

Math 1530 Final Exam Fall 2018

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 A

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 Fall 2018 MATH1530 class survey produced variables that are categorical?
- i. Approximately how many ounces of water do you drink a day?
 - ii. Do you think the use of marijuana should be made legal?
 - iii. When a person has a disease that cannot be cured, do you think doctors should be allowed by law to end the patient's life by some painless means if the patient and his or her family request it?
 - iv. What is your favorite color in a rainbow?
 - v. What is the length of your foot (in inches)?
- (A) iv. (B) i. v. (C) ii. iii. iv. (D) i. ii. iii. iv. v. (E) None

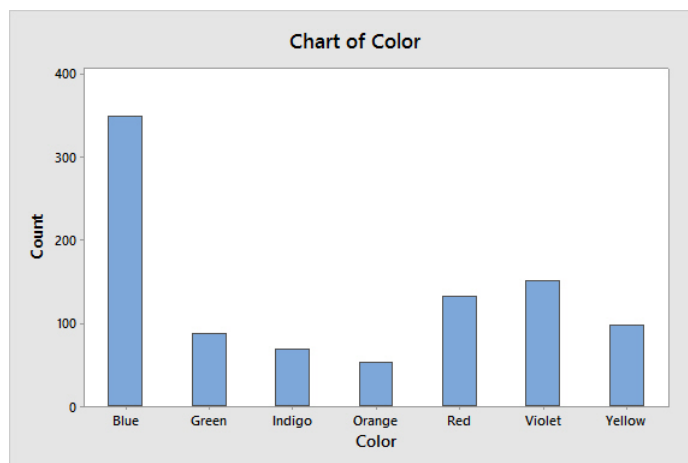
2. A student survey asked “What is your length of your forearm (in inches)?” Here is the data:

9 9 16 10 11 8 12 7 11

What is the median length?

- (A) 5 (B) 9 (C) 10 (D) 10.33 (E) 11
3. If a distribution is skewed to the left,
- (A) the mean and median are equal. (C) the mean is greater than the median.
 - (B) the mean is less than the median. (D) the standard deviation is small.
 - (E) the standard deviation is large.
4. For an exam given to a class, the students' scores ranged from 35 to 98, with a mean of 74. Which of the following is the most realistic value for the standard deviation?
- (A) 12 (B) 1 (C) -10 (D) 63 (E) 74

5. A mean is known as a statistic if it is computed from the
(A) population. (B) parameter. (C) center of the distribution. (D) sample. (E) normal curve.
6. _____ describe the overall pattern of the distribution of a quantitative variable. (Fill-in the blank.)
(A) Pie charts and bar charts (C) The mean and the median
(B) Outliers (D) The range and correlation
(E) Shape, center, and variability
7. You have sample data on two groups shown below and you want to estimate the population mean for each group. For which group will the estimate be more precise?
- | | | | | | | | | | | |
|---------|----|----|----|----|----|----|----|----|----|----|
| Group A | 1 | 6 | 3 | 17 | 21 | 13 | 2 | 4 | 5 | 10 |
| Group B | 57 | 56 | 49 | 52 | 53 | 51 | 55 | 54 | 52 | 53 |
- (A) Group A (C) equally precise for A and B
(B) Group B (D) neither since the numbers vary
(E) insufficient information to tell
8. The Fall 2018 MATH1530 survey asked “What do you think is the most important problem facing this country today?” The responses from 939 students were: Poverty 167; Economy 79; Education 65; Environment 82; Immigration 72; Race relations 149; Dissatisfaction with government 145; Healthcare 96; Other 84. To display this data you should construct a
(A) histogram. (B) stemplot. (C) time plot. (D) bar graph. (E) 5-number summary.
9. The Fall 2018 MATH1530 survey asked “What is your favorite color in a rainbow?” The following graph displays the information on this variable:



Which of the following best describes this distribution?

