

AP Statistics
Practice Test 2
Multiple-Choice

Name: _____

1. A candy company produces chocolate bars that have a mean weight of 8 ounces with a standard deviation of 0.27 ounces. Assume that the production process is independent and that the weights are approximately normally distributed. If two chocolate bars are randomly selected, what is the approximate probability that both will weigh less than 7.5 ounces?

- (A) 0.001
(B) 0.004
(C) 0.032
(D) 0.036
(E) 0.064

2. A large product-testing firm conducts interviews with people who come to the company offices to participate in focus groups for new products. After the participants have had a chance to view the product or participate in a taste test, an interviewer will spend time asking the participants questions. Interviews and focus groups are scheduled between noon and 8 p.m. to accommodate participants who work. The interviewers have a varying number of weeks of experience at their job, and the company wants to know if there is a relationship between the number of weeks of experience and the number of interviews that can be conducted during the course of a typical interview day. If there is such a relationship, it will help the company schedule the number of participants based on which interviewers are available. Below is some computer output from a linear regression analysis of data from 10 randomly selected interviewers with anywhere from 15 to 60 weeks of work experience. An examination of the residual plot indicates no apparent pattern.

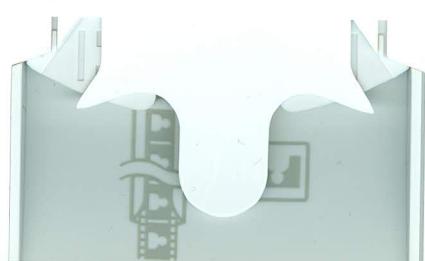
Predictor	Coef	SE Coef	T	P
Constant	2.1538	0.7694	2.80	0.023
Weeks of Experience	0.1637	0.0205	8.00	0.000
S = 0.873067 R-Sq = 88.9% R-Sq(adj) = 87.5%				

Which one of the following represents a 96% confidence interval for β , the true slope for the regression line relating weeks of interview experience and the number of interviews conducted in one day?

- (A) $0.1637 \pm 2.359(.8731)$
(B) $0.1637 \pm 2.398(.0205)$
(C) $0.1637 \pm 2.449(.0205)$
(D) $2.1538 \pm 2.398(.8731)$
(E) $2.1538 \pm 2.449(.7694)$

3. A researcher wants to determine the mean income of adults in a particular state. She decides to take a random sample of 1525 citizens with a driver's license from the registry maintained by the Department of Motor Vehicles in that state, and record their incomes. Which one of the following statements is NOT correct?
- (A) People who had recently moved from another state may not yet be included in the DMV registry.
(B) The sample would not be representative of adults, since teenagers would be included in the sample.
(C) The sample would be biased since some people without a license may be too poor to own a car and others may choose to use public transportation.
(D) The distribution of incomes in the sample would probably be skewed to the right.
(E) Taking a bigger sample from the registry would solve any problems with possible bias.
4. As the costs of college increase, financial counselors are becoming more concerned about the debt load of students upon graduation. A survey is being planned to determine how much debt from student loans, on average, undergraduate students have accumulated by the time they graduate. How many students should researchers plan to survey in order to be within \$500 of the true mean with 98% confidence? A preliminary sample indicates that the standard deviation of student loan debts is approximately \$2650.
- (A) 13
(B) 60
(C) 108
(D) 153
(E) 244
5. A certain candy has different color wrappers for various holidays. During holiday #1 the candy wrappers are 50% red and 50% green. During holiday #2 the wrapper colors are 30% red, 30% silver, and 40% pink. Fifty pieces of candy are randomly selected from the holiday #1 distribution and 50 pieces of candy are randomly selected from the holiday #2 distribution. Respectively, what is the expected number and standard deviation of the total number of red wrappers from both holiday distributions?
- (A) 40, 23
(B) 40, 4.796
(C) 40, 6.776
(D) 80, 6.776
(E) Cannot be determined from the given information.

6. In a recent study, a random sample of 250 trees that had been planted five years ago along the Oregon coast showed a significant negative relationship between the distance from the shoreline and the height of the tree. The distance from the shoreline is
- (A) A block design with the distance from the shoreline as the block.
(B) A response variable.
(C) An explanatory variable.
(D) A confounding variable.
(E) Random variation.
7. A college agricultural department wants to test the yields of four different varieties of tomato plant that it has developed. It also wants to determine how the type of soil in which the tomatoes are grown — sandy or clay — and which of two fertilizers used will affect yield. How many treatment groups would be required for this experiment?
- (A) 16
(B) 12
(C) 8
(D) 6
(E) 3
8. A large company is testing a new marketing strategy for increasing sales at its stores. Currently store sales average \$250,000 per month. Thus the company wishes to test the following hypotheses: $H_0 : \mu = \$250,000$ versus $H_a : \mu > \$250,000$, where μ = true mean monthly sales per store. Which one of the following sample sizes and significance levels would lead to the highest power for this test?
- (A) $n = 10$ and $\alpha = 0.01$
(B) $n = 20$ and $\alpha = 0.05$
(C) $n = 40$ and $\alpha = 0.05$
(D) $n = 20$ and $\alpha = 0.10$
(E) $n = 40$ and $\alpha = 0.10$



9. A random sample of 40 trees was used to estimate the mean height of all trees at a tree farm. Summary statistics for this sample are given below.

N	MEAN	MEDIAN	STDEV	MIN	Q1	Q3	MAX
40	18.50	18.0	5.169	10.0	14.50	22.0	29.0

Which one of the following stem-and-leaf plots most likely represents the actual sample data set?

