Math 1530 Final Exam Spring 2017

Name _____
Section # _____
Instructor _____

There are five possible responses to each of the following multiple choice questions. There is only one "BEST" answer. Be sure to read all possible choices before selecting your answer. You may mark on this examination. You can use a calculator but a calculator manual cannot be used.

Form B

Please circle your answer to each question and fill the blank sheet. After you finish the exam, log into D2L and input your answers under the right quiz item. There are a total of 43 questions.

- If your test is Form A, take Quiz item *Final Form A*.
- If your test is Form B, take Quiz item *Final Form B*.
- If your test is Form C, take Quiz item *Final Form C*.
- If your test is Form D, take Quiz item *Final Form D*.

CONFIDENCE LEVEL	90%	95%	99%
z^*	1.645	1.96	2.576

1. Which of these questions from the Spring 2017 MATH1530 class survey produced variables that are quantitative?

   i. Are you satisfied with America's law or policies on guns? (Yes or No)
   ii. Approximately, how many phone calls do you make per day?
   iii. What is your ideal number of children?
   iv. Are you satisfied with America's security with terrorism? (Yes or No)
   v. Are you satisfied with America's state of race relations? (Yes or No)

   (A) i, iv, v (B) ii, iii (C) iii (D) ii (E) i

   Use the following for the next 2 questions. The following data (sorted) are the percents of state residents who were born outside the United States.

1.2	1.8	1.9	2.1	2.2	2.7	2.7	2.8	2.9	3.2	3.3	3.6	3.8
3.8	3.9	3.9	4.1	4.2	4.4	4.9	5.1	5.4	5.6	5.6	5.9	6.3
6.6	6.9	7.0	8.1	8.3	9.2	9.7	10.1	10.1	10.3	12.2	12.4	12.6
12.7	12.9	13.8	14.1	15.1	15.9	16.3	18.9	19.1	20.1	21.6	27.2	

2. Which of the following best describes the distribution?

   (A) The distribution is right skewed with a peak which lies between 0% and 5% of residents that are foreign born. There is a possible outlier.
   (B) The distribution is strongly skewed left and has several peaks. The center of the distribution is around 10% with a spread of 20%.
   (C) Ignoring the outlier(s), it is close to symmetric. The mean will be close to the median.
   (D) The distribution is double-peaked and has an outlier.
   (E) Flat, resembles a uniform distribution.

3. The center of the distribution is close to
   (A) 50%. (B) 10%. (C) 6.3%. (D) 3.8%. (E) the modes: 2.7%, 3.8%, 3.9%, 5.6%, 10.1%.

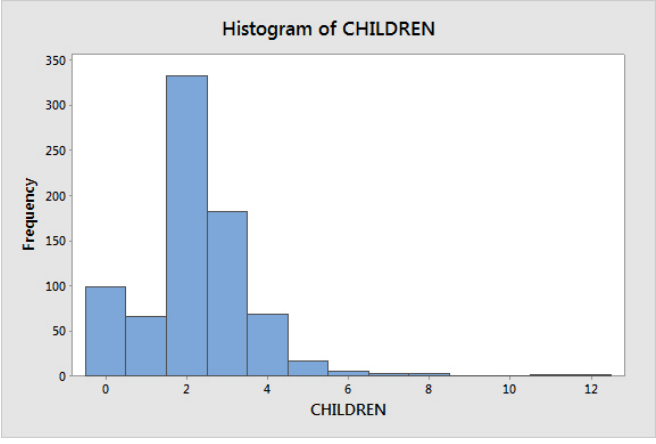
4. Here are the number of hours that each of a group of students studied for an exam:

   2 4 22 2 1 4 1 5 5 4

   What is the mean number of study hours?

   (A) 50 (B) 5 (C) 4.5 (D) 4 (E) 2.5

Use the following for the next 2 questions. The Spring 2017 MATH1530 survey asked "What is the ideal number of children?" The figure below represents the responses of 775 students.



5. Which of the following best describes this distribution? This distribution is

   (A) symmetric around 2, very little variation, and many outliers.

