

THE UNIVERSITY OF WESTERN ONTARIO
DEPARTMENT OF STATISTICAL AND ACTUARIAL SCIENCES
STATISTICS 1024A MAKE-UP EXAM

Friday, January 11th, 2013, 7:00 PM - 10:00 PM

EXAM CODE 200

INSTRUCTIONS:

- This is a closed book test. A formula sheet and statistical tables are attached.
- Use only an HB pencil for the Scantron sheet.
- Fill in NAME, INSTRUCTOR, SIGNATURE and COURSE on the Scantron sheet.
- Fill in and **code** your STUDENT NUMBER, SECTION and EXAM CODE on the Scantron sheet. Your EXAM CODE is 200.
- Leave the ANSWER SHEET NUMBER blank on the Scantron sheet.
- There are 50 multiple choice questions on pages 2 to 23. A blank page for rough work, and the formula sheet, are at the end of the paper.
- **Code** your answers on the Scantron sheet and submit it.
- Only non-programmable calculators are permitted.
- No other electronic devices are allowed. If any student is found with an electronic device on their person at any time during the test, that student will automatically receive a mark of 0 for the test.
- **NO EXTRA TIME WILL BE GIVEN TO CODE YOUR ANSWERS!!**

GOOD LUCK!

- 1) To get the best deal on a GPS, Tom called four electronics stores and asked the cost of a specific model. The prices he was quoted are listed below:

\$65 \$35 \$55 \$35

The standard deviation of the prices is:

- (A) \$15.00 (B) \$12.99 (C) \$12.50 (D) \$16.67
- 2) Here is a stemplot of the average travel times to work each day for workers in the United States for the 50 states and the District of Columbia among workers who are at least 16 years of age and don't work at home. The stems are whole minutes and the leaves are tenths of minutes.

Stem	Leaf
15	59
16	
17	6779
18	25
19	
20	017889
21	28
22	01333499
23	345669
24	01266
25	0012569
26	689
27	39
28	
29	12
30	69

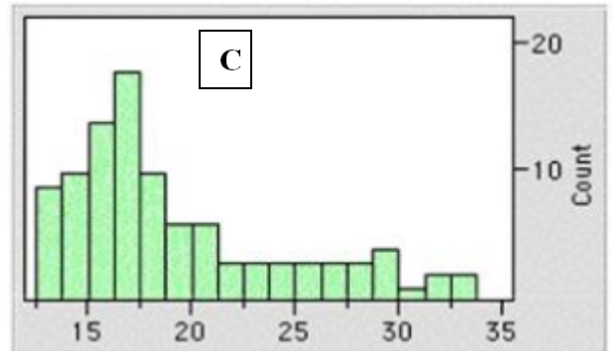
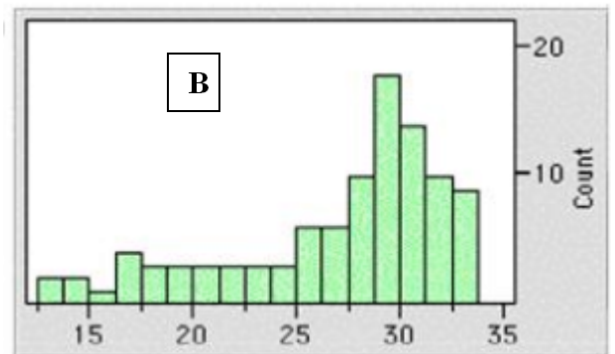
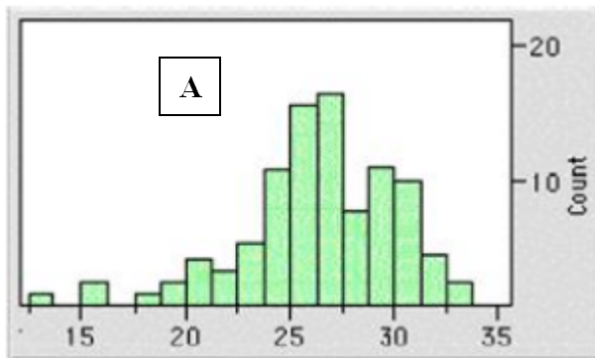
On average, what is the median of these average travel times to work each day?