NOTE : $1|4 = 14$ (stem = tens and leaf = ones)

(A)

1	11
1	3
1	4
1	666777
1	88888888999
2	0001
2	333333
2	45
2	66677
2	89

(B)

1	00011
1	23
1	4446
1	666777
1	88888999
2	01
2	3333
2	45
2	6677
2	899

(C)

1	0001
1	2333
1	4445
1	6677
1	888889999
2	0111
2	333
2	455
2	667
2	899

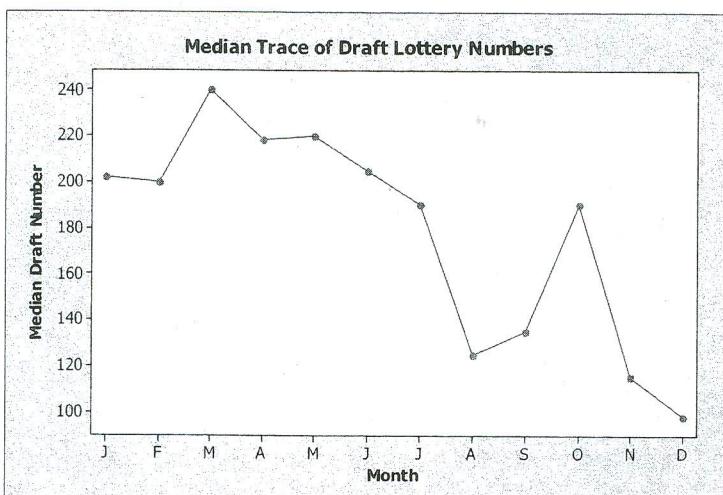
(D)

1	0011
1	233
1	4445
1	666777
1	888889999
2	011
2	333
2	4
2	667
2	889

(E)

1	00011
1	23
1	44
1	666777
1	88888999
2	01
2	3333
2	4
2	6677
2	8899
3	01

10. In 1970, the U.S. Armed Services instituted a lottery as a random selection process when 19-year-old males were drafted into service. Each of the 366 possible birth dates was placed in a separate capsule and each capsule was placed in a large plastic drum. Those with the lowest lottery numbers were drafted before others. For example, the May 9th number was 176. This meant that there would be 175 groups of males born on other days drafted before those who had been born on May 9th. The scatter plot below shows the median draft eligible number for each month of the year. The line that connects the points is sometimes called a median trace line. Which one of the following statements is TRUE?



- (A) There seems to be no relationship between birth month and median draft lottery number.
- (B) There is a very weak negative correlation between birth month and median draft lottery number.
- (C) Those people born near the end of the year tended to have lower median draft lottery numbers than those born at the beginning of the year.
- (D) Since the spread for each month was about the same (30 days), the median draft lottery number for each month is about the same.
- (E) Since each birth date is independent, the scatterplot indicates that you were just as likely to be drafted early in the process as late in the process.

11. It is known that 15% of the seniors in a large high school enter military service upon graduation. If a group of 20 seniors are randomly selected, what is the probability of observing at most one senior who will be entering military service?

- (A) $20(0.15)^1(0.85)^{19}$
- (B) $1 - 20(0.15)^1(0.85)^{19}$
- (C) $(0.85)^{20} + 20(0.15)^1(0.85)^{19}$
- (D) $(0.85)^{20}$
- (E) $1 - (0.85)^{20}$

12. A study done by the Duke University Medical School examined the effectiveness of St. John's Wort as a natural alternative to reducing depression. In a double-blind study, 100 people with mild to moderate depression were given daily doses of St. John's Wort. Another 100 people with similar symptoms were given a sugar pill. Assignments to the two treatment groups were made randomly. Each participant was asked to count the number of days they exhibited feelings of depression. A two-sample t test on the difference in the mean number of days of depression was performed. The P -value was 0.42. Which of the following is a correct interpretation of the P -value?
- (A) About 42% of the time, samples drawn from populations with no differences in mean depression count would show a difference at least as extreme as that found in the study.
(B) Forty-two percent of the participants did not experience any relief.
(C) There was a 42% drop in the mean number of days of depression in the control group.
(D) There was a 42% drop in the mean number of days of depression in the group that received the St. John's Wort.
(E) There was a 42% drop in the difference in the number of days of depression between the two groups.
13. A manufacturer of equipment for boats is interested in producing a new type of marine radio with an emergency homing beacon. It is expensive to set up production, so the manufacturer wants to determine the level of interest in purchasing such a device before money is invested in this venture. The company is located on the West Coast, which is its primary market. The company decides to randomly select 12 of the marinas on the West Coast and then demonstrate the prototype to all of the boat owners at these marinas and gauge the potential market for the device. What sampling method did the company use to gather its data?
- (A) Simple random sample.
(B) Systematic.
(C) Stratified.
(D) Cluster.
(E) Convenience.
14. A cereal maker claims that the weight of its new packages is 48 ounces. An investigator from the Division of Weights and Measures believes that the packages hold less than the stated amount. Let μ be the true mean weight in ounces of all packages produced by the company. Which one of the following represents the null and alternate hypotheses the investigator should test?
- (A) $H_0 : \mu = 48$ and $H_a : \mu < 48$
(B) $H_0 : \mu > 48$ and $H_a : \mu < 48$
(C) $H_0 : \mu = 48$ and $H_a : \mu \neq 48$
(D) $H_0 : \mu < 48$ and $H_a : \mu > 48$.
(E) $H_0 : \mu > 48$ and $H_a : \mu \leq 48$



15. A reporter believes that police officers were required to write a specific quota of traffic tickets during a month. In order to meet the alleged quota, he thinks officers would need to write more tickets during the last week of the month. To investigate the claim, the reporter collected the number of tickets written by the local police force in a month and organized them by week as shown in the table below.

Week	First Week	Second week	Third week	Fourth week	Total
Tickets written	133	112	154	165	564

Which one of the following inference procedures would be appropriate for testing the reporter's belief?

