Math 1530 Final Exam
May 2011

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 may use a calculator but a calculator manual cannot be used.

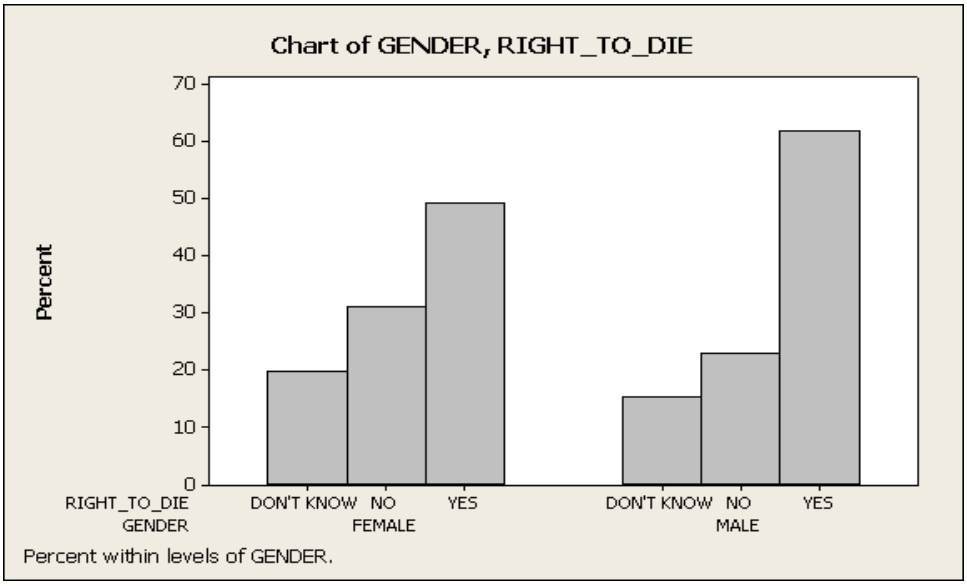
FORM A

1. Which of these questions from the class survey produced variables that are categorical?

   i. What is your sex?
   ii. Do you believe in heaven?
   iii. If you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole?
   iv. What was the cost (in dollars) of your last haircut, including tip?
   v. During a typical week, how many cans (or bottles) of beer do you consume?
   vi. What's the fastest you've ever driven a car (mph)?

   (A) None of these questions. (B) iv, v, vi (C) i, ii, iii (D) i only (E) All of these questions.

   Use the following for the next 2 questions. The figure below represents the class survey question on the right to die by gender. The right to die question was "A person has a disease that cannot be cured, do you think doctors should be allowed by law to end the patients life by some painless means if the patient and his/her family request it?"



2. Approximately what percent of the females responded "Yes" on right to die?
   (A) About 60% (B) About 30% (C) About 50% (D) Nearly 100% (E) Unable to determine from the figure.

3. Which of the following is a true statement?

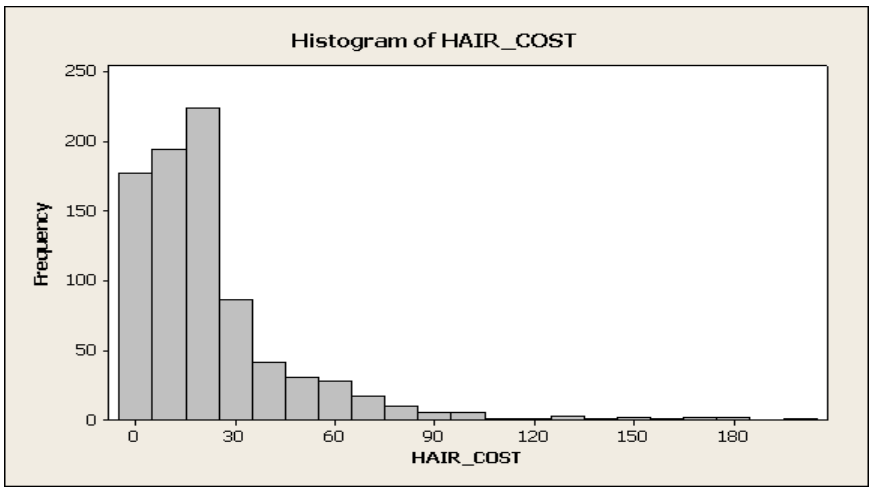
   (A) The bar graphs are bimodal.
   (B) There is a gap between the bar graphs which means we may have outliers.
   (C) It appears that both graphs are skewed left.
   (D) Overall, it appears that a larger percent of males believe that doctors should be allowed to end the patient's life than females do.
   (E) Side-by-side boxplots would be a better display to compare the responses of the females and the males.