- (A) 23.3 minutes (B) 23.35 minutes (C) 23.5 minutes (D) 23.4 minutes

- 3) A study examined the length of a certain species of fish from one lake. The plan was to take a random sample of 100 fish and examine the results. Numerical summaries on lengths of the fish measured in this study are given.

Mean	26.8 mm
Median	29.4 mm
Standard Deviation	5.0 mm
Minimum	12.0 mm
Maximum	33.4 mm

Which of the following histograms is most likely to be the one for these data?



(A) Histogram A.

(B) Histogram B.

(C) Histogram C.

- 4) A reporter wishes to portray baseball players as overpaid. Which measure of center should he report as the central value for salaries of major league players?
- (A) The standard deviation
 - (B) The mean
 - (C) The median
 - (D) Either the mean or the median; it doesn't matter since they will be equal.
- 5) Which of the following values is closest to the 20th percentile for the standard normal distribution?
- (A) 0.84 (B) -2.1 (C) 2.1 (D) -0.84
- 6) What value is closest to the interquartile range for the standard normal distribution?
- (A) 0.5 (B) 1.3 (C) 0 (D) 3
- 7) The scores of adults on an IQ test are approximately Normal with mean 100 and standard deviation 15. Clara scores 127 on such a test. She scores higher than what percent of all adults?
- (A) About 90% (B) About 4% (C) About 96% (D) About 10%

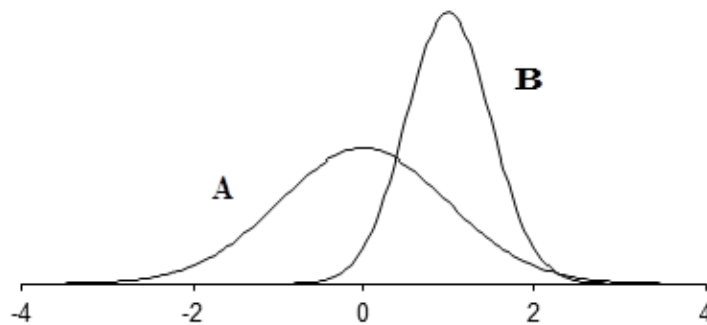
- 8) Suppose that the amount spent by students on textbooks this semester has a Normal distribution with mean $\mu = \$500$ and standard deviation $\sigma = \$100$. What proportion of the amounts spent on textbooks falls between \$450 and \$600?

(A) 0.3612 (B) 0.0796 (C) 0.5328 (D) 0.6826

- 9) The heights of women aged 20 to 29 follow approximately the $\mathcal{N}(64.3, 2.7)$ distribution. Men the same age have heights distributed as $\mathcal{N}(69.9, 3.1)$. What percent of young women are taller than the mean height of young men?

(A) 0.0192 (B) 0.9808 (C) 0.0351 (D) 0.9649

- 10) Two normal density curves A and B are given in the graph below:



If μ_A and σ_A are the mean and standard deviation from density curve A and μ_B and σ_B are the mean and standard deviation from density curve B, which of the following statements is true?

- (A) μ_A is greater than μ_B and σ_A is less than σ_B
(B) μ_A is less than μ_B and σ_A is less than σ_B
(C) μ_A is less than μ_B and σ_A is greater than σ_B
(D) μ_A is greater than μ_B and σ_A is greater than σ_B