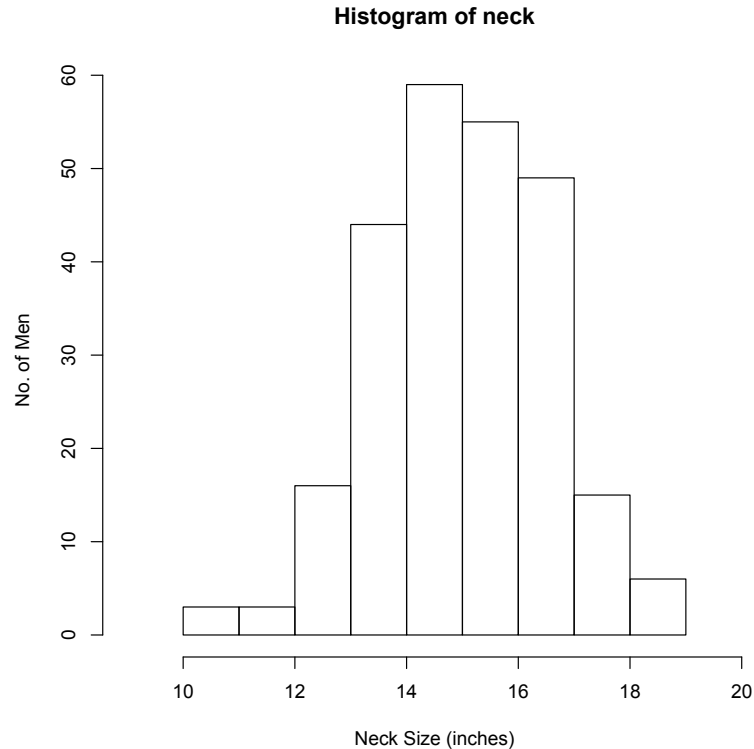
- (A) The distribution is evenly spaced from blue to yellow.
(B) The distribution is multimodal and skewed right.
(C) The distribution is skewed right, the center is orange and the range is from blue to yellow.
(D) The distribution is irregular and centered around orange.
(E) Blue is the favorite color of the rainbow while orange is the least favorite color of the rainbow.

Use the following for the next 4 questions. Here are the ages (sorted) of the 45 presidents of the United States at the start of their presidency.

42 43 46 46 47 47 48 49 49 50 51 51 51 51 51 52 52 54 54 54 54 54 55
55 55 55 56 56 56 57 57 57 57 58 60 61 61 61 62 64 64 65 68 69 70

10. The center of this distribution is
(A) 23 (B) 42 to 70 (C) 51 (D) 55 (E) 59
11. The five-number summary of the 45 ages is
(A) 42, 55, 6.59, 28, 70
(B) 42, 45, 55, 65, 70
(C) 42, 51, 55, 59, 70
(D) 42, 54, 57, 61, 70
(E) 42, 54, 55, 56, 70
12. The shape of this distribution is
(A) strongly skewed right.
(B) uniformly distributed from 42 to 70.
(C) fairly symmetric with possible outliers.
(D) strongly skewed left.
(E) multimodal with many centers and the spread is quite large.
13. Joe is writing a report on the backgrounds of American presidents. He finds that the mean age of the 45 presidents is 55 years with the standard deviation of 6.59 years. Because Joe took a statistics course, he calculates a 95% confidence interval for the mean age of all men who have been president to be 55 ± 1.98 . Comment on this result.
(A) The confidence interval makes no sense since Joe has the entire population of ages.
(B) Joe can be 95% confident that the mean age of all presidents is 55 years.
(C) 95% of all the ages of the presidents lie between 55 ± 1.98 .
(D) Joe is 95% confident that the mean age of all presidents lies between 53.02 and 56.98 years.
(E) Joe should have taken a larger sample to reduce the margin of error.

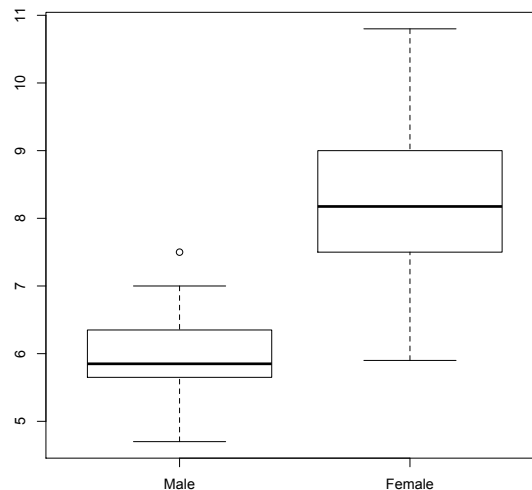
Use the following for the next 2 questions. The histogram shows the neck sizes (in inches) of 250 men recruited for a study.