4. The table below represents the student responses from the class survey to the question "If you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole?" Which type of graph is appropriate for these data?

General Happiness	Counts
Very Happy	299
Fairly Happy	461
Not Very Happy	53
Not At All Happy	12
Don't Know	10

(A) histogram (B) stem plot (C) scatterplot (D) bar chart (E) boxplot

Use the following for the next 2 questions. The histogram and descriptive statistics below summarize the student responses from the class survey to the question "What was the cost (in dollars) of your last haircut, including tip?"



Descriptive Statistics: HAIR_COST

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
HAIR_COST	832	6	21.584	0.886	25.561	0.000	7.000	15.000	25.000	200.00

---

5. Which of the following best describes the shape of the distribution?

(A) The data are strongly skewed right with many outliers.

(B) 21.58 $\pm$25.56 dollars

(C) The data seems to be somewhat symmetric.

(D) There is an outlier and it is 200.

(E) Who: 832 students, What: cost of last haircut, including tip

6. About what percent of the students paid at most \$7.00 for a haircut?

(A) 75% (B) 50% (C) 7% (D) 25% (E) $z = \frac{7-21.584}{25.561} = -.57$ (Area under Normal Curve = 28%)

Use the following for the next 3 questions. We all know that fruit is good for us. The data below represents the number of servings of fruit per day claimed by 74 seventeen-year-old girls in a study in Pennsylvania.

Servings of fruit per day	0	1	2	3	4	5	6	7	8
Number of subjects	15	11	15	11	8	5	3	3	3

7. Describe the shape of this distribution.

   (A) The distribution is skewed to the right.
   (C) The distribution is symmetric about 4.
   (B) The distribution is skewed to the left.
   (D) The scatterplot shows a negative relationship.
   (E) The distribution is evenly spaced from 0 to 8.

8. What is the median number of servings of fruit per day eaten by these 74 girls?

   (A) 2 (B) 3 (C) 11 (D) 15 (E) 4

9. What percent of these girls ate fewer than two servings per day?

   (A) 55% (B) 65% (C) 35% (D) 25% (E) 45%

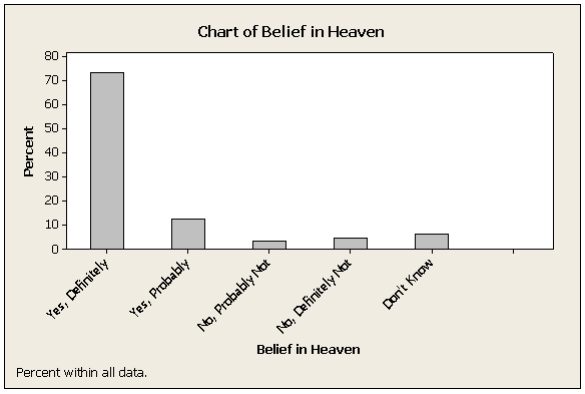
10. 100 students took a test and invested a great amount of time in solving the problems, except for a few students that solved the problems incredibly fast. If we would make a histogram for the amount of time (in minutes) that each student invested in solving the problems, which option would best describe the shape of the histogram?

    (A) skewed to the left
    (C) skewed to the right
    (B) symmetric
    (D) all the bars will be of the same height
    (E) bell-shaped

11. In a study of exercise, a large group of male runners walk on a treadmill for 6 minutes. Their heart rates in beats per minute at the end vary from runner to runner according to the Normal distribution with mean $\mu = 104$ and standard deviation $\sigma = 12.5$. About what percent of the runners have heart rates above 130?

    (A) 98.12% (B) .26% (C) 99.74% (D) 2.5% (E) 1.88%

12. From the class survey students were asked to choose one of the listed responses to the question "Do you believe in heaven?" The bar graph below displays the distribution of heaven belief.



Which of the following is an accurate description of the graph?