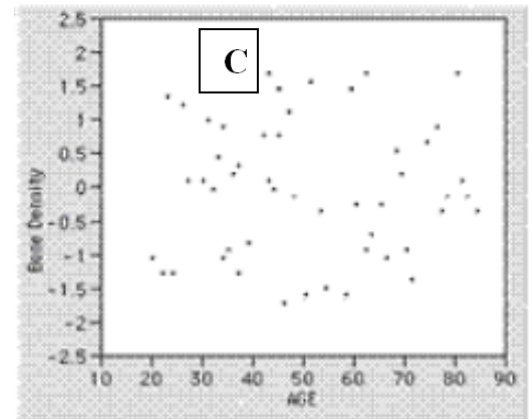
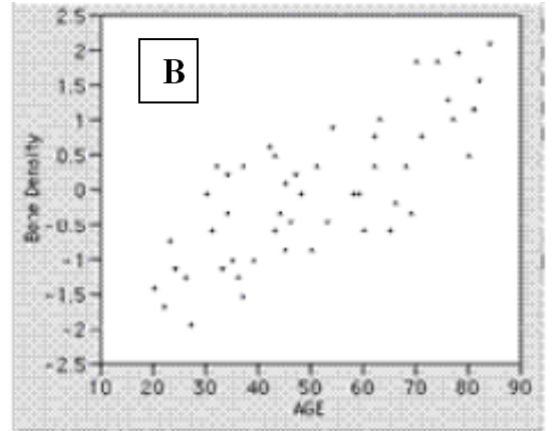
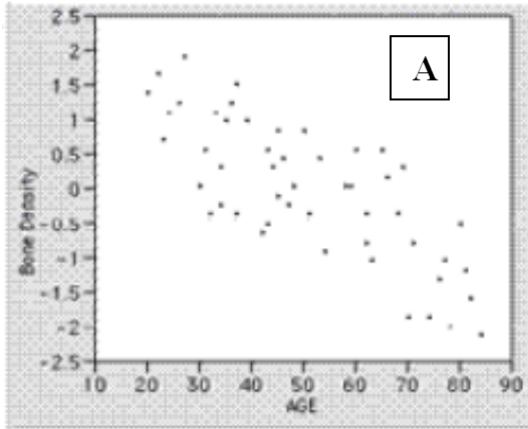
11) Which correlation is the strongest?

- (A) 0.1 (B) -1 (C) 0.9 (D) -0.95

12) A positive correlation is present when

- (A) one variable goes up and one goes down.
- (B) several variables never change.
- (C) two variables move in the same direction.
- (D) two variables move in opposite directions.

13) Bone density is typically measured as a standardized score with a mean of 0 and a standard deviation of 1. Lower scores correspond to lower bone density. Which of the following graphs shows that as women grow older they tend to have lower bone density?



(A) Graph A.

(B) Graph B.

(C) Graph C.

- 14) Researchers wished to determine whether individual differences in introspective ability are reflected in the anatomy of brain regions responsible for this function. They measured introspective ability (using a score on a test of introspective ability, with larger values indicating greater introspective ability) and gray-matter volume in milliliters (the Brodmann area) in the anterior prefrontal cortex of the brain of 29 subjects. The researchers wished to determine the equation of the least-squares regression line for predicting introspective ability (y) from gray-matter volume (x). To do this, they calculated the following summary statistics:

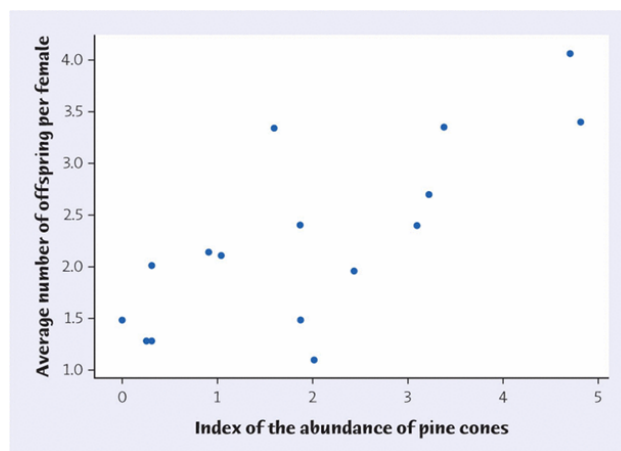
$$\bar{x} = 0.649, s_x = 0.045$$

$$\bar{y} = 65.86, s_y = 8.69$$

$$r = 0.448$$

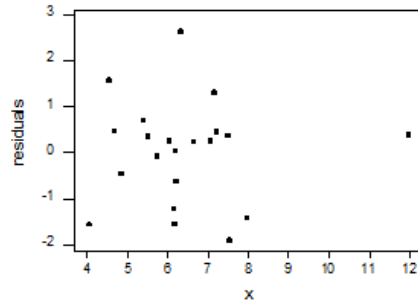
Based on the least-squares regression line, what would you predict introspective ability to be for someone with gray-matter volume 0.60?

