

Math 1530 Final Exam Fall 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 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 2017 MATH1530 class survey produced variables that are categorical?
- i. Do you prefer to read textbooks in print or on an electronic device (such as a smart phone, tablet, laptop, computer, e-reader)?
 - ii. Approximately how much did you spend on textbooks this semester?
 - iii. What is your age?
 - iv. How many U.S. states have you visited, not including the state you were born?
 - v. Which comes closest to your opinion about Confederate monuments in public spaces: “They should be removed.”, “They should remain.”, or “I do not know.”?
- (A) i (B) ii (C) v (D) i, v (E) ii, iii, iv

Use the following for the next 2 questions. Where do students go to school? Although 81.4 % of first time first-year students attended college in the state in which they lived, this percent varied considerably over the states. Here is the data (sorted) of the percent of first-year students in each of the 50 states who were from the state where they enrolled.

36 44 55 55 59 60 62 63 63 68 68 69 70
73 73 74 75 76 76 76 77 79 79 80 80 80
80 81 81 81 81 81 82 82 83 83 83 83 83
84 84 84 85 85 85 86 91 91 92 92 92

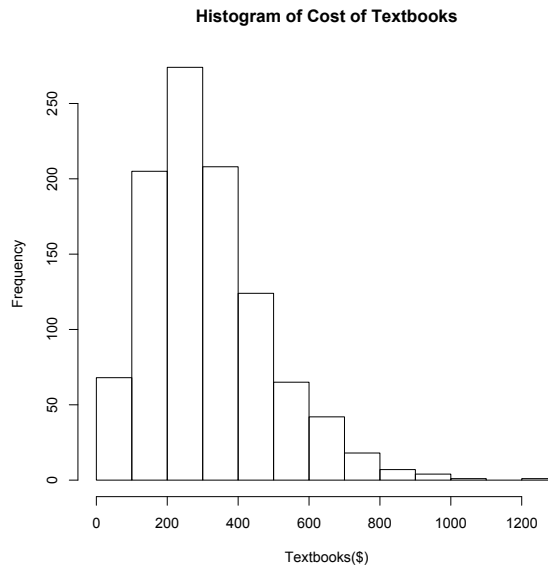
2. The shape of the distribution is
- (A) roughly symmetric.
 - (B) clearly skewed to the right with two low outliers.
 - (C) clearly bell-shaped with possible outliers.
 - (D) clearly multimodal, skewed right with outliers.
 - (E) clearly skewed to the left with two low outliers.
3. The center of the distribution is close to
- (A) 65% (B) 70% (C) 80% (D) 83% (E) 84%
4. A student survey asked “On average, how many hours of sleep did you get per night over the past 7 days?” Here is the data:

8 6.5 9 6.5 9 7.5 6 7 7.5

What is the mean number of sleep hours?

- (A) 7.00 (B) 7.44 (C) 7.50 (D) 9.57 (E) 67

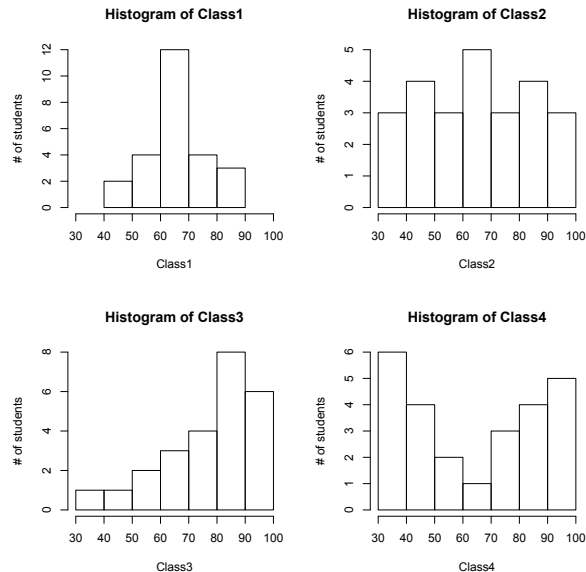
5. The Fall 2017 MATH1530 survey asked “Approximately how much did you spend on textbooks this semester?” The figure below represents the responses of 1017 students.