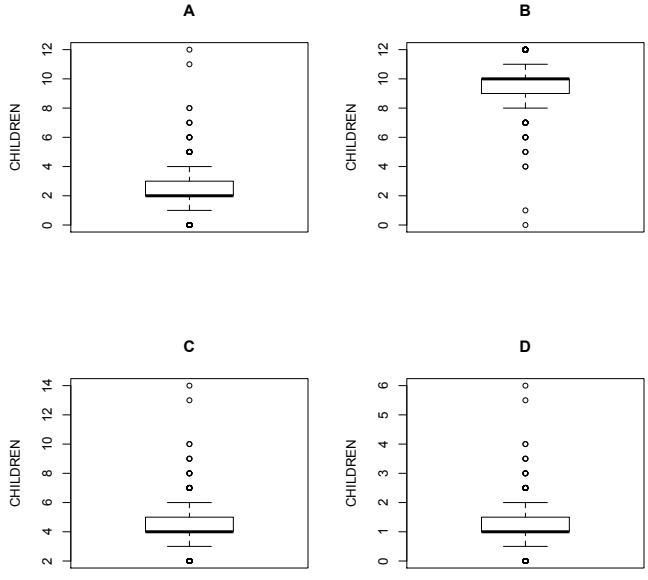
   (B) multi-peaked with the bulk of data from 2 to 4.

   (C) evenly-spaced from 0 to 12.

   (D) right-skewed, with the center around 2 children and spread from 0 to 12 children, with possible outlier(s).

   (E) left-skewed with outliers.

6. Which of the following boxplots corresponds to the above histogram?



7. The scores of an easy test range from 0 to 100. There are a few scores below 50, some scores ranging from 50 to 80, and many scores above 80. The distribution of test scores will be

   (A) roughly symmetric. (C) correlated to age of the student.

   (B) roughly normal. (D) skewed to the right.

   (E) skewed to the left.

8. The *Three Stooges* were a comedy act that made a series of 190 short films over the period from 1934 to 1958. Their comedy was centered around slapstick and cartoonish violence. If we count the average number of times per film that someone was slapped over a 24-year period, then we get the following numbers (listed in order from smallest to largest):

6.4	7.6	8.5	8.9	9.9	10.2	10.5	10.6	11.4	11.6	12.2	13.6
13.7	14.0	14.4	14.8	15.6	16.0	17.8	19.6	21.1	23.1	31.9	33.5

   The five-number summary of the data is

   (A) $1934, 1940.5, 1946.5, 1952.5, 1958$ (C) $6.4, 10.35, 22.95, 16.9, 33.5$

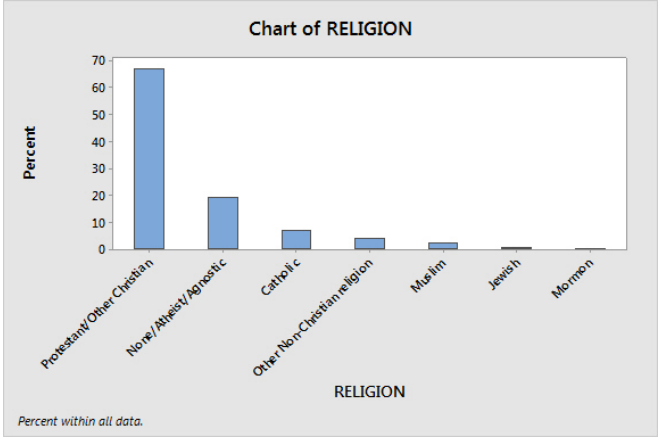
   (B) $6.4, 10.35, 13.65, 16.9, 33.5$ (D) $6.4, 10.35, 13.6, 13.7, 16.9, 33.5$

   (E) $1, 6.5, 12.5, 18.5, 24$

9. The number of deaths among persons aged 15 to 24 years in the United States in 2003 due to the leading causes of death for this age group were: accidents, 14,966; homicide, 5148; suicide, 3921; cancer, 1628; heart disease, 1083; congenital defects, 425. To display these data you should construct a

   (A) stemplot. (B) time plot. (C) histogram. (D) 5-number summary. (E) bar graph.