(A) The bar graph is skewed to the right.
(B) It appears that a majority of students responding to the survey have a strong belief in heaven.
(C) The bar graph is skewed to the left.
(D) The center of the graph is approximately "No, Probably Not."
(E) Exactly 73 students responded "Yes, Definitely" in the belief of heaven.

Use the following for the next 4 questions. The spring 2010 student survey of Math 1530 students found the proportion of students who believe in heaven to be about 0.86. We would like to know if current ETSU students have a different belief in heaven than what the 2010 student survey found. The following table is the summary information.

BELIEF IN HEAVEN	Count
YES, DEFINITELY	630
YES, PROBABLY	83
NO, PROBABLY NOT	33
NO, DEFINITELY NOT	35
DON'T KNOW	54
N=	835

13. The hypotheses for a test to answer this question are

(A) $H_0 : p = 0.86, H_a : p < 0.86$ (C) $H_0 : p = 0.86, H_a : p \neq 0.86$

(B) $H_0 : \hat{p} = 0.88, H_a : \hat{p} > 0.88$ (D) $H_0 : \hat{p} = 0.85, H_a : \hat{p} > 0.85$

(E) $H_0 : \mu = 713, H_a : \mu \neq 713$

14. Give the numerical value of the statistic $\hat{p}$ that estimates p.

(A) 0.8600 (B) 835 (C) 0.8539 (D) 713 (E) 0.8800

15. Find the $p-$value for the previous test.

(A) approximately 0.6 (B) approximately 1 (C) approximately 0 (D) between 0.01 and 0.05 (E) between 0.1 and 0.5

16. Interpret your results relative to the survey conducted in 2010 and comment on any assumptions that are needed for your conclusions to be accurate.

(A) There is enough evidence to suggest that the proportion of current ETSU students that believe in heaven is lower than the 2010 student survey. The sample size is large so that we can trust our results.

(B) There is not enough evidence to suggest that the proportion of current ETSU students that believe in heaven is different from the 2010 student survey. It is important that the students who responded to the survey represent all ETSU students.

(C) Reject H_0 since the $p-$value is small. The assumptions should be okay since the sample size is large ($n = 835$).

(D) Reject H_0 and conclude that the percentage of all ETSU students that believe in heaven is 86%. It is important that the data follow the Normal distribution and there are no outliers.

(E) There is no need to worry about the assumptions since the sample size is large ($n = 835$).

17. A study published in the *New England Journal of Medicine* (Aug. 2001) suggests that it's dangerous to enter a hospital on a weekend. During a 10-year period, researchers tracked over 4 million emergency admissions to hospitals in Ontario, Canada. Their findings revealed that patients admitted on weekends had a much higher risk of death than those who went to the emergency room on weekdays. The researchers said the difference in death rates was "statistically significant." What does this mean?

(A) They mean that the difference is higher than what was expected to occur by chance.

(B) Perhaps the results are attributable to some confounding variable (e.g., more serious accidents on the weekend).

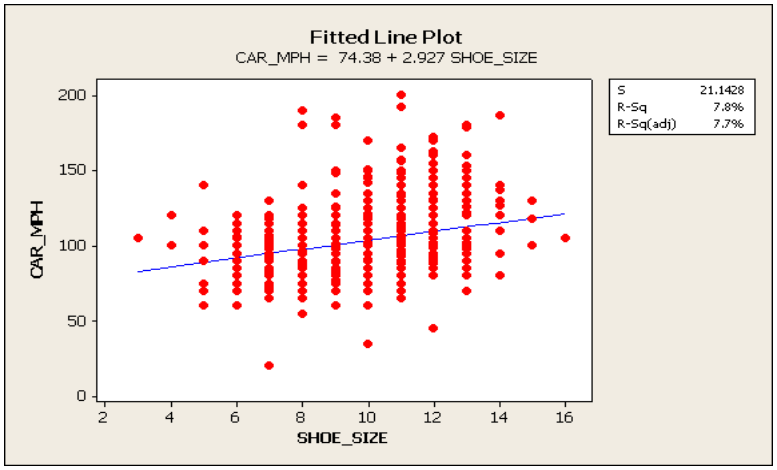
(C) The results are of practical importance and this can be stated with 95% confidence.

(D) The chance that the null hypothesis is true is very small.