- (A) A chi-square test for independence.
- (B) A chi-square goodness-of-fit test.
- (C) A linear regression t test on slope.
- (D) A two-sample z test for proportions.
- (E) A t test for means.

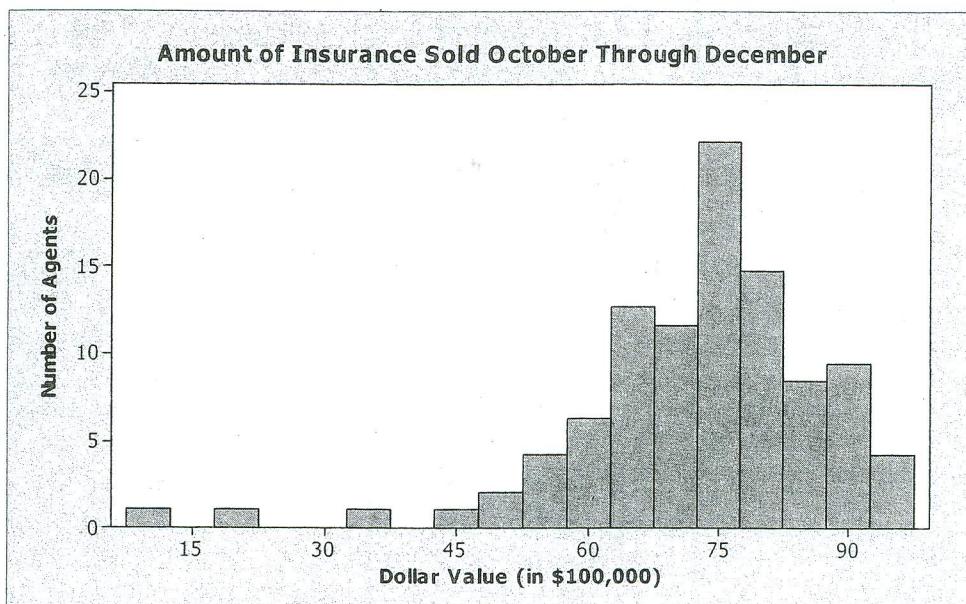
16. A study tested a recently advertised claim that a newly developed "gas pill" boosts mileage when added to a tank of gas. Ten randomly selected new cars were used in the study and would be run over an identical course twice, once with the additive and once without the additive. The order of the two treatments (pill and no pill) was randomized. Each car's tank was filled with a standard gasoline and one of the treatments applied. The car was then driven until the tank was empty. The car's tank was then refilled with the same standard gasoline and the process repeated for the other treatment. Professional drivers drove over the same course at identical speeds and did not know which treatment was used. The miles per gallon for each car were calculated under both treatments. The results from a computer analysis are shown below. Let C1 represent cars without the gas pill and C2 represent the same cars with the gas pill.

ONE SAMPLE T FOR C1 - C2				
	N	MEAN	STDEV	SE MEAN
C1 - C2	10	0.8	1.3	0.411
95 PCT CI FOR MU (C1 - C2): (-0.13, 1.73)				
TTEST MU (C1 - C2) (VS Not Equal): T = 1.946 P = 0.083 DF = 9				

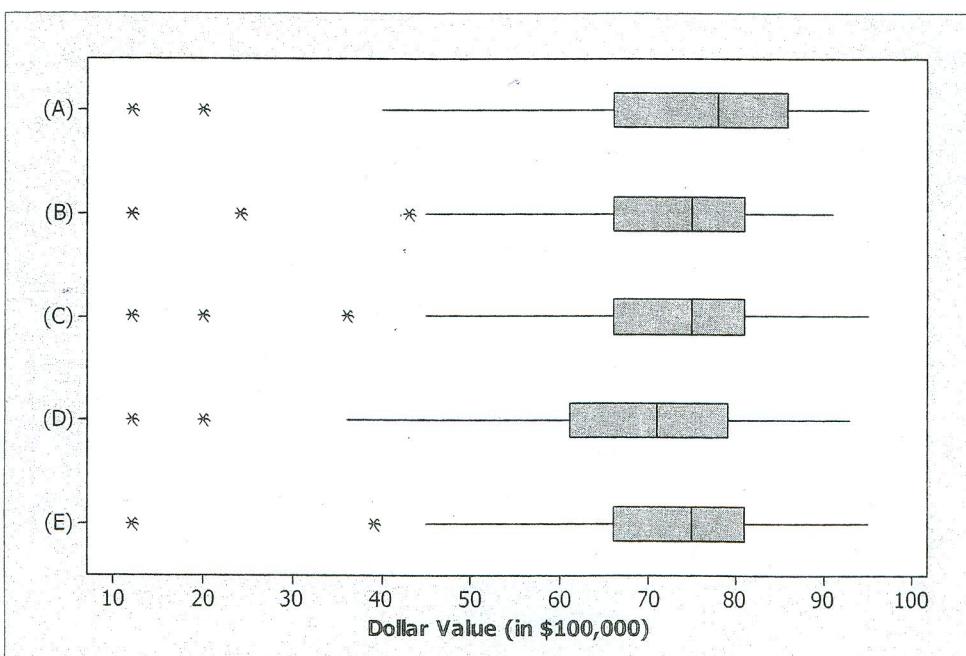
Based on the information, which one of the following statements is TRUE?

- (A) Ninety-five percent of the time, the confidence interval will contain 0.
- (B) A two-sample t test should have been used instead of a paired t test.
- (C) The true mean mileage is higher for cars without the additive than those with the additive since the mean difference is positive.
- (D) The sample size is too small to draw any conclusions.
- (E) Since the interval contains 0, we cannot conclude that using the gas pill made a difference in mileage.