14. Which summary statistics would you use to summarize the center and spread in these data? Why?
- (A) The 5-number summary since it is multimodal.
 - (B) The mean or median since they would be approximately equal.
 - (C) The 5-number summary since it is strongly skewed.
 - (D) The mean and standard deviation because the distribution is unimodal and symmetric.
 - (E) The mean and the interquartile range (IQR) because the distribution is symmetric with no outliers.
15. Approximately what percent of the men have neck sizes between 14 and 16 inches?
- (A) 15% (B) 46% (C) 55% (D) 60% (E) 100%

Use the following for the next 3 questions. A biologist is interested in comparing the lengths of the male and female green lynx spiders. Here are the summary statistics and boxplots of the lengths of 30 male and 30 female green lynx spiders that were measured in millimeters.

GENDER	n	Mean	Standard Deviation	Minimum	Q1	Median	Q3	Maximum
MALE	30	5.917	0.663	4.700	5.650	5.850	6.312	7.500
FEMALE	30	8.153	1.187	5.900	7.513	8.175	8.938	10.800



16. The distribution of female lengths is

- (A) strongly skewed left.
- (B) skewed left with a median of 8.175 mm and a standard deviation of 1.187 mm.
- (C) strongly skewed right since the boxplot extends up to nearly 11.
- (D) is heavy in the middle and no observations at the tails.
- (E) symmetric with a mean of 8.153 mm and a standard deviation of 1.187 mm.

17. The distribution of male lengths is

- (A) strongly skewed right since the boxplot extends up to 7.5.
- (B) uniformly distributed between 4.7 mm and 7.5 mm.
- (C) is heavy in the middle and has one observation at the tail.
- (D) somewhat symmetric with a mean of 5.917 mm and a standard deviation of 0.663 mm with a potential outlier of 7.5 mm.
- (E) strongly skewed left since the mean is much larger than the median.

18. The lengths of the females tend to be

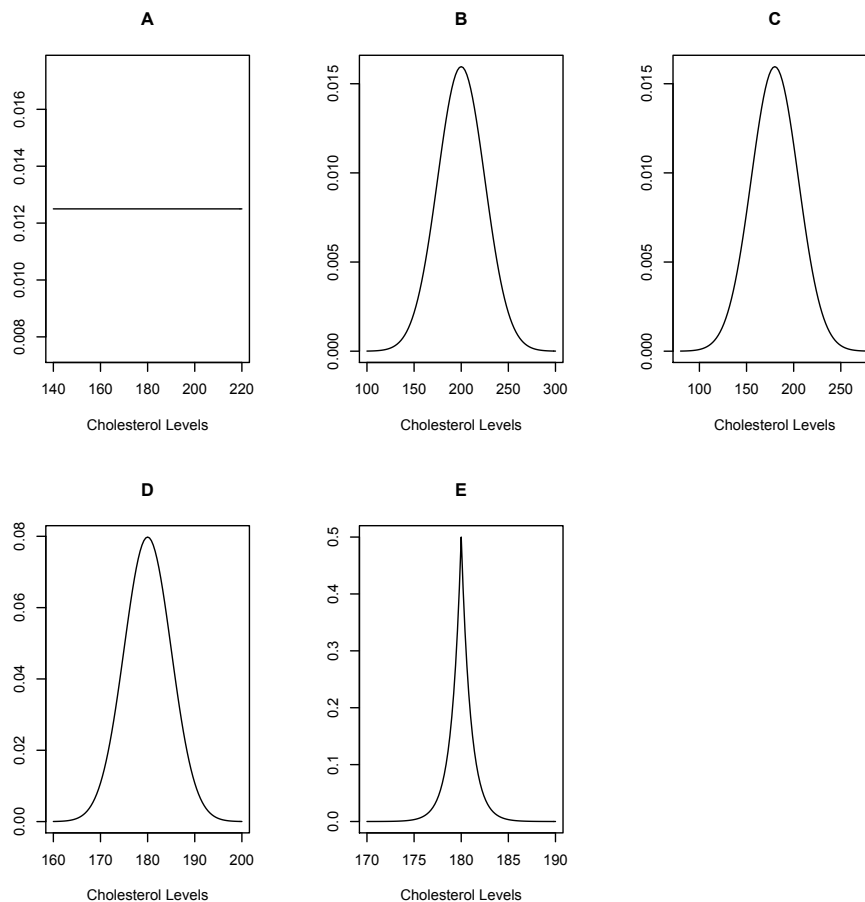
- (A) shorter than the lengths of the males. The lengths of the females vary less than the lengths of the males.
- (B) about the same as the lengths of the males since the summary statistics are nearly the same.
- (C) longer than the lengths of the males since the standard deviation of the female lengths is larger than the standard deviation of the male lengths.
- (D) less varying than the lengths of the males because of the outlier.
- (E) longer than the lengths of the males. The 25th percentile (Q1), which is about 7.513 mm for the females, is longer than nearly all of the male lengths. The lengths of the females vary more than the lengths of the males.