(E) The researchers found evidence to suggest that risk of death was the same no matter which day they went to the emergency room.

Use the following for the next 3 questions. Big Lead Foot? From the survey students were asked to give their shoe size and their response to the question "What's the fastest you've ever driven a car (mph)?" In other words, can we predict fastest-speed driven for students based on their shoe size? The *CAR MPH = 0* cases were deleted and the remaining observations along with the least-squares regression line appear in the scatterplot below. The correlation between the two variables is $r = 0.279$ and the least-squares regression line for predicting fastest-speed driven (*CAR MPH*) from shoe size (*SHOE SIZE*) is

$$\texttt{CAR MPH} = 74.38 + 2.927 \times \texttt{SHOE SIZE}.$$



18. Which of the following best describes the relationship between shoe size and fastest-speed driven?

    (A) There seems to be a weak negative linear relationship between shoe size and fastest-speed driven with possible outliers.

    (B) The association is very strong since the points on the scatterplot are positive and are very close to the line.

    (C) We see that low values of shoe size go with high values of car mph. High values of shoe size go with low values of car mph.

    (D) Since $r = 0.279$ we know that for every shoe size increase this causes the student to drive faster.

    (E) There seems to be a weak positive linear relationship between shoe size and fastest-speed driven with possible outliers.

19. Which of the following is the correct interpretation of the slope of the least-squares regression line?

    (A) Since the correlation is 0.279, MPH increases by about 27.9% on the average.

    (B) MPH increases by about 74.4 on the average.

    (C) As MPH increases, shoe size will increase by nearly 3 sizes on the average.

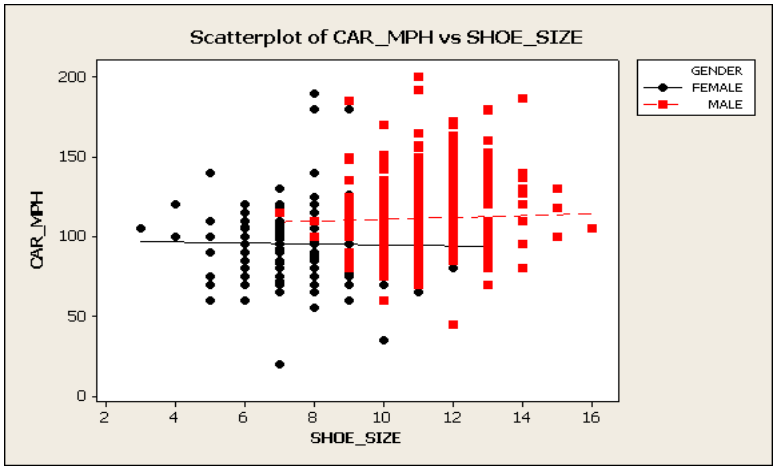
    (D) For every shoe size increase, the fastest-speed driven increases by 74.38 MPH on the average.

    (E) For every shoe size increase a student drives about 2.93 MPH faster on the average.

20. Use the least-squares regression line to predict fastest-speed driven when a student has a shoe size of 10.

    (A) 160 (B) 104 (C) 40 (D) 77 (E) 84

21. **Big Lead Foot continued.** Does it make sense to claim that people with bigger feet will drive faster? The scatterplot includes the regression lines for the males and the females. For the males: $\texttt{CAR MPH} = 105 + .58 \times \texttt{SHOE SIZE}, r = .03$ and for the females: $\texttt{CAR MPH} = 97.3 - .29 \times \texttt{SHOE SIZE}, r = -.03$



(A) There is a mistake, since the correlation must be zero in the combined data because we found a positive correlation for the males and a negative correlation for the females.

(B) It is well know that a person with a large foot will drive fast.

(C) Basically there is no relationship between shoe size and fastest speed within either sex. The observed correlation in the combined data occurs only because men tend to have higher values than women do for both variables. Gender is a lurking variable that impacts both variables.

(D) The faster drivers will tend to have larger feet since both variables tend to be related to a higher level of self esteem.

(E) There is a relationship between shoe size and fastest speed driven and for each size increase in shoe this will cause a driver to increase their speed by an average of 2.93 MPH.

Use the following to answer the next 3 questions about gender and right to die. Question 7 from the 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/her family request it?" Choose a student at random from this group.