17. A large insurance company has 95 agents within a mid-size state. The histogram below shows the amount of insurance sold (in \$100,000) for the period October through December in a recent year. The highest bar, for example, indicates that 23 agents sold between \$7.25 million and \$7.75 million of insurance during the three-month period.



Which one of the following represents a boxplot of the same data?



18. A national consumer organization has conducted a poll among purchasers of desktop and laptop computers concerning the consumer's level of satisfaction with the selection of computers that were available for purchase, prices, knowledge of the machines by the salespeople, helpfulness of sales staff, and the availability of support services and technical advice offered when there was a problem. Sources of computers were divided into two groups: sales originating at retail/electronic stores or sales from manufacturer-direct stores and Web catalog retailers. For each type of sales outlet, a random sample of consumers who actually purchased a computer was asked to complete a survey, which consisted of a series of questions and ratings to be checked off. Scores for the survey ranged from 0 to 100. Each score in the stemplot below represents the average for at least 200 responses for that store or Web site. A score of 80 was considered to be customers who were, on average, very satisfied with the experience. A score of 60 indicated that customers were fairly satisfied. A score of 40 would indicate, on average, dissatisfied customers.

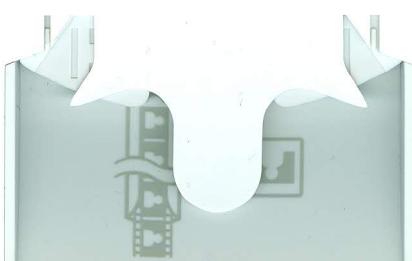
Retail Stores		Manufacturer Direct and Web Sales
	9	2, 4
0, 1, 2	8	1, 2, 3, 5, 6, 7
3, 3, 4, 5, 7, 7, 9	7	4
2, 3, 5, 6	6	8
3, 7	5	
2	4	

Which one of the following statements is TRUE?

- (A) The ranges for the two ratings distributions are the same.
(B) The ratings distribution for retail/electronic stores is skewed to the right.
(C) The median of the ratings distribution for retail/electronic stores is higher than the median of the ratings distribution for direct and Web sales.
(D) The IQR for the ratings distribution of retail/electronic stores is smaller than the IQR for the ratings distribution of manufacturer direct and Web site stores.
(E) The ratings for retail/electronic stores are, on average, lower than the ratings for manufacturer direct and Web sales.
19. A hospital administrator wants to determine a 94% confidence interval for the mean number of days that a patient stays at the hospital. The data given to her by the hospital's business manager lists the length of stay for each patient for the previous calendar year. Which one of the following is the best reason the administrator cannot determine the confidence interval?
- (A) A confidence interval cannot be constructed when population data are known.
(B) Since most stays are just a few days long, the data are heavily skewed.
(C) A 90% or 95% confidence interval can be computed but not a 94% confidence interval since the tables do not show the required critical values.
(D) Since many patients were treated in the emergency room and then released, their stay would be zero days. This cannot be factored into the confidence interval.
(E) Some patients may have been treated at the hospital more than once, which invalidates the independence requirement for a confidence interval.



20. The distribution of ages of females in the United States is strongly skewed to the left with a mean of 80.2 years. A random sample of $n = 20$ females is taken from this population and the mean age of the sample is calculated. This is repeated 500 times. Which one of the following best describes the shape of the sampling distribution?
- (A) Cannot be determined because the standard deviation is unknown.
(B) Skewed to the left with a mean of $\frac{80.2}{\sqrt{20}}$.
(C) Skewed to the left with a mean of 80.2.
(D) Approximately normally distributed with a mean of 80.2.
(E) Approximately normally distributed with a mean of $\frac{80.2}{\sqrt{20}}$.
21. A simple random sample of 100 batteries is selected from a process that produces batteries with a mean lifetime of 32 hours and a standard deviation of 3 hours. Thus the standard deviation of the sampling distribution, $\sigma_{\bar{x}}$, is equal to 0.3. If the sample size had been 400, how would the value of $\sigma_{\bar{x}}$ change?
- (A) It is one-fourth as large as when $n = 100$.
(B) It is one-half as large as when $n = 100$.
(C) It is twice as large as when $n = 100$.
(D) It is four times as large as when $n = 100$.
(E) The value of $\sigma_{\bar{x}}$ does not change.
22. An advertisement for a new ten-week reading program claims that the method significantly improves a person's reading comprehension without reducing his or her reading speed. To test this claim, a random sample of 15 candidates is enrolled in the program. A reading comprehension test is administered before the program starts and an equivalent test is administered at the end. The order of the two tests is randomized for all subjects. The individual scores range from 0 (lowest comprehension) to 100 (highest comprehension). Which one of the following procedures should be used to test the claim made in the advertisement?
- (A) A chi-square test for homogeneity.
(B) A two-sample z test.
(C) A one-sample z test.
(D) A two-sample t test.
(E) A one-sample t test.



23. Which of the following are resistant to outliers?

- (A) The mean and standard deviation.
- (B) The median and standard deviation.
- (C) The median and the range.
- (D) The median and the *IQR*.
- (E) The range and *IQR*.

24. Consumers frequently complain that there is a large variation in prices charged by different pharmacies and drug stores for the same medication. A survey of a large sample of the pharmacies in a major city revealed that the prices charged for one bottle containing 50 tablets of a popular pain reliever were approximately normally distributed.

The charge of \$10.48 for this bottle was at the 90th percentile and 20% of the bottles cost less than \$8.61. What are, respectively, the approximate mean and standard deviation of the price distribution for bottles of this pain reliever?