- (A) 67.81 (B) 61.62 (C) 60.05 (D) 55.45 (E) 64.3
- 15) The following is a scatterplot of average number of offspring per female against cone index. Which of the following is a plausible value of the correlation, r , between average number of offspring per female and cone index?



- (A) 0.03 (B) 0.6 (C) 0.95 (D) -0.65

- 16) A regression was carried out on data and below is a plot of the residuals versus the explanatory variable.



Which of the following statements is true?

- (A) There is an outlier.
 - (B) The data exhibit a nonlinear relationship.
 - (C) There is an influential observation.
 - (D) The response variable has more spread for larger values of the explanatory variable.
- 17) Researchers compared 665 men who had been admitted to a hospital with their first heart attack to 772 men in the same age group (21 to 54 years old) who had been admitted to a hospital for other reasons. There were a total of 35 hospitals in the research study. The researchers found that the percentage of men who showed some degree of pattern baldness was substantially higher for those who had a heart attack (42%) than for those who had not (34%). This is an example of
- (A) a stratified random sampling design with hospitals as strata.
 - (B) a comparative experiment.
 - (C) an observational study.
 - (D) the elimination of all confounding variables by design, since the researchers picked men in the same age group who had been admitted to hospital group as the control.

- 18) Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers give the new medication to 50 adult volunteers who suffer from allergies; 44 of these volunteers report a significant reduction in their allergy symptoms without any drowsiness. This study could be improved by
- (A) including people who do not suffer from allergies in the study in order to represent a more diverse population.
 - (B) repeating the study with only the 44 volunteers who reported a significant reduction in their allergy symptoms without any drowsiness, and giving them a higher dosage this time.
 - (C) using a control group.
 - (D) all of the above.
- 19) A college official conducted a survey to estimate the proportion of students currently living in dormitories about their preference for single rooms, double rooms, or multiple (more than two people) rooms in the dormitories on campus. Which of the following does NOT affect the college official's ability to generalize the survey results to all dormitory students?
- (A) Five thousand students live in dormitories on campus. A random sample of only 500 were sent the survey.
 - (B) The survey was sent to only first-year students.
 - (C) Of the 500 students who were sent the survey, only 160 responded.
 - (D) All of the above present a problem for generalizing the results.

- 20) As part of a database on the new births at a hospital, some variables recorded are the age of the mother, marital status of the mother (single, married, divorced, other), blood type of the mother, weight of the baby, and sex of the baby. Of these variables, which of the following is true?
- (A) sex, marital status, and age are categorical variables.
 - (B) sex, marital status, and blood type are categorical variables.
 - (C) age, blood type, and weight are quantitative variables.
 - (D) age and weight are categorical variables.
- 21) In a study of human development, investigators showed two movies that were of different types to a group of children. Crackers were available in a bowl at each movie, and the investigators compared the number of crackers eaten by children watching each movie. One movie was shown at 8 A.M. (right after the children had breakfast) and the other at 11 A.M. (right before the children had lunch). It was found that during the movie shown at 11 A.M., more crackers were eaten than during the movie shown at 8 A.M. The investigators concluded that the different types of movies had different effects on appetite. The results cannot be trusted because
- (A) the investigators should have used several bowls of crackers randomly placed in the room.
 - (B) the time each movie was shown is a confounding variable.
 - (C) the study was not double-blind. Neither the investigators nor the children should have been aware of which movie was being shown.
 - (D) the investigators were biased. They knew beforehand what the study would show.
 - (E) None of the other answers given.