Tabulated statistics: Gender, Right to Die

Rows: Gender Columns: Right to Die

	Yes	No	Don't Know	All
Female	236	148	95	479
Male	222	82	55	359
All	458	230	150	838

Cell Contents: Count

22. The conditional probability that the student was female, given that the student believes that doctors should be allowed by law to end the patients life by some painless means if the patient and his/her family request it

(A) 0.2816 (B) 0.4927 (C) 0.5716 (D) 0.5465 (E) 0.5153

23. The probability that the student is a female and believes that doctors should be allowed by law to end the patients life by some painless means if the patient and his/her family request it is about

(A) 0.5153 (B) 0.5716 (C) 0.5465 (D) 0.2816 (E) 0.4927

24. The conditional probability that the student believes that doctors should be allowed by law to end the patients life by some painless means if the patient and his/her family request it given that the student was a female, is about

(A) 0.5716 (B) 0.2816 (C) 0.4927 (D) 0.5465 (E) 0.5153

25. The weather reporter predicts that there is a 90% chance of rain tomorrow for a certain region. What is meant by this phrase?

    (A) Rain occurs 90% of the time in this region.

    (B) It will rain 90% of the day tomorrow.

    (C) 90% of the time it rains on this date.

    (D) In circumstances "like this," rain occurs 90% of th

    (E) The occurrences of rain is "truly random" and will occur 90% of the time.

26. Consider the following example from research on statistical reasoning (Nisbett, et al., 1987 in *Science*). There are two hospitals: in the first one, 120 babies are born every day, in the other, only 12. On average, the ratio of baby boys to baby girls born every day in each hospital is 50/50. However, one day, in one of those hospitals twice as many baby girls were born as baby boys. In which hospital was it more likely to happen? Explain.

    (A) It is much more likely to happen in the larger hospital since the number of births is larger. This is called the law of large numbers.

    (B) It is much more likely to happen in the small hospital. The reason for this is that technically speaking, the probability of a random deviation of a particular size (from the population mean), decreases with the increase in the sample size.

    (C) In the larger hospital 80/120 is greater than 8/12.

    (D) Equally likely 80/120 is equal to 8/12.

    (E) It is equally likely to be 20 babies far from 60 than 2 babies far from 6.

27. Grocery store receipts show that customer purchases have a skewed distribution with a mean of $32 and a standard deviation of $20. Is it likely that the next 50 customers will spend an average of at least $40?

    (A) No. The probability is approximately 0.0023

    (B) No. The probability is approximately .40

    (C) No. The probability is approximately .34

    (D) Yes. The probability is approximately .66

    (E) Yes. The probability is approximately 0.9977

28. Administrators at a hospital are concerned about the possibility of drug abuse by people who work there. They decide to check on the extent of the problem by having a random sample of employees undergo a drug test. Each employee has a 4-digit ID number. Randomly choose 40 numbers. Name the sampling strategy.

    (A) simple random sampling (B) stratified sampling (C) observational study (D) experiment (E) voluntary response

29. Music and Memory. Is it a good idea to listen to music when studying for a big test? In a study conducted by some Statistics students, 62 people were randomly assigned to listen to rap music, Mozart, or no music while attempting to memorize objects pictured on a page. They were asked to list all the objects they could remember. *Who* are the individuals in this survey?

    (A) Statistics students

    (B) 62 people

    (C) Rap, Mozart, no music

    (D) Music and Memory

    (E) Number of objects correctly memorized and type of music.

30. To check the effect of cold temperatures on a battery's ability to start a car researchers purchased a battery from Sears and one from NAPA. They disabled a car so it would not start, put the car in a warm garage, and installed the Sears battery. They tried to start the car repeatedly, keeping track of the total time that elapsed before the battery could no longer turn the engine over. Then they moved the car outdoors where the temperature was below zero. After the car had chilled there for several hours the researchers installed the NAPA battery and repeated the test. Is this a good experimental design?

    (A) Yes, because the researchers used control, randomization, and replication.

    (B) No, because temperature is confounded by brand.

    (C) No, because the car and the batteries were not chosen at random.

    (D) No, because they should have tested other brands of batteries, too.

    (E) No, because they should have tested more temperatures.