19. A town's January high temperatures average 36°F with a standard deviation of 10°F , while in July the mean high temperature is 74°F and the standard deviation is 8°F . In which month is it more unusual to have a day with a high temperature of 55°F ? Explain.

- (A) In January, a high of 55 is not quite 2 standard deviations above the mean, whereas in July a high of 55 is more than 2 standard deviations lower than the mean. So it's less likely to happen in July.
- (B) In July, a high of 55 is more than 2 standard deviations lower than the mean, whereas in January a high of 55 is not quite 2 standard deviations above the mean. So it's less likely to happen in January.
- (C) Unlikely in January since the standard deviation is larger than the standard deviation is in July.
- (D) It is more likely to have a temperature of 55 in July since it is below the mean.
- (E) It is equally likely to happen in either month since 55 is 19 from 36 and 74.

Use the following for the next 3 questions. Assume that cholesterol levels of adult American women can be described by a Normal model with a mean of 180 mg/dL and a standard deviation of 25.

20. Which of the following graphs best represents the cholesterol levels of adult American women?



21. What percent of adult women do you expect to have cholesterol levels between 130 and 230 mg/dL?

- (A) 50% (B) 68% (C) 95% (D) 99.7% (E) 100%

22. Above what value are the highest 5% of the women's cholesterol levels?

- (A) 131 (B) 139 (C) 212 (D) 221 (E) 229

23. Does prayer reduce anxiety? A study was done to see if people who were being prayed for on a regular basis felt less anxious than those who were not. In this study, the explanatory variable is
- (A) the level of anxiety of the people in the study. (C) the number of people used in the study.
(B) whether a person is being prayed for on a regular basis. (D) the researchers who conducted the study.
(E) the variable on the x-axis.

24. The gas mileage (miles per gallon) and the speed (miles per hour) of an automobile have been measured:

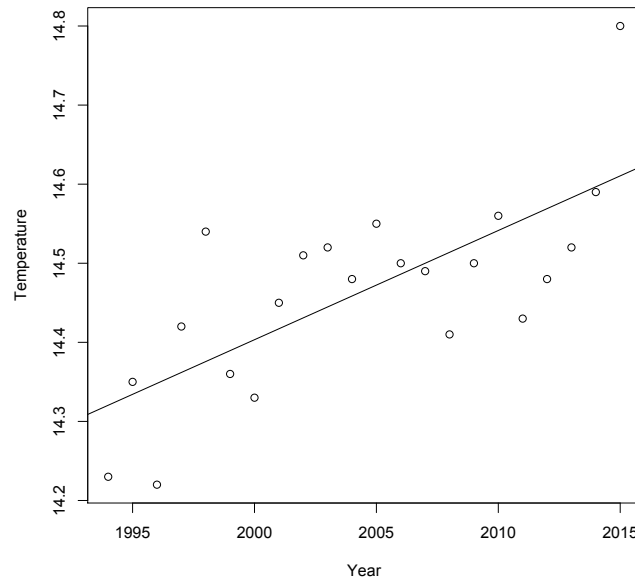
Speed	30	40	50	60	70
Mileage	24	28	30	28	24

Which of the following best describes the relationship between speed and mileage?

- (A) The side-by-side boxplots show that speed is higher than mileage.
(B) The median speed is larger than the median mileage.
(C) There is a strong linear relationship between speed and mileage and the correlation is close to 1.
(D) There is no relationship between speed and mileage.
(E) The correlation is close to 0 but there is a strong curved relationship between speed and mileage.
25. Assuming that we have a linear relationship between two quantitative variables, a large positive correlation shows that
- (A) as one variable increases, the other variable tends to increase.
(B) as one variable increases, the other variable tends to decrease.
(C) as one variable decreases, the other variable tends to increase.
(D) as one variable increases, the other variable tends to stay the same.
(E) there is a weak relationship between the two variables.
26. 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 35 ounces in a liter.)
- (A) 0.3 (B) $0.3/35 = 0.009$ (C) $(0.3)(35) = 10.5$ (D) $35/0.3 = 116.7$ (E) $(0.3)(10)/35 = 0.086$
27. You can summarize the data for two categorical variables x and y by
- (A) drawing a scatterplot of the x - and y -values. (C) calculating the correlation between x and y .
(B) constructing a two-way table for the x - and y -values. (D) constructing a box plot for each variable.
(E) calculating the least-squares regression line.