10. The Spring 2017 MATH1530 survey asked "What is your religious identification?" The figure below represents the responses of 775 students.



   Which of following best describes this distribution?

   (A) The distribution is strongly skewed right.

   (B) The distribution is strongly skewed left.

   (C) The distribution is evenly spaced from Protestant/Other Christian to Mormon.

   (D) The distribution is skewed right, the center is Other/Non-Christian religion and the range is from Protest/Other Christian to Mormon.

   (E) Approximately 2 times the number of students identify as Protestant/Other Christian than all other religious classes combined.

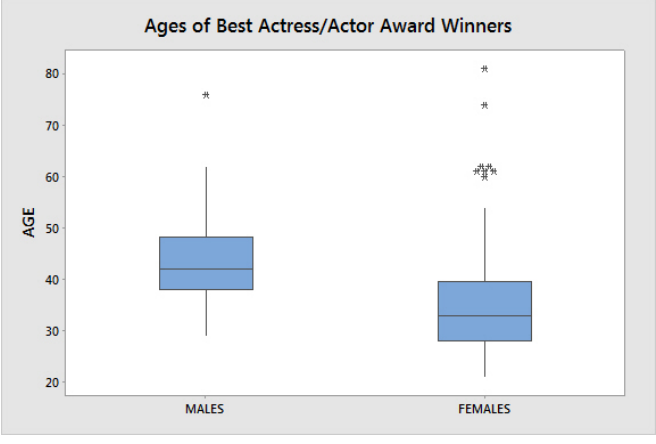
11. The Spring 2017 MATH1530 survey asked students to identify their classification in college and whether or not they are satisfied with America's law or policies on guns. The survey results are given in the following table:

	Year			
Satisfied	Freshman	Sophomore	Junior	Senior
YES	289	101	45	16
NO	200	62	31	31

To compare the four classes (year in school) with respect to their opinion regarding America's law or policies on guns, which distribution should we calculate?

(A) The joint distribution of year in school and opinion.
(B) The marginal distribution of year.
(C) The conditional distribution of opinion given year in school.
(D) The conditional distribution of year in school given opinion.
(E) Freshman students are more in favor of current gun laws/policies than any other class.

12. Here are the boxplots of the ages of all Oscar winners for the Best Actress and Best Actor.



Which of the following best describes the plot?

(A) There is less diversity in age among the Best Actresses than among the Best Actors.
(B) The oldest person winning a Best Actor/Actress award was a man.
(C) The youngest man winning a Best Actor award is younger than the youngest woman winning a Best Actress award.
(D) The Best Actresses are generally younger than the Best Actors.
(E) The margin of error in age among the Best Actresses is larger than the margin of error in age among the Best Actors.

Use the following for the next 2 questions. Suppose that the blood alcohol count (BAC) of students who drink 5 beers varies from student to student according to a normal distribution with mean 0.08 and standard deviation 0.01.

13. What percent of students who drink 5 beers have a BAC above 0.08 (the legal limit for driving in all states)?
(A) 50% (B) 32% (C) 16% (D) 5% (E) 2.5%

14. The middle 95% of students who drink 5 beers have a BAC between
(A) 0.07 and 0.09. (B) 0.05 and 0.11. (C) 0.04 and 0.12. (D) 0.06 and 0.10. (E) 0% and 8%.

15. You see a student standing in front of the student center, now and then stopping other students to ask them questions. She says that she is collecting student opinions for a class assignment. Identify the type of sampling method.