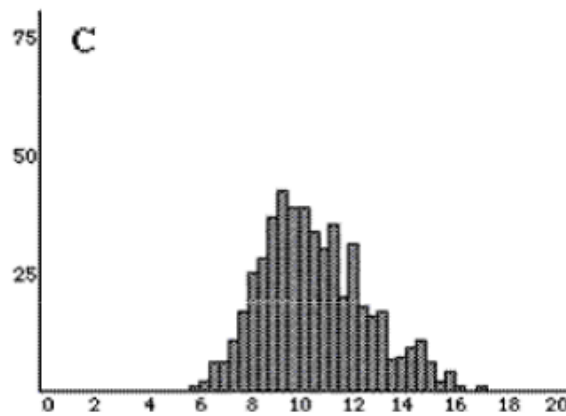
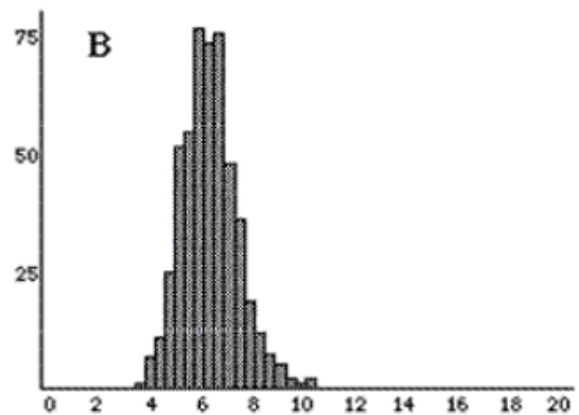
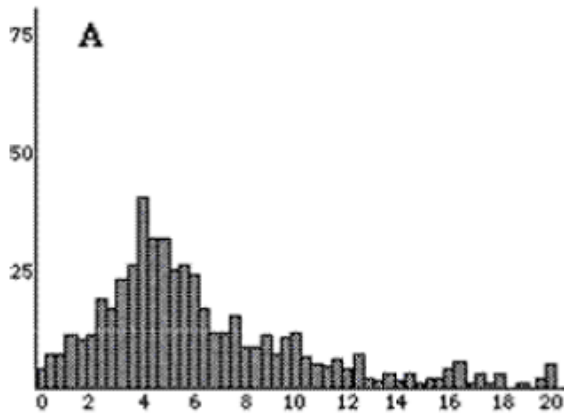
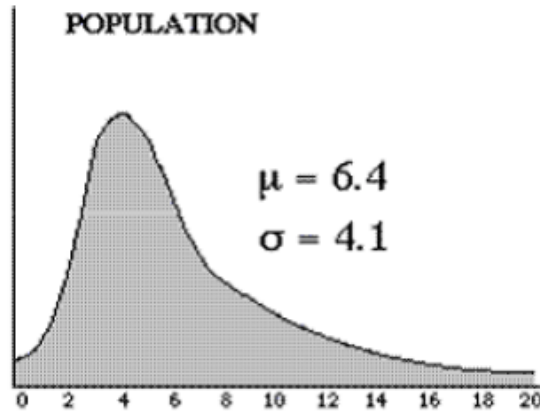
- 22)** A study attempts to determine whether a football filled with helium travels farther when kicked than one filled with air. Each subject kicks twice, once with a football filled with helium and once with a football filled with air. The order of the type of football kicked is randomized. This is an example of
- (A) a randomized controlled experiment.
 - (B) a stratified experiment.
 - (C) a matched pairs experiment.
 - (D) a randomized block experiment.
 - (E) the placebo effect.
- 23)** City records show that 35% of all residents are over the age of 60. To test this claim, 50 callers to a radio talk show with the theme “our aging population” are asked their age. Of these callers, 40% were over the age of 60. Assume all the callers responded and cooperated. Bias may arise in the data as a result of:
- (A) the sample being too small.
 - (B) using a voluntary response sample.
 - (C) nonresponse.
 - (D) undercoverage.

- 24) A student participates in a Coke versus Pepsi taste test. She correctly identifies which soda is which four times out of six tries. She claims that this proves that she can reliably tell the difference between the two soft drinks. You have studied statistics and you want to determine the probability of anyone getting at least four right out of six tries just by chance alone. Which of the following would provide an accurate estimate of that probability?
- (A) Have the student repeat this experiment many times and calculate the percentage time she correctly distinguishes between the brands.
 - (B) Simulate this on the computer with a 50% chance of guessing the correct soft drink on each try, and calculate the percent of times there are four or more correct guesses out of six trials.
 - (C) Repeat this experiment with a very large sample of people and calculate the percentage of people who make four correct guesses out of six tries.
 - (D) All of the methods listed above would provide an accurate estimate of the probability.
- 25) C and G are two events with $P(C) = 0.35$ and $P(G) = 0.25$. Assuming that C and G are independent, what is the $P(C \text{ and not } G)$?
- (A) 0.1 (B) 0.3534 (C) 0.4035 (D) 0.2625 (E) None of the other answers given.
- 26) A roulette wheel has 38 slots on which the ball can land. Two of the slots are green, 18 are red, and 18 are black. The ball is equally likely to land in any slot. The roulette wheel is going to be spun twice and the outcomes of the two spins are independent. The probability that it lands one time on red and another time on black in any order is
- (A) 0.2244 (B) 0.25 (C) 0.5 (D) 0.4488