Use the following for the next 3 questions. A recent study gave evidence that the drug lovastatin can reduce fat deposits in arteries that are linked to heart attacks. In the study, 270 subjects with moderate blood cholesterol levels were assigned at random to take either lovastatin or a placebo. Fat deposits in arteries were measured using X-rays. After two years, the lovastatin group had less arterial fat.

31. The design of this study is called a

(A) simple random sample. (C) least-squares regression.

(B) an observational study. (D) matched-pair design.

(E) randomized comparative experiment.

32. The response variable in the study is

(A) X-ray. (B) lovastatin or placebo. (C) blood cholesterol. (D) heart attacks. (E) how much fat is deposited in arteries.

33. The study was *double-blind*. This means that

(A) a placebo was used.

(B) the treatment lovastatin was used.

(C) subjects were randomly (blindly) assigned to groups.

(D) neither the subjects nor the medical personnel working with them knew who took lovastatin and who took the placebo.

(E) results about individual subjects were not released, only statistical summaries for the groups.

Use the following for the next 4 questions. A nutrition laboratory tests 9 "reduced sodium" hot dogs. Here are the sodium contents for 9 hot dogs.

353 333 363 315 352 299 315 312 363

34. Find the sample mean of the sodium content for these hot dogs.

(A) 352 (B) 300.5 (C) 340 (D) 315 (E) 333.9

35. Find a 95% confidence interval for the mean sodium content of this brand of hot dog. It is known that the standard deviation of the sodium content for all hot dogs of this brand is $\sigma = 36$ mg.

(A) $(326.1, 341.7)$ (B) $(330.0, 337.8)$ (C) $(310.4, 357.4)$ (D) $(322.5, 345.3)$ (E) $(261.9, 405.9)$

36. What does the calculated confidence interval mean?

(A) There is a 95% probability that the true μ for all hot dogs of this type is within the 95% confidence interval.

(B) A different sample of the same size and same $C = 95\%$ would have the exact same confidence interval as the first sample did.

(C) We are certain the true value of μ is in the interval that was calculated.

(D) We are 95% confident that all sample means will be within the calculated confidence interval.

(E) With 95% confidence, the population mean sodium content for this brand of "reduced sodium" hot dogs is between the calculated interval.

37. What assumptions have you made in this inference?

(A) The level of confidence is normally distributed. (C) Sample size is small.

(B) Successes and failures are each greater than 10. (D) Random sample and Normal population

(E) The sodium contents must vary.

38. It has been reported that taking Vitamin C can reduce your chance of getting the common cold. Suppose that the added benefit of taking Vitamin C reduces your chances by .1% for getting the common cold. A statistical test is more likely to find a significant decrease in the percent of colds if

(A) it is based on a very large sample. (C) The size of the sample doesn't have any effect on the significance of the test.

(B) it is based on a very small sample. (D) the $p-$value is large.

(E) the $p-$value is smaller than .1%.

39. A medical researcher has tested a new treatment for dandruff against a popular brand of treatment. She concludes that the new treatment, with a $p-$value of 4%, is more effective. What does the $p-$value mean in this context?

    (A) Do not Reject H_0 since the $p-$value is smaller than 5% and conclude that the popular brand is more effective in treating dandruff.

    (B) The observed effect has a practical significance.

    (C) Reject H_0 and conclude that new treatment and the popular brand are equivalent in treating dandruff.

    (D) The chance that the new treatment equals a popular brand in treating dandruff is 4%.

    (E) If there is no difference between the new treatment and a popular brand for treating dandruff, the chance of seeing an observed difference this large or larger is 4% by natural sampling variation.

40. On April 20-23, 2011, the Gallup Poll obtained a random sample, based on telephone interviews, of 1,013 adults nationwide. When asked, "Would you vote for Donald Trump if he ran for president in 2012?" 64% said "Definitely not". Gallup reported their margin of error to be under $\pm 4\%$. It is standard among pollsters to use 95% confidence level unless otherwise stated. Given that, what does the Gallup Poll mean by claiming a margin of error of $\pm 4\%$ in this context?

    (A) Gallup dialed landline telephone numbers at random and so missed all people without landline phones, including people whose only phone is a cell phone.

    (B) If this polling were done repeatedly, 95% of all random samples would yield estimates that come within $\pm 4\%$ of the true proportion of adults who would not vote for Donald Trump if he ran for president in 2012.