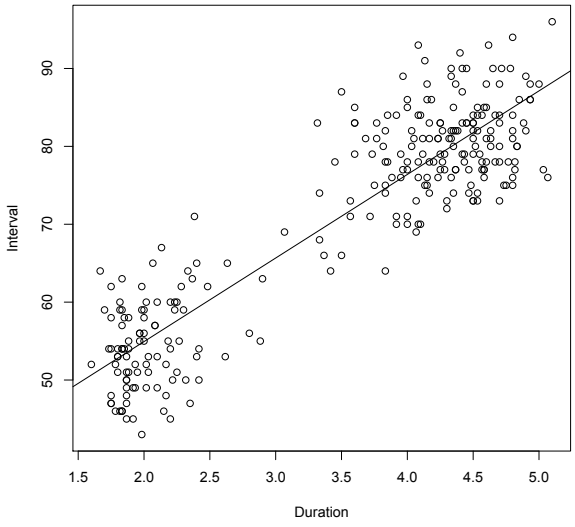
    (A) simple random sampling (C) convenience sampling

    (B) bias random sampling (D) stratified random sampling

    (E) census

    Use the following for the next 3 questions. Can we predict the next eruption of the Old Faithful geyser by knowing the duration of the previous eruption? The observations and least-squares regression line appear in the scatterplot. The correlation between the two variables is $r = 0.901$ and the least-squares regression line for predicting the *Interval* (min.) between eruptions and the *Duration* (min.) of the previous eruption is

$$\texttt{Interval} = 33.47 + 10.73 \times \texttt{Duration}$$



16. Describe the association between *Duration* and *Interval* of the Old Faithful geyser.

    (A) There appears to be a fairly strong positive linear relationship with two distinct clusters of data.

    (B) There appears to be a fairly strong negative linear relationship with two distinct clusters of data.

    (C) There appears to be weak relationship since $r = 0.901$ is smaller than 1.

    (D) There appears to a weak relationship because of the two clusters of data.

    (E) The relationship would be stronger if r^2 were less than 81%.

17. Explain what the slope of the line means in this context.

    (A) Since the correlation is 0.901, time between eruptions increases by about 90%.

    (B) The time between eruptions increases by about 33.47 minutes on the average.

    (C) The duration of a previous eruption lasts on the average of about 10.73 minutes.

    (D) We can predict the time between eruptions will be 10.73 minutes and this can be stated with 0.901 probability.

    (E) If duration increases by one minute, the interval between eruptions increases by 10.73 minutes on average.

18. If you just witnessed an eruption that lasted 2 minutes, how long do you predict that you will have to wait until the next eruption?

    (A) about 88 mins. (B) about 78 mins. (C) about 62 mins. (D) about 55 mins. (E) about 44 mins.

19. The Community Intervention Trial for Smoking Cessation asked whether a community-wide advertising campaign would reduce smoking. The researchers located 11 pairs of communities, each pair similar in location, size, economic status, and so on. One community in each pair participated in the advertising campaign and the other did not. To decide which community in each pair should get the advertising campaign the researchers tossed a coin. This is

    (A) a completely randomized experiment.
    (B) a stratified random sample.
    (C) an observational study.
    (D) an uncontrolled experiment.
    (E) a matched pairs experiment.

20. Make _____ to display the relationship between two quantitative variables measured on the same subjects.

    (A) side-by-side boxplots
    (B) a scatterplot
    (C) back-to-back stemplots
    (D) the five-number summary
    (E) a two sample t-test

21. Does regular exercise reduce the risk of a heart attack? Here are two ways to study this question.

    (1) A researcher finds 2000 men over 40 who exercise regularly and have not had heart attacks. She matches each with a similar man who does not exercise regularly, and she follows both groups for 5 years.

    (2) Another researcher finds 4000 men over 40 who have not had heart attacks and are willing to participate in a study. She assigns 2000 of the men to a regular program of supervised exercise. The other 2000 continue their usual habits. The researcher follows both groups for 5 years.

    Which of the following is correct?

    (A) The first design is an observational study, and the second is an experiment.
    (B) The first design is an experiment, and the second is an observational study.
    (C) The first design is an SRS and the second is a stratified random sample.
    (D) Both designs are experiments.
    (E) Both designs are observational studies.

    Use the following for the next 2 questions. Can changing diet reduce high blood pressure? Vegetarian diets and low-salt diets are both promising. There are 240 men with high blood pressure to serve as subjects and they are to be assigned at random to four diets: (1) normal diet with unrestricted salt; (2) vegetarian with unrestricted salt; (3) normal with restricted salt; and (4) vegetarian with restricted salt. The study ran for 8 weeks.