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

- (A) left-skewed with outliers.
 - (B) right-skewed, with the center around \$300 and spread from \$0 to \$1300, with possible outlier(s).
 - (C) symmetric around \$300, quite a bit of variation, and many outliers.
 - (D) evenly-spaced from \$0 to \$1300.
 - (E) evenly-spaced from \$0 to \$300.
6. The Educational Testing Service reported that the mean Graduate Record Exam (GRE) for all individuals who have taken the exam was 529. A student with a GRE score of 600 wants to know her relative standing in relation to the mean GRE. A numerical summary that would be useful for this purpose is the
- (A) correlation coefficient.
 - (B) median.
 - (C) interquartile range.
 - (D) the number of individuals that took the exam.
 - (E) standard deviation.
7. The Fall 2017 MATH1530 survey asked “Which comes closest to your opinion about Confederate monuments in public spaces?” The responses from 1017 students were: They should remain, 563; I do not know, 254; They should be removed, 200. To display these data you should construct a
- (A) histogram.
 - (B) stemplot.
 - (C) time plot.
 - (D) bar graph.
 - (E) 5-number summary.

Use the following for the next 4 questions. Four Statistics classes all took the same test. Histograms of the scores for each class are shown below.



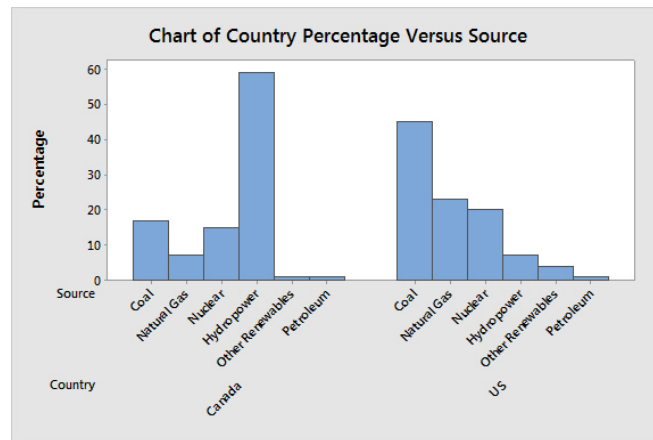
8. Which class had the highest mean score?
(A) Class 1 (B) Class 2 (C) Class 3 (D) Class 4 (E) Unable to be determined.
9. Which class had the highest median score?
(A) Class 1 (B) Class 2 (C) Class 3 (D) Class 4 (E) Unable to be determined.
10. Which class had the smallest standard deviation?
(A) Class 1 (B) Class 2 (C) Class 3 (D) Class 4 (E) Unable to be determined.
11. Overall, which class do you think performed better on the test?
(A) Class 1 (B) Class 2 (C) Class 3 (D) Class 4 (E) Unable to be determined.
12. The data below represents the sodium values for 20 breakfast cereals.

0	50	70	100	130	140	140	150	160	170
180	180	190	200	200	210	210	220	290	340

The five-number summary of the data is

- | | |
|------------------------------|------------------------------|
| (A) 0, 135, 175, 205, 340 | (C) 0, 166.5, 175, 290, 340 |
| (B) 0, 70, 180, 210, 340 | (D) 1, 5.75, 10.5, 15.25, 20 |
| (E) 0, 87.5, 175, 257.5, 340 | |

13. The figure below represents a percentage breakdown by source of all electricity generated in the United States and Canada in 2009.



Which of the following best describes these distributions?

- (A) The distribution for Canada is irregular and the distribution for the US is skewed right.
- (B) The main source of electricity for the US is coal and hydropower is the main source of electricity for Canada.
- (C) The center for both distributions is between Nuclear and Hydropower.
- (D) US: Skewed right, Center = Nuclear, Spread = Coal to Petroleum. Canada: Bimodal, Center = Nuclear, Spread = Coal to Petroleum.
- (E) The distribution is bimodal. There is a gap between Petroleum (Canada) and Coal (US).

Use the following for the next 2 questions. You are asked to do a study of shelters for abused and battered women to determine the necessary capacity in your city to provide housing for most of these women. After recording data for a whole year, you find that the mean number of women in shelters each night is 250, with a standard deviation of 75. The distribution of the number of women in shelters each night follows the normal distribution.