- (A) \$9.35 and \$0.88
- (B) \$9.41 and \$2.67
- (C) \$9.55 and \$0.88
- (D) \$9.55 and \$1.70
- (E) \$9.73 and \$1.70

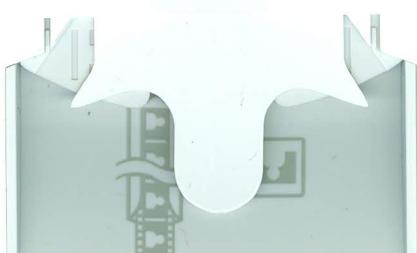
25. An electronics company is testing whether a new assembly process will increase the number of circuits that can be produced daily by each worker. Nine people are randomly selected to participate in the study. They are trained in the new assembly process. When training is completed, the number of circuits that they produce in a day is recorded. The number of circuits assembled in one day using the new method was then compared to the number of circuits assembled in one day using the old method. The results from both methods are shown in the following table:

Person	1	2	3	4	5	6	7	8	9
Old method (units per day)	67	68	73	77	69	65	70	69	70
New method (units per day)	69	68	74	77	72	68	71	72	71

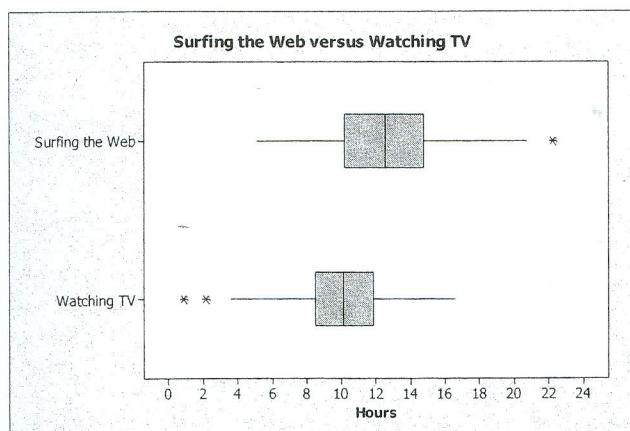
The null hypothesis H_0 : The training had no effect on the mean number of circuits that could be assembled daily was tested versus the alternate hypothesis H_a : The training increased the mean number of circuits that could be assembled daily. Using the proper test statistic, which one of the following is the correct decision rule if $\alpha = 0.05$?

- (A) Reject H_0 if the test statistic is greater than 1.645.
- (B) Reject H_0 if the test statistic is greater than 1.812.
- (C) Reject H_0 if the test statistic is greater than 1.833.
- (D) Reject H_0 if the test statistic is greater than 1.860.
- (E) Reject H_0 if the test statistic is greater than 2.262.

26. Golf courses have a wide range of difficulty. Similarly, players differ in ability. In order to adjust for variations between players, they are often assigned a handicap score. To adjust for variations between courses, a handicapper decides to compare the golfer's score against the data from the course. Suppose that course A plays at a mean score of 76 with a standard deviation of 8 strokes with an approximately normal distribution of scores. The mean score for course B is 80 with a standard deviation of 6 strokes and the scores are also approximately normally distributed. If a golfer regularly shoots an 80 on course A, what should be the comparable score on course B?
- (A) 80
(B) 83
(C) 84
(D) 86
(E) 88
27. About 78% of the U.S. population claims that they visit the dentist at least once a year. A simple random sample of 200 people is selected for a health study and the sample proportion of those who visited the dentist at least once during the year is computed. Which one of the following gives, respectively, the mean and standard deviation of the sampling distribution of the sample proportion who claim they visit the dentist at least once a year?
- (A) 156, 12.49
(B) 156, 5.86
(C) 0.78, 0.012
(D) 0.78, 0.029
(E) You cannot determine the standard deviation from the given information.
28. After a recent chemical spill into a large bay, warnings were issued about eating oysters taken from this body of water. It was declared that oysters would be unsafe to eat if the polychlorinated biphenol (PCB) concentration is at least 5 parts per billion (ppb). The state Department of Natural Resources is testing oysters brought in by commercial fishermen to determine their suitability to be sold to the public. The Department's working hypotheses are $H_0 : \mu = 5$ and $H_a : \mu < 5$, where μ = true mean concentration of PCB in the oysters. Which of the following is a description of a Type II error and its consequence?
- (A) The department declares the oysters safe to eat and, in fact, they are safe. The public gets to eat oysters and the fishermen get paid for their catch.
(B) The department declares the oysters safe to eat when, in fact, they are unsafe. The public is put in jeopardy from eating contaminated oysters.
(C) The department declares the oysters safe to eat when, in fact, they are not. Fishermen get paid for their catch and the public gets to eat oysters.
(D) The department cannot find the evidence to declare the oysters safe to when, in fact, they are unsafe. The public is protected from eating contaminated oysters.
(E) The department cannot find the evidence to declare the oysters safe to eat when, in fact, they are safe to eat. Fishermen lose out on their income and the public doesn't get to eat oysters.



29. In a recent social research poll, 4800 randomly selected adults aged 18 to 26 were asked to record the number of hours they spent watching television and the number of non-work related hours they spent surfing the Web during a typical week. Parallel boxplots of the data are given below.



Based on the plots, which one of the following statements is FALSE?

- (A) The range of the distribution of hours spent surfing the Web is higher than the range for the distribution of hours spent watching television.
- (B) On average, adults aged 18 to 26 spent more time surfing the Web than watching television.
- (C) The median number of hours spent surfing the Web is higher than the median number of hours spent watching television.
- (D) The IQR for hours spent surfing the Web is smaller than the IQR for hours spent watching TV.
- (E) Both distributions of hours are a slightly skewed.

30. A dice game pays a player \$5 for rolling a 3 or a 5 with a single die. The player has to pay \$2 for any other roll. If a person plays the game 30 times, what is the approximate probability that the person will win at least \$15?