22. An important response variable in this experiment is

    (A) control, randomization and replication.
    (B) the randomization of the subjects to the treatments.
    (C) the change in blood pressure after 8 weeks on the assigned diet.
    (D) normal/vegetarian diet or unrestricted/restricted salt.
    (E) the length of the study.

23. It turned out that the group of men that were on the vegetarian diet with restricted salt had reduced their high blood pressure the most. The researchers said the results were statistically significant at the 0.05 significance level. What does "statistically significant" mean in this context?

    (A) Perhaps the results are attributable to some confounding variable (e.g., exercise).
    (B) They mean that the change in blood pressure for the men on the vegetarian diet with restricted salt was larger than what they would expect to occur by chance alone.
    (C) There is evidence at the 0.05 significance level that there is a difference in the mean reduction of blood pressure between the diets.
    (D) The results are of practical importance and this can be stated with 95% confidence.
    (E) The chance that the null hypothesis is true is very small.

24. Researchers asked mothers how much soda (in ounces) their kids drank in a typical day. They also asked these mothers to rate how aggressive their kids were on a scale of 1 to 10, with larger values corresponding to a greater degree of aggression. The correlation between amount of soda consumed and aggression rating was found to be $r = 0.3$. If the researchers had measured amount of soda consumed in liters instead of ounces, what would be the correlation? (There are about 34 ounces in a liter.)

    (A) 0.3

    (B) $0.3/34 = 0.009$

    (C) $(0.3)(34) = 10.2$

    (D) 0.3 ounces per point on scale

    (E) Cannot be determined without converting ounces to liters.

    Use the following for the next 2 questions. Here is the probability model for the classification in college of a randomly chosen student from the Spring 2017 survey.

Classification	Freshman	Sophomore	Junior	Senior
Probability	0.63	0.21	0.10	?

25. The probability that a randomly chosen student from the survey is a Senior must be

    (A) 0.60 (B) 0.06 (C) 0.94 (D) unable to be determined (E) any number between 0 and 1

26. What is the probability that a randomly chosen student from the survey is not a Freshman?

    (A) 0.63 (B) 0.31 (C) 0.37 (D) unable to be determined (E) any number between 0 and 1

27. In Pick 3 lottery game, the player chooses 3 numbers from 0 to 9. For the straight bet, this is simply matching all 3 numbers in the correct order. The probability of getting each number correct is $1/10$ so the probability of matching all 3 is $(1/10)^3 = 1/1000$. This means that

    (A) if you play the Pick 3 lottery 1000 times, exactly 1 of them will be a perfect match of all 3 numbers.

    (B) since there are only two possibilities (win or lose) your chances of winning will be very close to $50 - 50$ in the long run.

    (C) if you played the Pick 3 lottery 999 times and lost every time, the next time that you play you should have a perfect match of all 3 numbers.

    (D) if you played the Pick 3 lottery and lost then your chances of winning will increase the next time you play.

    (E) if you play the Pick 3 lottery thousands of times, the fraction of them that would have a perfect match with your 3 numbers will be very close to $1/1000$.

28. An analysis of published medical studies about heart attacks noted that in the studies having randomization and strong control for bias, the new therapy provided improved treatment 9% of the time. In studies without randomization or other controls for bias, the new therapy provided improved treatment 58% of the time. How should we interpret these reported findings?

    (A) The randomized studies were poorly conducted, or they would have found the new treatment to be better much more than 9% of the time.

    (B) These results suggest it is better not to use randomization in medical studies, because it is harder to show that new ideas are beneficial.

    (C) In this case, it is better to use the uncontrolled study results instead of the controlled experiment results.

    (D) Some articles that suggest a particular food, drug, or environmental agent is harmful or beneficial should be viewed skeptically, unless we learn more about the statistical design and analysis for the study.

    (E) The findings of these studies suggest that the answer will most likely lie in the middle.

Use the following for the next 3 questions. Question 2 from the Spring 2017 Math 1530 survey asked students to identify their grade classification and question 12 from the survey asked students "Are you satisfied with America's state of race relations?" The distribution of counts is shown in the table below.