- 27) A multiple choice exam consists of 10 questions, each having 5 possible answers to choose from. If you guess at every question independently, the probability that you get at least one correct answer is
- (A) 0.998 (B) 0.107 (C) 0.002 (D) 0.893
- 28) A standard deck of 52 playing cards will have cards in four suits (clubs, diamonds, hearts and spades) each in 13 denominations (ace, two, etc. to ten, jack queen and king). Five cards are dealt off the top of a well-shuffled deck. The chance that the 5th card is a club given that the first 4 were clubs is
- (A) 0.1875 (B) 0.1702 (C) 0.1731 (D) 0.1042
- 29) In order to estimate the average price of TV's in London, a researcher decided to increase the size of his random sample from 100 to 400 TV's. The effect of this increase
- (A) has no effect because the population size is the same.
(B) reduces the bias of the estimate.
(C) increases the standard error of the estimate.
(D) reduces the variability of the estimate
(E) increases the confidence interval width for the parameter.

Questions 30 and 31 refer to the following situation.

Four graphs are presented below. The graph at the top is a distribution for a population of test scores. The mean score is 6.4 and the standard deviation is 4.1.



30) Which graph (A, B, or C) do you think represents a single random sample of 500 values from this population?

(A) Graph A

(B) Graph B

(C) Graph C

31) Which graph (A, B, or C) do you think represents a distribution of 500 sample means from random sample each of size 9?

(A) Graph A

(B) Graph B

(C) Graph C

32) Let the random variable \bar{x} be the mean weight of a random sample of 100 runners randomly selected from a large population of runners with mean weight, $\mu = 63.3kg$, and standard deviation, $\sigma = 3kg$. The Central Limit Theorem states that the sampling distribution of \bar{x} is approximately

(A) $\mathcal{N}(63.3, 0.03)$

(B) $\mathcal{N}(63.3, 0.3)$

(C) $\mathcal{N}(0, 1)$

(D) $\mathcal{N}(6.33, 0.3)$

33) A research article reports the results of a new drug test. The drug is to be used to decrease vision loss in people with Macular Degeneration. The article gives a p-value of .04 in the analysis section. Indicate if the following interpretation of this p-value is valid or invalid.

The probability of getting results as extreme as or more extreme than the ones in this study if the drug is actually not effective.

(A) Valid.

(B) Invalid.

- 34) A first-year statistics class wants to estimate the average number of chocolate chips in a generic brand of chocolate chip cookies. They collect a random sample of cookies, count the chips in each cookie, and calculate a 95% confidence interval for the average number of chips per cookie (18.6 to 21.3). Indicate if the following interpretation is valid or invalid.

We would expect about 95% of all possible sample means from this population to be between 18.6 and 21.3 chocolate chips.

- (A) Valid.
- (B) Invalid.
-
- 35) The following situation models the logic of a hypothesis test. An electrician uses an instrument to test whether or not an electrical circuit is defective. The instrument sometimes fails to detect that a circuit is good and working. The null hypothesis is that the circuit is good (not defective). The alternative hypothesis is that the circuit is not good (defective). If the electrician rejects the null hypothesis, which of the following statements is true?
- (A) The electrician decides that the circuit is defective, but it could be good.
- (B) The circuit is definitely not good and needs to be repaired.
- (C) The circuit is most likely good, but it could be defective.
- (D) The circuit is definitely good and does not need to be repaired.
-
- 36) A paint manufacturer advertises that one gallon of its paint will cover 400 square feet of an interior wall. Some local painters suspect the average coverage is considerably less and decide to conduct an experiment to find out. Which of the following are the correct null and alternative hypotheses to be tested?
- (A) $H_0 : \mu = 400, H_a : \mu \neq 400$
- (B) $H_0 : \mu = 400, H_a : \mu > 400$
- (C) $H_0 : \mu \neq 400, H_a : \mu < 400$
- (D) $H_0 : \mu = 400, H_a : \mu < 400$