- (A) 0.0030
- (B) 0.0643
- (C) 0.2767
- (D) 0.3085
- (E) 0.3910

31. A car rental agency has two locations, one by the airport and the other downtown. The agency rents 70% of its cars at the airport and the rest downtown. At the airport, 20% of the rentals are Sport Utility Vehicles (SUVs). At the downtown location, 40% are SUVs. If a customer is selected at random, what is the probability that she did NOT rent an SUV?
- (A) 0.74
(B) 0.70
(C) 0.66
(D) 0.40
(E) 0.26
32. Which one of the following statements about a confidence interval is TRUE?
- (A) Changes in the required level of confidence, $1 - \alpha$, do not affect the width of the confidence interval.
(B) The width of a confidence interval is increased when the sample size is increased.
(C) A confidence interval cannot be used to make a decision for a test of significance.
(D) The wider the confidence interval, the more precise the estimate of the population parameter.
(E) The width of the confidence interval is affected by the variation in the original population.
33. A large sales company recruits many graduating students from universities for its workforce. Thirty percent of those hired for management positions come from private universities and colleges and the rest from public colleges and universities. It is very expensive and time-consuming to train new managers, so the company is examining its retention rate (those still working for the company after six years) of these hires. Over the past six years, 35% of those managers who were hired from private schools had left for other jobs, while 20% of those from public schools had done so. What is the probability that a randomly selected person, who left the company within the past six years, was hired from a private university or college?
- (A) 0.105
(B) 0.4286
(C) 0.2308
(D) 0.35
(E) 0.2450

34. A family-run convenience store/gas station in a resort town experiences a lot of variation in sales, depending whether it is during the summer vacation season or the regular year. Sales during the off-season do not vary that much, but the vacation season can be very unpredictable. The family is interested in predicting the total daily sales from the number of customers who buy gas and various other items in the store each day during the summer vacation season. A random sample of 20 days from the past two summer seasons is taken. Some of the summary statistics about the number of customers per day and daily sales (in \$1000) are given below.

mean number of customers = 721.47

standard deviation of customers = 182.0

mean daily sales (\$1000) = 38.778

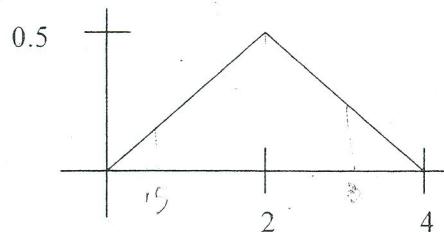
standard deviation of daily sales (\$1000) = 7.442

$$r^2 = 0.743$$

Which one of the following would be closest to the predicted daily sales if there were 800 customers that day?

- (A) \$10,820
- (B) \$14,555
- (C) \$16,770
- (D) \$41,135
- (E) \$41,545

35. A continuous random variable X has the probability distribution as shown in the graph below. What is $P(0.5 < X < 3)$?

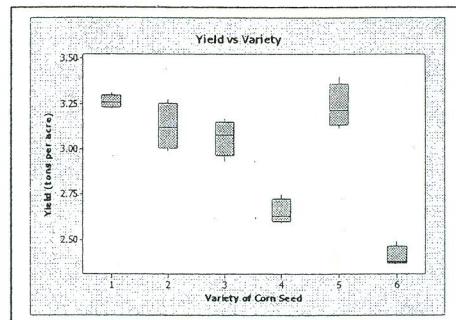
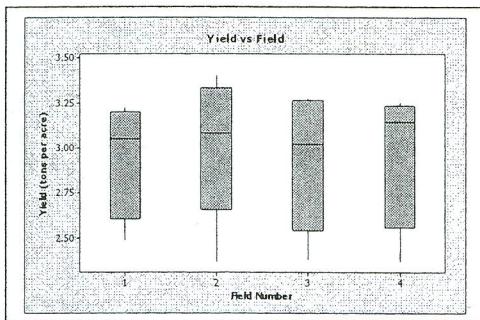


- (A) 3/4
- (B) 13/16
- (C) 7/8
- (D) 27/32
- (E) 15/16

36. An agricultural station is testing the yields for six different varieties of seed corn. The station has four large fields available, which are located in four distinctly different parts of the county. The agricultural researchers consider the climatic and soil conditions in the four parts of the county as being unequal but are reasonably confident that the land in each field is fairly uniform. The researchers divide each field into six sections and then randomly assign one variety of corn seed to each section in that field. This procedure is done for each field. At the end of the growing season, the corn will be harvested, and the yield measured in tons per acre will be compared. Which one of the following statements about the design is correct?

- (A) This is an observational study since the researchers are watching the corn grow.
- (B) This is a block design with fields as blocks and seed types as treatments.
- (C) This is a block design with seed types as blocks and fields as treatments.
- (D) This is a completely randomized design since the six seed types were randomly assigned to the four fields.
- (E) This is a completely randomized design with 24 treatments—6 seed types and 4 fields.

37. The researchers in question #36 are now analyzing the results of their work. As a part of their analysis they have constructed several boxplots of the data. They are interested in examining the affect of the field and the variety of seed on the yield per acre. Was there a significant difference in the yield as measured by tons of corn per acre due to the fields or due to the variety of the seed?