Use the following for the next 4 questions. Have average global temperatures been increasing in recent years? Data has been collected on the annual average global temperatures for the last 22 years in degrees Celsius. The observations and the least-squares regression appear in the scatterplot. The correlation between the two variables is $r = 0.72$ and the least-squares regression equation is

$$\text{Temperature} = -13.22 + 0.0138 \times \text{Year}.$$



28. Which of the following statements is true regarding the data?
- (A) The association between temperature and year is negative, curved, and weak.
 - (B) The association between temperature and year is positive, linear, and fairly strong.
 - (C) The association between temperature and year is positive and weak.
 - (D) There is no association between temperature and year.
 - (E) The association between temperature and year is nearly perfect.
29. Which is the most appropriate interpretation of the slope?
- (A) For every degree increase, years increased 0.0138.
 - (B) For every year increase, the annual average global temperature decreased by 13.22°C .
 - (C) For every year increase, the annual average global temperature decreased by 0.0138°C .
 - (D) For every year increase, the annual average global temperature increased by $r = 0.72^{\circ}\text{C}$.
 - (E) For every year increase, the annual average global temperature increased by 0.0138°C .
30. According to the least-squares regression equation, the predicted average global temperature for the Year 2000 is
- (A) 13.2°C (B) 14.38°C (C) 14.78°C (D) 28°C (E) 1986.78°C
31. An environmentalist uses the line, based on the 22 years, to predict average global temperatures in 2050. Which of the following is correct?
- (A) The predicted average global temperature in 2050 will be 41.9°C .
 - (B) The predicted average global temperature in 2050 will be 28.7°C .
 - (C) The predicted average global temperature in 2050 will be 2036.8°C .
 - (D) The predicted average global temperature in 2050 will be 15°C .
 - (E) This is extrapolation and we shouldn't use the least-squares regression line for prediction for years beyond 2015.

32. A question posted on a web site asked visitors to the site to say whether they thought marijuana should be legally available. Are there any potential sources of bias and what conclusions can be drawn from this survey?
- (A) Yes. Since this is a voluntary response sample, those who visit the site and respond may be predisposed to a particular answer. It would be difficult to generalize the results to the population.
 - (B) We would need to know the margin of error before we can draw a conclusion from this survey.
 - (C) We should be able to estimate the percentage of the population who think marijuana should be legalized from this survey.
 - (D) No. The results will produce a yes or no response and the binomial distribution can be used to draw conclusions about the population.
 - (E) No. As long as the sample size is large we are able to draw conclusions about the population.
33. Does exposure to aircraft noise increase the risk of hospitalization for cardiovascular disease in older people (≥ 65 years) residing near airports? Selecting a random sample of approximately 650,000 Medicare claims, it was found that about 75,000 people had zip codes near airports and the remaining 575,000 did not. The proportions of hospital admissions related to no cardiovascular diseases were computed for those with zip codes near airports. A large proportion of admissions for cardiovascular disease was found for older people living in zip codes near airports. Which of the following statements is correct?
- (A) Because of the large sample sizes from each group, we can claim that living in a zip code near an airport is causing the increase in the proportion of admissions for cardiovascular disease.
 - (B) Since this is an experiment, but not a randomized experiment, we can still conclude that living in a zip code near an airport is causing the increase in the proportions of admissions for cardiovascular disease.
 - (C) Since this is an observational study, living in a zip code near an airport may or may not be causing the increase in the proportions of admissions for cardiovascular disease.
 - (D) The results are statistically significant and older people should not live near airports.
 - (E) The results are of practical importance and this can be stated with 95% confidence.

Use the following for the next 2 questions. Each month the U.S. Census Bureau's American Community Survey mails survey forms to 300,000 households asking questions about demographic, social, economic, and housing characteristics such as mortgage and utility costs. Telephone calls are made to households that don't return the form. In one month, responses were obtained from 295,000 of the households contacted.

34. The sample is
- (A) all U.S. households.
 - (B) all households with mortgages.
 - (C) the 5,000 households that did not respond.
 - (D) the 295,000 households that responded.
 - (E) the 300,000 households initially contacted.
35. The population of interest is
- (A) all U.S. households.
 - (B) all households with mortgages.
 - (C) only U.S. households with phones.
 - (D) the 295,000 households that responded.
 - (E) the 300,000 households contacted.