14. The middle 95% of the number of women in shelters each night is between
 (A) 175 and 325. (B) 25 and 475. (C) 250. (D) 100 and 400. (E) 250 ± 2 .
15. If the city's shelters have a capacity of 350, what percentage of nights will there NOT be enough places for abused women?
 (A) 0.1% (B) 2.5% (C) 9% (D) 16% (E) 90%
16. If the correlation between two variables is close to 0, you can conclude that a scatterplot would show
 (A) nothing.
 (B) outliers.
 (C) a strong straight-line pattern.
 (D) all points lying on a straight line.
 (E) no straight-line pattern, but there might be a strong pattern of another form.

17. Match each of the five scatterplots to the descriptions of its regression line and correlation coefficient. The scales on the axes of the scatterplots are the same.

$$(I) \quad r = .83, \quad \hat{y} = -5.20 + 1.47x$$

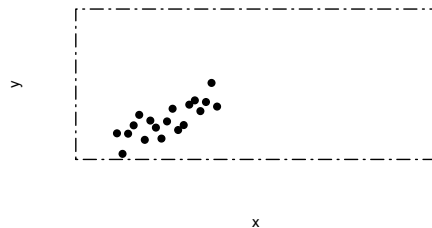
$$(II) \quad r = -.26, \quad \hat{y} = 30.30 - 0.34x$$

$$(III) \quad r = .95, \quad \hat{y} = -5.20 + 1.47x$$

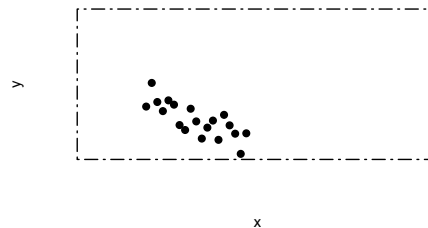
$$(IV) \quad r = -.83, \quad \hat{y} = 50.47 - 1.47x$$

$$(V) \quad r = .59, \quad \hat{y} = -3.09 + 1.47x$$

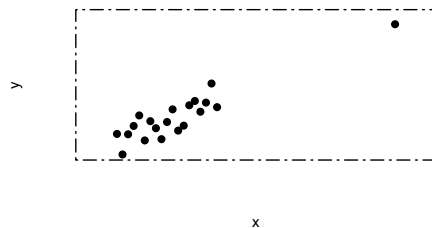
Plot 1



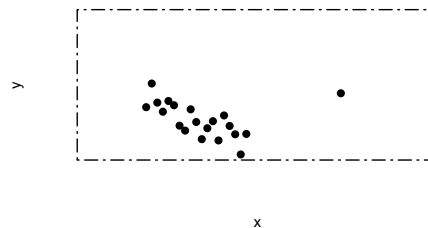
Plot 4



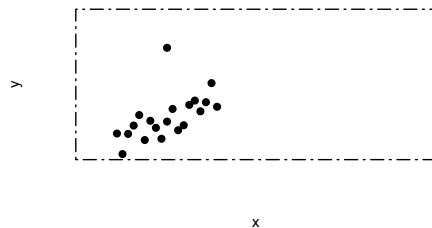
Plot 2



Plot 5



Plot 3



- (A) I = Plot 1, II = Plot 5, III = Plot 2, IV = Plot 4, V = Plot 3 (C) I = Plot 1, II = Plot 4, III = Plot 2, IV = Plot 5, V = Plot 3
 (B) I = Plot 2, II = Plot 4, III = Plot 1, IV = Plot 5, V = Plot 3 (D) I = Plot 3, II = Plot 4, III = Plot 2, IV = Plot 5, V = Plot 1
 (E) Unable to be determined.