- 37) It has been established that under normal environmental conditions, adult largemouth bass in Silver Lake have an average length of 12.3 inches with a standard deviation of 3 inches. People who have been fishing Silver Lake for some time claim that this year they are catching smaller than usual largemouth bass. A research group from the Department of Natural Resources took a random sample of 100 adult largemouth bass from Silver Lake and found the mean of this sample to be 11.2 inches. Which of the following is the most appropriate statistical conclusion?
- (A) The researchers cannot conclude that the fish are smaller than what is normal because 11.2 inches is less than one standard deviation from the established mean (12.3 inches) for this species.
 - (B) The researchers can conclude that the fish are smaller than what is normal because the sample mean should be almost identical to the population mean with a large sample of 100 fish.
 - (C) The researchers can conclude that the fish are smaller than what is normal because the difference between 12.3 inches and 11.2 inches is much larger than the expected sampling error.
- 38) You are thinking of employing a t-procedure to test hypotheses about the mean of a population using a significance level of 0.05. You suspect the distribution of the population is not Normal and may be moderately skewed. Which of the following statements is correct?
- (A) You should not use a t-procedure, because the population does not have a Normal distribution.
 - (B) You may not use the t-procedure, because t-procedures are robust to non-Normality in the data.
 - (C) You may use the t-procedure, but you should probably claim the significance level is only 0.10.
 - (D) You may use the t-procedure, provided your sample size is large, say, at least 50.

- 39) Suppose a confidence interval for the population mean was 62.84 to 69.46. The population standard deviation was assumed to be 6.50, and a sample of 100 observations was used. The mean of the sample was
- (A) 66.2 (B) 56.34 (C) 62.96 (D) 66.15
- 40) The owner of a local nightclub has recently surveyed a random sample of $n = 300$ customers of the club. She would now like to determine whether or not the mean age of her customers is over 35. If so, she plans to alter the entertainment to appeal to an older crowd. If not, no entertainment changes will be made. Suppose she found that the sample mean was 35.5 years and the standard deviation was 5 years. What is the approximate p-value associated with the test statistic?
- (A) 0.0418 (B) 0.9582 (C) 0.0836 (D) 0.4602
- 41) A random sample of the weights of 20 bags of chips was taken and the sample mean was calculated to be $\bar{x} = 404$ grams and the standard deviation $s = 1.2$ grams. A 95% confidence interval for the population mean is
- (A) 404 ± 0.709 (B) 404 ± 0.441 (C) 404 ± 0.091 (D) 404 ± 0.562

42) An industrial plant claims to discharge no more than 1000 liters of wastewater per hour, on average into a river beside the plant. An environmental action group decides to monitor the plant, in case this limit is being exceeded. Doing so is expensive and only a small sample is possible. A random sample of four hours is selected over a period of a week with the results that the sample mean $\bar{x} = 2000$ and the sample standard deviation $s = 816.5$. For testing $H_0 : \mu = 1000$ against $H_a : \mu > 1000$, the results of the study are

- (A) significant at the 0.20 level but not at the 0.10 level.
- (B) significant at the 0.10 level but not at the 0.05 level.
- (C) significant at the 0.01 level but not at the 0.001 level.
- (D) significant at the 0.05 level but not at the 0.01 level.

43) To assess the accuracy of a laboratory scale, a standard weight that is known to weigh 1 gram is repeatedly weighed a total of n times and the mean \bar{x} of the weighings is computed. Suppose the scale readings are Normally distributed with an unknown mean μ and standard deviation $\sigma = 0.01$ grams. How large should n be so that a 95% confidence interval for μ has a margin of error of 0.0001 or less?