	Relations		
Classification	**Yes**	**No**	**Total**
Freshman	166	323	489
Sophomore	41	122	163
Junior	17	59	76
Senior	8	39	47
Total	232	543	775

29. What percent of the students are Freshman and are not satisfied with race relations in America?
(A) 70% (B) 66% (C) 63% (D) 59% (E) 42%

30. Given the group of students that are not satisfied with race relations in America, what percent are Freshman?
(A) 70% (B) 66% (C) 63% (D) 59% (E) 42%

31. Given the group of students that are Freshman, what percent are not satisfied with race relations in America?
(A) 70% (B) 66% (C) 63% (D) 59% (E) 42%

32. A person is testing whether a coin that a magician uses is biased. After analyzing the results from his coin flipping, the $P-$value ends up being 0.20, so he concludes that there is no evidence that the coin is biased. Based on this information, which of these is a possible 95% confidence interval on the population proportion of times heads comes up?
(A) $(.43, .55)$ (B) $(.32, .46)$ (C) $(.76, .98)$ (D) $(.14, .45)$ (E) Any interval is possible.

Use the following for the next 3 questions. Fox News/Anderson Robbins Research/Shaw & Company conducted a poll between March 12-14 on the president's approval rating. They sampled 1008 registered voters and found 434 registered voters approved of the president.

33. The sample proportion $\hat{p}$ that approve of the president's job is
(A) 1008. (B) 434. (C) 95%. (D) 0.43. (E) 0.57.

34. A 95% confidence interval for the proportion of all registered voters that approve of the president during this time period is
(A) $(0.41, 0.45)$. (B) $(0.40, 0.46)$. (C) $(0.43, 0.44)$. (D) $(0.54, 0.60)$. (E) $(403, 465)$.

35. How many registered voters must be interviewed to estimate the proportion who approve of the president within ± 0.02 with 95% confidence using the large-sample confidence interval? Use 0.5 as the conservative guess for p.
(A) 5000 (B) 4802 (C) 2401 (D) 1008 (E) 25

Use the following for next 2 questions. As a safety recommendation to recreational fishers, the EPA recommended screening value for Mirex is 0.08 ppm. Fish with higher values are not safe for human consumption. Researchers tested three farm-raised salmon for organic contaminants. They found the mean concentration of the carcinogenic insecticide Mirex to be 0.16 parts per million, with standard deviation 0.06 ppm. Does the sample give evidence that farm-raised fish do not meet EPA guidelines?

36. The hypotheses of interest are

(A) $H_0 : \mu = 0.08 \; vs. \; H_a : \mu > 0.08$ (C) $H_0 : \mu = 0.16 \; vs. \; H_a : \mu < 0.16$

(B) $H_0 : \mu = 0.08 \; vs. \; H_a : \mu < 0.08$ (D) $H_0 : \mu = 0.16 \; vs. \; H_a : \mu > 0.16$

(E) $H_0 : \bar{x} = 0.08 \; vs. \; H_a : \bar{x} > 0.08$

37. A one-sample t-test was computed and the $P-$value was found to be 0.07. Although the results _____ _____.

(A) are statistically significant at $\alpha = 0.05$, the explanation is that the sample size is small and it isn't surprising to find the results to be statistically significant.

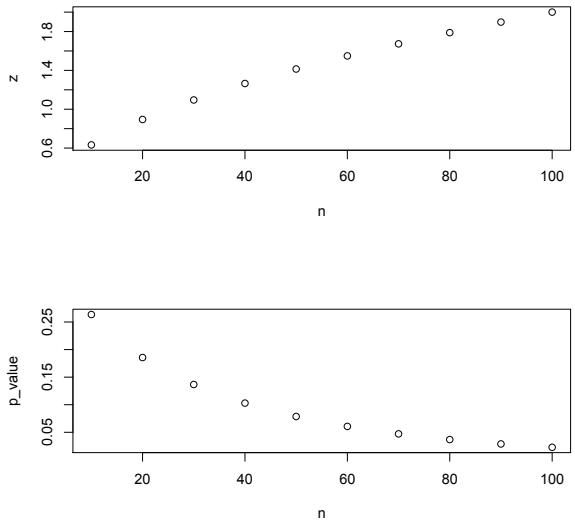
(B) are statistically significant because the mean is equal to 0.16, they are not of practical importance.