    (C) Some people whose numbers were chosen never answered the phone in several calls or answered but refused to participate in the poll.

    (D) The respondents did not tell the truth about how they felt about Donald Trump.

    (E) There will always be mistakes in conducting polls and the margin error tells us the percent of all these errors to expect.

41. From the Math 1530 survey students were asked for their opinion on the following question on marijuana: "Which argument do you find to be more persuasive? *Legalizing marijuana will result in increased health care costs OR Legalizing marijuana will increase local and federal tax revenue*." Do males and females view the question of about the legalization of marijuana the same? The Minitab output shows the two-way table and related information to answer this question.

Tabulated statistics: Gender, Marijuana
Using frequencies in Counts

Rows: Gender	Columns: Marijuana		
	0	1	All
Female	320	156	476
	67.23	32.77	
Male	295	63	358
	82.40	17.60	
All	615	219	834

Cell Contents: Count
% of Row

Pearson Chi-Square = 24.300, DF = 1, P-Value = 0.000

Note: 1 = Legalizing marijuana will result in increased health care costs
0 = Legalizing marijuana will increase local and federal tax revenue

What is the correct conclusion of this analysis? (Use $\alpha = 0.05$.)

(A) It appears that females and males have the same opinion about the legalization of marijuana.

(B) The number of females who responded to the survey is statistically different from the number of males that responded to the survey.

(C) There does appear to be strong evidence to suggest that females and males differ in their opinion about the legalization of marijuana.

(D) We see that 0 is statistically different from 1.

(E) Do not reject the null hypothesis.

42. Based on the class survey, we want to know if there is good evidence that all female ETSU students and all male ETSU students differ in their mean number of work hours per week? Assume that the students that responded to the survey represent all ETSU students. Below you see four outputs from Minitab.

```
Output 1:
Two-Sample T-Test and CI: WORK_HOURS, GENDER

Two-sample T for WORK_HOURS

GENDER    N  Mean  StDev  SE Mean
FEMALE  478  13.5   14.1     0.65
MALE    343  14.2   14.9     0.80

Difference = mu (FEMALE) - mu (MALE)
Estimate for difference:  -0.75
95% CI for difference:  (-2.78, 1.27)
T-Test of difference = 0 (vs not =): T-Value = -0.73  P-Value = 0.465  DF = 714
------------------------------------------------------------------------------
Output 2:
One-Sample T: WORK_HOURS

Variable      N    Mean   StDev  SE Mean       95% CI
WORK_HOURS  821  13.777  14.447    0.504  (12.787, 14.767)
------------------------------------------------------------------------------
Output 3:
One-Sample T: FEMALE

  N    Mean   StDev  SE Mean       95% CI
478  13.500  14.100    0.645  (12.233, 14.767)

One-Sample T: MALE

  N    Mean   StDev  SE Mean       95% CI
343  14.200  14.900    0.805  (12.618, 15.782)
------------------------------------------------------------------------------
Output 4:
Test and CI for Two Proportions

Sample    X    N  Sample p
1       478  821  0.582217
2       343  821  0.417783

Difference = p (1) - p (2)
Estimate for difference:  0.164434
95% CI for difference:  (0.116724, 0.212144)
Test for difference = 0 (vs not = 0):  Z = 6.76  P-Value = 0.000
------------------------------------------------------------------------------
```

Which of the following options answers the research question?

(A) Output 1; A 95% confidence interval for the difference in the true mean number of work hours per week is anywhere from -2.78 to 1.27 hours. Since 0 is between these two numbers there does not appear to be good evidence to suggest that all female students and all male students differ in their mean number of work hours per week.

(B) Output 2; We are 95% confident that the mean number of work hours for ETSU students ranges from about 13 to 15 hours. Hence, since this interval doesn't contain 0 the results are statistically significant.

(C) Output 3; We are 95% confident that the mean number of work hours per week for females is about 12 to 15 hours. We are 95% confident that the mean number of work hours per week for males is about 13 to 16 hours. Hence, there is statistical evidence that males work one hour per week longer than females.

(D) Output 4; There is evidence to suggest that the proportion of females is larger than the proportion of males. Hence, there is a larger percent of females that work longer than the males.

(E) All outputs tell us to reject H_0. Hence, the research question has been verified.