- (A) 10000 (B) 100 (C) 38416 (D) 196

44) The plasma ascorbic acid levels of pregnant women were compared for smokers and nonsmokers. Ten women, five smokers and five nonsmokers, who were in the last three months of pregnancy, free of major health disorders, and ranging in age from 15 to 32, years were selected for the study. Prior to the collection of 20 ml of blood, the participants were told to avoid breakfast, forego their vitamin supplements, and avoid foods high in ascorbic acid content. From the blood samples, the following plasma ascorbic acid values of each subject were determined in milligrams per 100 milliliters:

Plasma Ascorbic Acid Level			
	Mean	s	n
Nonsmokers	0.964	0.230	5
Smoker	0.726	0.180	5
Difference	0.238	0.258	5

Using the conservative choice for the degrees of freedom, what is the 95% confidence interval for the difference between the plasma levels between nonsmokers and smokers?

- (A) -0.125 to 0.601
- (B) -0.082 to 0.558
- (C) -0.059 to 0.535
- (D) -0.098 to 0.574
- (E) None of the Above

45) IQ scores were available for 113 men who at birth had very low birth weight (VLBW) and for 106 men in the control group. The mean IQ for the 113 men in the VLBW group was 87.6, and the standard deviation was 15.1. The 106 men in the control group had mean IQ 94.7, with standard deviation 14.9. To test that there is a good evidence that mean IQ is lower among VLBW men than among controls from similar backgrounds, the test statistic would be

- (A) $t = -3.50$ (B) $t = -1.72$ (C) $t = -4.20$ (D) $t = -5.00$ (E) $t = -7.10$

46) A survey claims that 9 out of 10 doctors recommend aspirin for their patients with headaches. To test this claim against the alternative that the actual proportion of doctors who recommend aspirin is less than 0.90, a random sample of 100 doctors' results in 83 who indicate that they recommend aspirin. The value of the test statistic is (A) -1.86 (B) -2.33 (C) -2.18 (D) -1.67

47) If we want to estimate p , the population proportion of likely voters who believe the state of the economy is the most urgent national concern, with 99% confidence and a margin of error no greater than 2%, how many likely voters need to be surveyed? Assume that you have no idea of the value of p .

- (A) 3483 (B) 4148 (C) 4147 (D) 3484 (E) 4146

- 48) A college president is hoping that the proportion of students enrolled at her college who binge drink is actually lower than the national proportion of 0.41. In a commissioned study, 348 students are selected randomly from a list of all students enrolled at the college. Of these, 132 admit to having engaged in binge drinking. To testing her hope that the proportion of students at her college who engage in binge drinking is lower than the national proportion of 0.41, her staff tests the hypotheses $H_0 : p = 0.41$ versus $H_a : p < 0.41$. The p-value for this test is
- (A) below 0.01.
 - (B) between 0.05 and 0.10.
 - (C) between 0.15 and 0.20.
 - (D) between 0.01 and 0.05.
 - (E) between 0.10 and 0.15.
- 49) A local board of education conducted a survey of residents in the community concerning a property tax levy on the coming local ballot. They randomly selected 850 residents in the community and contacted them by telephone. Of the 850 residents surveyed, 410 supported the property tax levy. Let p represent the proportion of residents in the community that support the property tax levy. Using the plus 4 rule, a 90% confidence interval for p is
- (A) 0.4489 to 0.5159.
 - (B) 0.4543 to 0.5106.
 - (C) 0.4487 to 0.5161.
 - (D) 0.4463 to 0.5185.

- 50) A government survey randomly selected 68 female high school students and 71 male high school students. Of these students, 10 females and 19 males played video or computer games for three or more hours a day. Using the plus 4 rule, a 90% confidence interval for the difference in proportions of male and female high school students who play video or computer games for at least three hours a day is
- (A) 0.007 to 0.234.
 - (B) 0.009 to 0.232.
 - (C) 0.002 to 0.232.
 - (D) 0.005 to 0.229.

Use this page for rough work

Answer Key

1 6 11 16 21 26 31 36 41 46
ADBCD BCCAC BCABB ACCAB BCBBB DDADA BBABA DCDDA DDCAA BBEBD