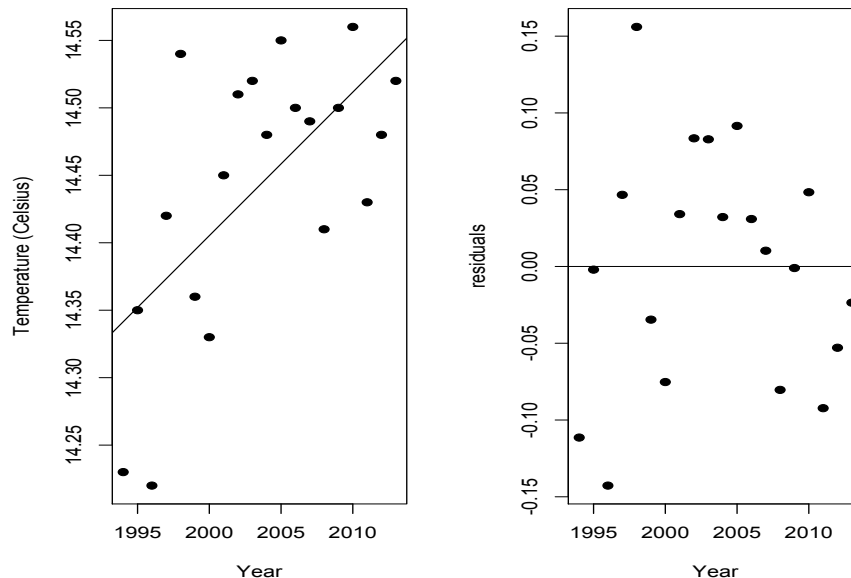
Use the following for the next 2 questions. An owner of a home in the Midwest installed solar panels to reduce heating costs. After installing the solar panels, he measured the amount of natural gas used y (in cubic feet) to heat the home and outside temperature x (in degree-days, where a day's degree-days are the number of degrees its average temperature falls below $65^\circ F$) over a 23-month period. He then computed the least-squares regression line for predicting y from x and found it to be

$$\hat{y} = 85 + 16x$$

18. How much, on average, does gas used increase for each additional degree-day?
 (A) 16 cubic feet (B) 23 cubic feet (C) 85 cubic feet (D) 65° (E) 101 cubic feet
19. According to the regression line, the predicted amount of gas used when the outside temperature is 20 degree-days is about
 (A) 85 cubic feet. (B) 105 cubic feet. (C) 320 cubic feet. (D) 405 cubic feet. (E) 1125 cubic feet.

Use the following for the next 2 questions. Have average global temperatures been increasing in recent years? The observations, least-squares regression line, and residual plot appear in the figure below. The correlation between the two variables is $r = 0.633$ and the least-squares regression line for predicting the annual average global *Temperature(Celsius)* over *Time(Years)* is

$$\text{Temperature} = -6.87293 + 0.01064 \times \text{Year}$$



20. If Temperature had been measured in Fahrenheit ($^{\circ}\text{F}$) instead of Celsius ($^{\circ}\text{C}$), what would be the correlation?
 ($^{\circ}\text{F} = ^{\circ}\text{C} \times 9/5 + 32$)
- (A) $0.633(9/5) + 32 = 4.139$
 (B) $0.633(5/9) - 32 = -31.648$
 (C) $0.633^2 = 0.4009$
 (D) 0.633
 (E) Cannot be determined without converting the temperatures from Celsius to Fahrenheit.
21. Describe the association between Year and Temperature.
- (A) There is no association between the two variables.
 (B) There appears to be a very strong linear relationship between the two variables.
 (C) The relationship between the two variables would be stronger if $R^2 = 0.4009$ was less than 40%.
 (D) The plots suggest that temperatures have been decreasing overall, but there seems to have been a slowing in the past few years; this graph looks curved.
 (E) The plots suggest that temperatures have been increasing overall, but there seems to have been a slowing in the past few years; this graph looks curved. Correlation may not be a useful measure here.
22. A simple random sample of size n is one in which
- (A) each possible sample of size n has the same chance of being selected.
 (B) every n th member is selected from the population.
 (C) there is exactly the same proportion of women in the sample as is in the population.
 (D) you keep sampling until you have a fixed number of people having various characteristics (e.g., males, females).
 (E) the sample size is large.

23. 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 cardiovascular disease were computed for those with zip codes near airports and those who did not have zip codes near airports. A larger 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 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.
 - (C) 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.
 - (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.
24. Many studies have found an association between red meat consumption and an increased risk of chronic diseases. What is the relationship between red meat consumption and mortality? A large study followed 120,000 men and women who were free of coronary heart disease and cancer at the beginning of the study. Participants were asked detailed questions about their eating habits every four years, and the study spanned almost 30 years. It was found that the risk of dying at an early age from heart disease, cancer, or any other cause, rises with the amount of red meat that they consumed. This is
- (A) a completely randomized experiment.
 - (B) a matched pairs experiment.
 - (C) a simple random sample survey.
 - (D) a double-blind experiment.
 - (E) an observational study.
25. The General Social Survey asked 827 people how many days they would wait to seek medical treatment if they were suffering pain that interfered with their ability to work. The results are presented in the following table.

Number of days	0	1	2	3	4	5
Frequency	27	436	263	72	19	10

Consider these 827 people to be a population. Find the probability that a randomly selected person would wait for 3 days.

