D) 0.07 ounce below the mean
 (E) 0.25 ounce below the mean

33. In general, how does tripling the sample size change the confidence interval size?
- (A) It triples the interval size.
 - (B) It divides the interval size by 3.
 - (C) It multiplies the interval size by 1.732.
 - (D) It divides the interval size by 1.732.
 - (E) This question cannot be answered without knowing the sample size.
34. Which of the following statements is *false*?
- (A) Like the normal distribution, the t -distributions are symmetric.
 - (B) The t -distributions are lower at the mean and higher at the tails, and so are more spread out than the normal distribution.
 - (C) The greater the df , the closer the t -distributions are to the normal distribution.
 - (D) The smaller the df , the better the 68–95–99.7 Rule works for t -models.
 - (E) The area under all t -distribution curves is 1.

35. A study on school budget approval among people with different party affiliations resulted in the following segmented bar chart:



- Which of the following is greatest?
- (A) Number of Democrats who are for the proposed budget
 - (B) Number of Republicans who are against the budget
 - (C) Number of Independents who have no opinion on the budget
 - (D) The above are all equal.
 - (E) The answer is impossible to determine without additional information.

36. The sampling distribution of the sample mean is close to the normal distribution
- (A) only if both the original population has a normal distribution and n is large.
 - (B) if the standard deviation of the original population is known.
 - (C) if n is large, no matter what the distribution of the original population.
 - (D) no matter what the value of n or what the distribution of the original population.
 - (E) only if the original population is not badly skewed and does not have outliers.
37. What is the probability of a Type II error when a hypothesis test is being conducted at the 10% significance level ($\alpha = .10$)?
- (A) .05
 - (B) .10
 - (C) .90
 - (D) .95
 - (E) There is insufficient information to answer this question.



Above is the dotplot for a set of numbers. One element is labeled X . Which of the following is a correct statement?

- (A) X has the largest z -score, in absolute value, of any element in the set.
- (B) A modified boxplot will plot an outlier like X as an isolated point.
- (C) A stemplot will show X isolated from two clusters.
- (D) Because of X , the mean and median are different.
- (E) The IQR is exactly half the range.

39. A 2008 survey of 500 households concluded that 82% of the population uses grocery coupons. Which of the following best describes what is meant by the poll having a margin of error of 3%?
- (A) Three percent of those surveyed refused to participate in the poll.
- (B) It would not be unexpected for 3% of the population to begin using coupons or stop using coupons.
- (C) Between 395 and 425 of the 500 households surveyed responded that they used grocery coupons.
- (D) If a similar survey of 500 households were taken weekly, a 3% change in each week's results would not be unexpected.
- (E) It is likely that between 79% and 85% of the population use grocery coupons.

40.

	College Plans	
	Public	Private
Taking AP Statistics:	18	27
Not taking AP Statistics:	26	40

The above two-way table summarizes the results of a survey of high school seniors conducted to determine if there is a relationship between whether or not a student is taking AP Statistics and whether he or she plans to attend a public or a private college after graduation. Which of the following is the most reasonable conclusion about the relationship between taking AP Statistics and the type of college a student plans to attend?

- (A) There appears to be no association since the proportion of AP Statistics students planning to attend public schools is almost identical to the proportion of students not taking AP Statistics who plan to attend public schools.
- (B) There appears to be an association since the proportion of AP Statistics students planning to attend public schools is almost identical to the proportion of students not taking AP Statistics who plan to attend public schools.
- (C) There appears to be an association since more students plan to attend private than public schools.
- (D) There appears to be an association since fewer students are taking AP Statistics than are not taking AP Statistics.
- (E) These data do not address the question of association.



If there is still time remaining, you may review your answers.