(C) are statistically significant at $\alpha = 0.05$, the explanation is that P-value was computed incorrectly.

(D) are not of practical significance they are statistically significant at $\alpha = 0.05$.

(E) are not statistically significant at $\alpha = 0.05$, the explanation is that the sample size is small and it isn't surprising to find the results to be not statistically significant.

38. You are testing the null hypothesis that $\mu = 0$ versus the alternative $\mu > 0$ using $\alpha = 0.05$. Suppose that $\bar{x} = 1$ and $\sigma = 5$. The following plots represent the $z-$test statistic values versus the sample size (n) and the $P-$values versus the sample size (n). Which of the following summarizes the effect of the sample size on significance testing?



(A) The plots show that as sample size increases the results become very important.

(B) The plots show that the null hypothesis is true as sample size increases.

(C) The plots indicate that as sample size increases the test of hypotheses will become statistically significant.

(D) The plots are correlated with each other.

(E) The plots indicate there is a problem with the test of hypotheses since one is increasing and the other is decreasing. Both plots should be increasing since n is getting larger.

39. After surveying students at ETSU, a 95% confidence interval for the mean cost of books for one term was computed to be $(\$380, \$406)$. Which of the following is the correct interpretation of the computed interval?

    (A) 95% of all students pay between \$380 and \$406 for books.

    (B) With 95% confidence, the average amount all students pay is between \$380 and \$406.

    (C) 95% of the sampled students paid between \$380 and \$406.

    (D) We are 95% sure that students in this sample averaged between \$380 and \$406 for books.

    (E) 95% of the samples of students will have average books costs between \$380 and \$406.

40. Inference is _____.

    (A) the process of deducing sample values for populations.

    (B) the process of selecting a sample.

    (C) valid for voluntary response samples.

    (D) the process of drawing conclusions about the population from sample information.

    (E) drawing conclusions about the sample mean and the sample standard deviation.

41. A recent survey of students found that 70% of the students in the sample said that they are not satisfied with the state of race relations in the U.S. The poll's margin of error for 95% confidence was 3%. This means that

    (A) the survey used a method that gets an answer within 3% of the truth about the population 95% of the time.

    (B) we can be sure that the percent of all students who are not satisfied with the state of race relations in the U.S. is between 67% and 73%.

    (C) another survey of students using the same method would produce results that will lie between 67% and 70%.

    (D) there is a 95% chance that 70% will be within $\pm 3\%$.

    (E) the survey wasn't conducted correctly 3% of the time and this can be stated with 95% probability.

Use the following for the next 2 questions. How much sleep do we need? Several sleep studies have found that seven hours is the optimal amount of sleep. A survey asked students "Usually, how many hours sleep do you get a night?" Assume that the 1216 students that responded to this question are an SRS of all students. We suspect that college students sleep less than seven hours. To verify the suspicion, we test the hypotheses

$$H_0 : \mu = 7.0 \; vs. \; H_a : \mu < 7.0$$

We found the following: $\bar{x} = 6.58$ hours, $s = 1.33$ hours, $t = -176$, $p - value \approx 0$. Hence, we reject H_0, and conclude that there is strong evidence that, on average, college students sleep less than seven hours per night.

42. The average number of hours sleep per night was 6.58 hours for this sample of 1216 students. The number 6.58 is a

    (A) parameter. (B) population mean. (C) significant number. (D) statistic. (E) practical number.

43. Which of the following would be most helpful in assessing the practical significance of the results?

    (A) Test the hypotheses again, using a smaller significance level.

    (B) Take a larger sample and retest the hypotheses.

    (C) Take a random sample of non-college students and retest the hypotheses.

    (D) Compute the probability of how many students slept less than 7.0 hours.

    (E) Plot the data to display the effect we are seeking, and use a confidence interval to estimate the actual value of μ.