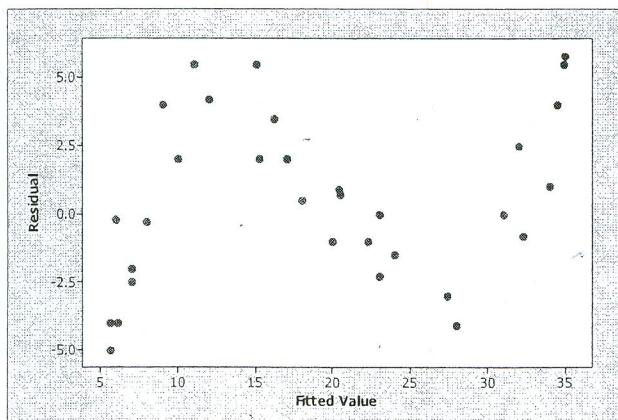


Which of the following is a correct conclusion based on the above graphs?

- (A) There appears to be a significant difference in yield per acre due to the fields.
- (B) There appears to be a significant difference in yield per acre due to the varieties of corn.
- (C) Both field and variety of corn appear to lead to significant differences in yield per acre.
- (D) Neither fields nor variety of corn appear to lead to significant differences in yield per acre.
- (E) These graphs give you no information about the effect of field and corn variety on the yield per acre

38. There is an ever-increasing number of sources for getting the daily news—traditional newspapers, online, radio, nightly television newscasts, comedy shows, cell phones, podcasts, etc. In a recent telephone survey, 3204 randomly selected adults were asked to cite their primary source of daily news. Four in 10 adults said that they read a newspaper, either in print or online, almost every day. A 98% confidence interval to estimate the true proportion of adults who would read either a print newspaper or its online equivalent for their daily news is given by $(0.38, 0.42)$. Which of the following is a correct interpretation of the confidence level?
- (A) Ninety-eight percent of all samples of this size would yield a confidence interval of $(0.38, 0.42)$.
(B) There is a 98% chance that the true proportion of readers who would read either a print newspaper or its online equivalent for their daily news is in the interval $(0.38, 0.42)$.
(C) The procedure used to generate this interval will capture the true proportion of readers who would read either a print newspaper or its online equivalent for their daily news 98% of the time.
(D) Ninety-eight percent of all of the samples of size 3204 lie in the confidence interval $(0.38, 0.42)$.
(E) There is a 98% chance that a randomly selected reader is one of the 40% who would read a print newspaper or its online equivalent for their daily news.

39. The residual plot for a least-squares regression model is given below.



Which of the following statements about the residual plot is TRUE?

- (A) The plot shows that the least-squares model is a good fit.
(B) The plot shows that the residual errors are independent of the x -value.
(C) The plot shows that the residual errors follow a normal distribution.
(D) The plot shows an equal variance in the residuals.
(E) An appropriate model is something other than a linear one.

40. There is intense competition among Internet service providers to get their home customers to switch from dial-up service to high-speed broadband service. A survey of 1758 randomly selected U.S. households conducted in 2006 asked whether or not they had Internet service. If the response was yes, then they were asked about household income level and whether the Internet service was dial-up or high-speed broadband. The results in 2006 were compared to a similar poll of 1494 randomly selected households in 2005. The table below summarizes the data for household income and the number of households with high-speed broadband service for the two years of the poll.

Household Income	2006	2005
Under \$30,000	196	184
\$30,000 – under \$50,000	414	316
\$50,000 – under \$75,000	488	382
At least \$75,000	660	612

A chi-square test was performed, which resulted in a P -value of 0.0760. Which of the following conclusions is correct.

- (A) At the 5% level of significance, we have strong evidence that the rate of broadband use across income levels has not changed between 2005 and 2006.
- (B) At the 5% level of significance, we can conclude that the rate of broadband use across various income levels has changed between 2005 and 2006.
- (C) At the 5% level of significance, there is sufficient evidence to conclude that an association exists between rate of broadband use across income levels and year.
- (D) At the 10% level of significance, we can conclude that the rate of broadband use across income groups is independent of year.
- (E) At the 10% level of significance, we can conclude that the rate of broadband use across income levels has changed between 2005 and 2006.

Free-Response

1. The Better Business Council of a large city has concluded that students in the city's schools are not learning enough about economics to function in the modern world. These findings were based on test results from a random sample of 20 twelfth-grade students who completed a 46-question multiple-choice test on basic economic concepts. The data set below shows the number of questions that each of the 20 students in the sample answered correctly.

12	16	18	17	18	33	41	44	38	35
19	36	19	13	43	8	16	14	10	9

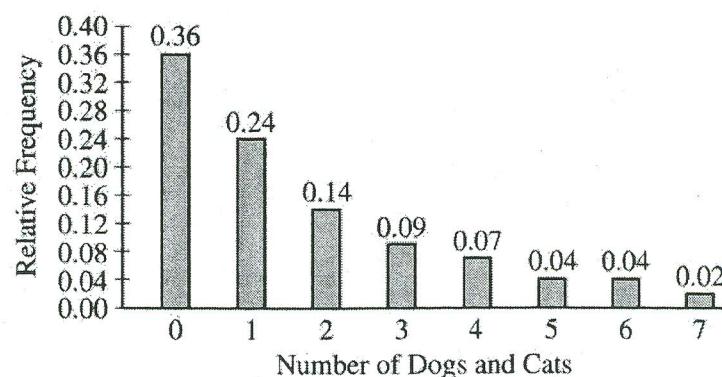
- (a) Display these data in a stemplot.
- (b) Use your stemplot from part (a) to describe the main features of this score distribution.
- (c) Why would it be misleading to report only a measure of center for this score distribution?