36. A large representative random sample of 6906 U.S. adults collected over 20 years showed that “parents reported higher levels of life satisfaction than non-parents,” with the observed difference in life satisfaction between the two groups being statistically significant. We can conclude that
- (A) having children leads to higher levels of life satisfaction. We can reach this conclusion because we have a representative sample.
 - (B) having children leads to higher levels of life satisfaction. We can reach this conclusion because we have both a large and a representative sample.
 - (C) parents should have many children to have even higher satisfaction in their lives.
 - (D) there is a positive association between having children and life satisfaction.
 - (E) parents tend to have higher satisfaction in their lives than non-parents.
37. Researchers investigated the effectiveness of oral zinc, as compared to a placebo, in reducing the duration of the common cold when taken within 24 hours of the onset of symptoms. The researchers found those taking oral zinc had a statistically significantly shorter duration ($P - value < 0.05$) than those taking the placebo. This means that
- (A) the probability that the null hypothesis is true is less than 0.05.
 - (B) the results are of practical significance.
 - (C) there is evidence to suggest that taking oral zinc can shorten the duration of the common cold when taken within 24 hours of the onset of symptoms.
 - (D) we do not reject the null hypothesis at significance level $\alpha = 0.05$.
 - (E) there is a lack of evidence to suggest that taking oral zinc can shorten the duration of the common cold when taken within 24 hours of the onset of symptoms.
38. You read in a book on poker (card game) that the probability of being dealt a straight flush in a five-card poker hand is $1/64,974$. This means that
- (A) if you deal 64,974 poker hands, exactly one of them will contain a straight flush.
 - (B) if you deal 6,497,400 poker hands, exactly 100 of them will contain a straight flush.
 - (C) if you deal 64,973 poker hands and none of them contain a straight flush, the next hand dealt will be a straight flush.
 - (D) if you deal millions of poker hands, the fraction of them that contain a straight flush will be very close to $1/64,974$.
 - (E) the probability that you get a straight flush is $5/64,974$.
39. The Fall 2018 MATH1530 class survey asked “When a person has a disease that cannot be cured, do you think doctors should be allowed by law to end the patient’s life by some painless means if the patient and his or her family request it?” In the sample survey 749 students responded “Yes” and 188 students responded “No.” Assume that this sample represents all college students. A large-sample 95% confidence interval for estimating the proportion of all college students who think doctors should be allowed by law to end the patient’s life by some painless means if the patient’s disease cannot be cured and his or her family request it is
- (A) 22% to 28% (B) 77.4% to 82.5% (C) 79.9% to 80.0% (D) 80% (E) 749 ± 188

Use the following for the next 3 questions. Two questions asked on a survey were “What gender do you identify with?” and “Do you personally believe that, in general, pornography is morally acceptable or morally wrong?” Here is a two-way table of the results:

	Pornography		
Gender	Morally Acceptable	Morally Wrong	Total
Female	203	316	519
Male	200	198	398
Other	10	1	11
Total	413	515	928

40. The percent of students who think pornography is morally wrong is
 (A) 38.4% (B) 44.6% (C) 49.7% (D) 55.5% (E) 61.4%
41. What percent of students who think pornography is morally wrong are males?
 (A) 21.3% (B) 38.4% (C) 49.7% (D) 55.5% (E) 55.9%
42. What percent of male students think pornography is morally wrong?
 (A) 21.3% (B) 38.4% (C) 49.7% (D) 55.5% (E) 55.9%
43. The 49 students in a class at the university made blinded evaluations of pairs of cola drinks. For the 49 comparisons of Coke and Pepsi, Coke was preferred 29 times. The table shows results of using MINITAB to do large-sample analysis.

Test of $p = 0.50$ vs. $\text{not} = 0.50$

X	N	Sample p	95.0% CI	Z-Value	P-Value
29	49	0.5918	(0.454, 0.729)	1.286	0.1985

In the population that this sample represents, what does the 95% confidence interval tell you that the test does not?

- (A) The 95% confidence interval tells us the range of plausible values, whereas the test merely tells us that 0.50 is plausible.
 (B) The probability that 0.50 is in between 0.454 and 0.729 is .95, whereas the probability that 0.50 is true is only 0.1985.
 (C) The test tells us how important the results are and interval gives a range of possible values.
 (D) The 95% confidence interval is computed from the sample proportion but the P-Value has nothing to do with the sample proportion.
 (E) They are equal in giving us important information about this study.