- (A) 0.004 (B) 0.042 (C) 0.087 (D) 0.6 (E) 72
26. A bag contains 12 red marbles, 5 yellow marbles, and 15 green marbles. How many additional red marbles must be added to the 32 marbles already in the bag so that the probability of randomly drawing a red marble is $\frac{3}{5}$?
- (A) 13 (B) 18 (C) 28 (D) 32 (E) 40
27. You flip a fair coin four times and get heads every time. What are the chances of flipping a tail on the fifth time?
- (A) 0 (B) $\frac{1}{5}$ (C) $\frac{1}{2}$ (D) $\frac{4}{5}$ (E) 1
28. An airport screens bags for forbidden items, and an alarm is supposed to be triggered when a forbidden item is detected. Suppose that 5% of bags contain forbidden items. If a bag contains a forbidden item, there is a 98% chance that it triggers the alarm. If a bag doesn't contain a forbidden item, there is an 8% chance that it triggers the alarm. What is the probability that the alarm is triggered? (Hint: Draw a tree diagram.)
- (A) $(0.05)(0.98) + (0.95)(0.08)$ (B) $0.98 + 0.08$ (C) 0.05 (D) 0.98 (E) $0.98 + 0.05$

29. You can summarize the data for two categorical variables x and y by
- (A) drawing a scatterplot of the x - and y -values.
 - (B) calculating the correlation between x and y .
 - (C) constructing a box plot for each variable.
 - (D) constructing back-to-back stem plots.
 - (E) constructing a two-way table of counts for the x - and y -values.

Use the following for the next 3 questions. The Fall 2017 survey asked 1017 students “What is your current smoking status?” and “Have you ever smoked marijuana?” The distribution of counts is shown in the table.

	Marijuana		
Smoking Status	Yes	No	Total
Non-smoker	346	510	856
Former	51	9	60
Current	86	15	101
Total	483	534	1017

30. What percent of students have smoked marijuana?
- (A) 8% (B) 18% (C) 47% (D) 84% (E) 85%
31. Given the group of students that are current smokers, what percent have smoked marijuana?
- (A) 8% (B) 18% (C) 47% (D) 84% (E) 85%
32. Given the group of students that have smoked marijuana, what percent are current smokers?
- (A) 8% (B) 18% (C) 47% (D) 84% (E) 85%

Use the following for the next 2 questions. An opinion poll asks an SRS of 1000 college seniors how they view their job prospects. In all, 530 say “Good.”

33. The large-sample 95% confidence interval for estimating the proportion of all college seniors who think their job prospects are good is
- (A) 0.53 ± 0.02 (B) 0.53 ± 0.0005 (C) 0.53 ± 0.45 (D) 0.53 ± 0.03 (E) 0.53 ± 0.95
34. Suppose only a random sample of 100 college seniors rather than 1000 was taken. Which of the following statements is true?
- (A) The margin of error for our 95% confidence interval would decrease.
 - (B) The margin of error for our 95% confidence interval would stay the same, since the level of confidence has not changed.
 - (C) The standard error would decrease.
 - (D) The margin of error for our 95% confidence interval would increase.
 - (E) The margin of error for our 95% confidence interval could decrease or increase depending on the value of the sample proportion of students that responded with “Good”.
35. An observed effect so large (or so different from our assumptions) that it could rarely occur by chance, is called what?
- (A) Practical significance
 - (B) Least-squares linear regression
 - (C) Statistically significant
 - (D) A Level C confidence interval
 - (E) A completely randomized experiment