2. As dogs age, diminished joint and hip health may lead to joint pain and thus reduce a dog's activity level. Such a reduction in activity can lead to other health concerns such as weight gain and lethargy due to lack of exercise. A study is to be conducted to see which of two dietary supplements, glucosamine or chondroitin, is more effective in promoting joint and hip health and reducing the onset of canine osteoarthritis. Researchers will randomly select a total of 300 dogs from ten different large veterinary practices around the country. All of the dogs are more than 6 years old, and their owners have given consent to participate in the study. Changes in joint and hip health will be evaluated after 6 months of treatment.
- (a) What would be an advantage to adding a control group in the design of this study?
- (b) Assuming a control group is added to the other two groups in the study, explain how you would assign the 300 dogs to these three groups for a completely randomized design.
- (c) Rather than using a completely randomized design, one group of researchers proposes blocking on clinics, and another group of researchers proposes blocking on breed of dog. How would you decide which one of these two variables to use as a blocking variable?

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.
- (a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?
- A random sample of 15 fish having a mean length that is greater than 10 inches
or
 - A random sample of 50 fish having a mean length that is greater than 10 inches
- Justify your answer.
- (b) Suppose the standard deviation of the sampling distribution of the sample mean for random samples of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.
- (c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b) ?
Justify your answer.



4. The graph below displays the relative frequency distribution for X , the total number of dogs and cats owned per household, for the households in a large suburban area. For instance, 14 percent of the households own 2 of these pets.



- According to a local law, each household in this area is prohibited from owning more than 3 of these pets. If a household in this area is selected at random, what is the probability that the selected household will be in violation of this law? Show your work.
- If 10 households in this area are selected at random, what is the probability that exactly 2 of them will be in violation of this law? Show your work.
- The mean and standard deviation of X are 1.65 and 1.851, respectively. Suppose 150 households in this area are to be selected at random and \bar{X} , the mean number of dogs and cats per household, is to be computed. Describe the sampling distribution of \bar{X} , including its shape, center, and spread.

5. There are 4 runners on the New High School team. The team is planning to participate in a race in which each runner runs a mile. The team time is the sum of the individual times for the 4 runners. Assume that the individual times of the 4 runners are all independent of each other. The individual times, in minutes, of the runners in similar races are approximately normally distributed with the following means and standard deviations.

	Mean	Standard Deviation
Runner 1	4.9	0.15
Runner 2	4.7	0.16
Runner 3	4.5	0.14
Runner 4	4.8	0.15

- (a) Runner 3 thinks that he can run a mile in less than 4.2 minutes in the next race. Is this likely to happen? Explain.
- (b) The distribution of possible team times is approximately normal. What are the mean and standard deviation of this distribution?
- (c) Suppose the team's best time to date is 18.4 minutes. What is the probability that the team will beat its own best time in the next race?



Investigative Task

6. A study was designed to explore subjects' ability to judge the distance between two objects placed in a dimly lit room. The researcher suspected that the subjects would generally overestimate the distance between the objects in the room and that this overestimation would increase the farther apart the objects were.

The two objects were placed at random locations in the room before a subject estimated the distance (in feet) between those two objects. After each subject estimated the distance, the locations of the objects were rerandomized before the next subject viewed the room.

After data were collected for 40 subjects, two linear models were fit in an attempt to describe the relationship between the subjects' perceived distances (y) and the actual distance, in feet, between the two objects.

$$\text{Model 1: } \hat{y} = 0.238 + 1.080 \times (\text{actual distance})$$

The standard errors of the estimated coefficients for Model 1 are 0.260 and 0.118, respectively.

$$\text{Model 2: } \hat{y} = 1.102 \times (\text{actual distance})$$

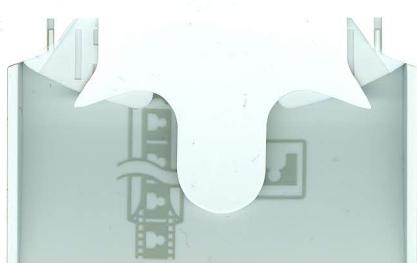
The standard error of the estimated coefficient for Model 2 is 0.393.

- Provide an interpretation in context for the estimated slope in Model 1.
- Explain why the researcher might prefer Model 2 to Model 1 in this context.
- Using Model 2, test the researcher's hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)

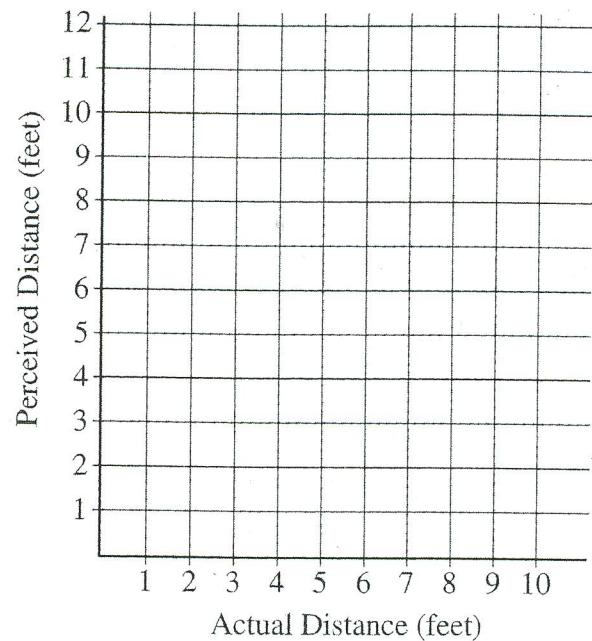
The researchers also wanted to explore whether the performance on this task differed between subjects who wear contact lenses and subjects who do not wear contact lenses. A new variable was created to indicate whether or not a subject wears contact lenses. The data for this variable were coded numerically (1 = contact wearer, 0 = noncontact wearer), and this new variable, named "contact," was included in the following model.

$$\text{Model 3: } \hat{y} = 1.05 \times (\text{actual distance}) + 0.12 \times (\text{contact}) \times (\text{actual distance})$$

The standard errors of the estimated coefficients for Model 3 are 0.357 and 0.032, respectively.



- (d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.



- (e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.