36. You turn your web browser to the online Harris Interactive Poll. Based on 2234 responses, the poll reports that 45% of U.S. adults believe that global climate change exists and humans are the main cause, 30% believe global climate change exists but that its causes are mainly not related to humans, 13% do not believe global climate change exists and 12% are undecided. A researcher computed a 95% confidence interval for the proportion of all U.S. adults who do not believe in global climate change to be from 12% to 14%. Interpret this result.
- (A) Anywhere from 12% to 14% of all U.S. adults do not believe in global climate change.
 - (B) The computed interval can't be trusted since it was computed from a voluntary response sample.
 - (C) We are 95% confident that 13% of all U.S. adults do not believe in global climate change give or take 1%.
 - (D) Given the data, there is a 95% probability that the true proportion of U.S. adults that do not believe in global climate change is between 12% and 14%.
 - (E) We are 95% confident that the sample of 2234 U.S. adults that do not believe in global climate change is between 12% and 14%.
37. A test of significance is based on computing a number that compares the value of the parameter stated by the null hypothesis with an estimate of the parameter from the sample data. In general, what is this number called?
- (A) The null hypothesis (C) The P-value
 - (B) The critical value (D) The sample mean \bar{x}
 - (E) The test statistic
38. A probability that is computed, for a statistical test, by using a probability distribution that models the chance, if H_0 is true, that a test statistic would take a value at least as extreme as that actually observed from sample data is called what?
- (A) The significance level (α) (C) The population mean (μ)
 - (B) The P-value (D) The probability of "success" on each trial.
 - (E) A two-sided test
39. There has been recent evidence to suggest that American adults weigh more, on average, than they did a quarter of a century ago. A biostatistician at the Quillen College of Medicine is curious whether the mean weight (μ) of all college students has increased, as well. Suppose that the mean weight of college students was 150 pounds in 1980. The researcher plans to examine the weights of a random sample of college students in 2017 to see if the mean weight of all college students has increased significantly in 37 years. What is the appropriate null hypothesis for the test?
- (A) $H_0 : \mu = 150$ (B) $H_0 : \mu \pm 150$ (C) $H_0 : \mu > 150$ (D) $H_0 : \mu \neq 150$ (E) $H_0 : \mu < 150$
40. In a large population of college-educated adults, the mean IQ score is 112 with standard deviation of 25. Suppose 300 adults from this population are randomly selected for a market research campaign. The distribution of the sample mean IQ is
- (A) approximately Normal, mean 112, standard deviation 25.
 - (B) approximately Normal, mean 112, standard deviation 1.443.
 - (C) approximately Normal, mean 112, standard deviation 0.083.
 - (D) approximately Normal, mean equal to the observed value of the sample mean, standard deviation 25.
 - (E) Cannot determine based on the information provided.

41. Researchers would like to estimate the mean cholesterol level μ of a particular variety of monkey that is often used in laboratory experiments. They would like their estimate to be within 0.5 milligrams per deciliter (mg/dl) of the true value of μ at a 90% confidence level. A previous study involving this variety of monkey suggests that the standard deviation of cholesterol level is about 4.68 mg/dl. What sample size is needed to ensure the margin of error is no more than 0.5 mg/dl at 90% confidence?
- (A) 15 (B) 16 (C) 237 (D) 238 (E) 337

Use the following for the next 2 questions. A wedding planner read the article in Business Insider about using a mathematical theory to determine the best age to get married (<http://www.businessinsider.com/best-age-to-get-married-is-26-2017-2>). This story reported the best age to get married was 26. While this number is determined by mathematical theory, the wedding planner feels that males and females may have different views on the best age to get married. She feels that females will think the best age to get married is less than what males consider the best age to get married. Question 10 from the survey asked students "At what age do you think is the best to get married in order to have a successful marriage?" Assume that the students who responded the survey are a SRS of all ETSU students. Is there good evidence to support the idea that female students at ETSU think the best age to get married is less, on average, than male students?

GENDER	N	Mean	StDev
Female	559	25.68	4.20
Male	449	26.90	6.18

42. To perform the hypothesis test, which analysis should we conduct?
- (A) One sample t test. (C) Two sample t test.
- (B) A matched pairs t test. (D) The z -test.
- (E) A Chi-square test.
43. The appropriate analysis was performed and the P-value was 0.0001. Which of the following correctly answers the research question? Use $\alpha = 0.05$.
- (A) The P-value is smaller than α so we will reject the null hypothesis. There is strong evidence that, on average, females think the best age to get married is less than male students at ETSU.
- (B) The P-value is so large that we will not reject the null hypothesis. There is not strong enough evidence to indicate that, on average, females think the best age to get married is less than male students at ETSU.
- (C) The P-value is smaller than α so we will reject the null hypothesis. There is not strong enough evidence to indicate that, on average, females think the best age to get married is less than male students at ETSU.
- (D) The P-value is smaller than α so we will not reject the null hypothesis. There is not strong enough evidence to indicate that, on average, females think the best age to get married is less than male students at ETSU.
- (E) Cannot draw a valid conclusion since the number of male students is not equal to the number of female students.