

Math 1530 Final Exam Spring 2014

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 Spring 2014 MATH1530 class survey produced variables that are quantitative?

- What is your average one-way commute time to campus in minutes?
- Have you ever had your teeth examined by a dentist? (Yes, No)
- Approximately how many dental cavities have you ever had?
- Are you a first generation college student? (Yes, No)
- What type of water do you drink? (Bottled, Tap, Filtered, Other)

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

2. In the Spring 2014 MATH1530 class survey asked “Approximately how many dental cavities (either active or filled) have you ever had?” The table below represents the responses of 651 students.

Number of Cavities	0	1	2	3	4	5	6 or more
Number of Students	154	75	98	83	82	49	110

The median number of cavities for these 651 students is

(A) 2 (B) 3 (C) 83 (D) 98 (E) Unable to determine from the given information.

Use the following for the next 2 questions. The table below represents the responses of 651 students to the MATH1530 survey question “What is your Zodiac sign?”

Zodiac Sign	Aries	Taurus	Gemini	Cancer	Leo	Virgo	Libra	Scorpio	Sagittarius	Capricorn	Aquarius	Pisces
Counts	45	65	57	68	65	51	64	43	38	52	50	53

3. Which type of graph is appropriate for these data?

(A) Pie Chart (B) Histogram (C) Stem Plot (D) Scatterplot (E) Boxplot

4. What proportion of the 651 students were born under the Zodiac sign of Cancer?

(A) $p = 0.08$ (B) $p = 0.104$ (C) $\bar{x} = 54.25$ (D) median = 52.5 (E) frequency = 68

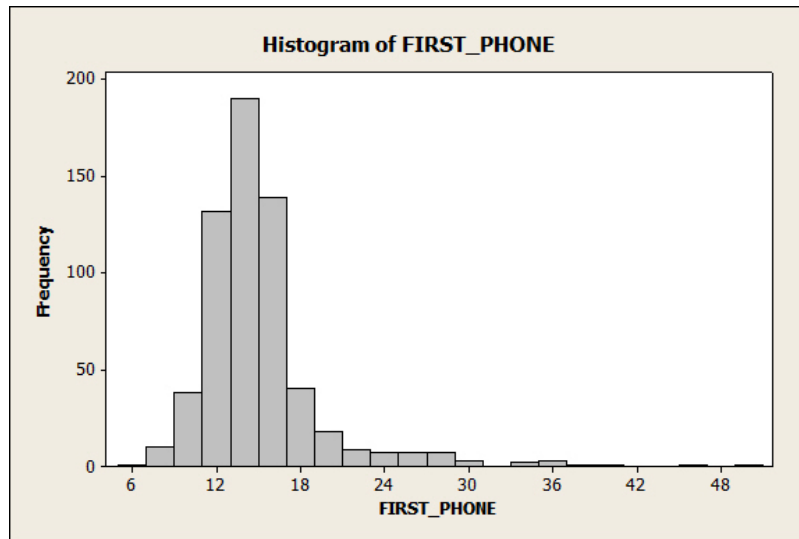
5. You have calculated that the value of the variance for a set of data is -6.8 . You can conclude that

(A) the data is left skewed. (C) there is very little variation in the data.

(B) the data is right skewed. (D) there is a lot variation in the data.

(E) you made an error in your calculations.

6. The Spring 2014 MATH1530 class survey asked “At what age did you get your first cell?” Assume that the students who responded to the survey are from an SRS of all ETSU students. The histogram below is the distribution of age from the survey of 610 students.



Which of the following best describes the distribution?

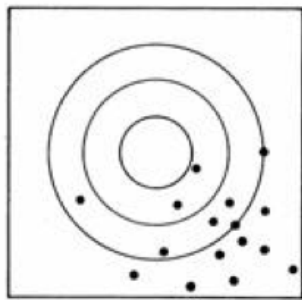
- (A) The distribution is bell-shaped. (C) The distribution is skewed left with outliers.
 (B) The distribution is skewed right with possible outliers. (D) The center is 14.
 (E) The distribution is fairly uniform from 6 to 50.
7. Environmental Protection Agency fuel economy estimates for automobile models tested predicted a mean of 24.8 miles per gallon (mpg) and a standard deviation of 6.2 mpg for highway driving. Assume that mpg can be approximated by the Normal distribution. About what percent of automobiles should get between 18.6 and 31 mpg?
- (A) 32% (B) 50% (C) 68% (D) 95% (E) 99.7%

Use the following information for the next 2 questions. Elephants sometimes damage crops in Africa. It turns out that elephants dislike bees. They recognize beehives in areas where they are common and avoid them. Can this be used to keep elephants away from crops that have trees nearby? A group in Kenya placed active beehives in some trees, empty beehives in others, and none in others. Will elephant damage of the crops be less that have trees with hives? Will even empty hives keep elephants away?

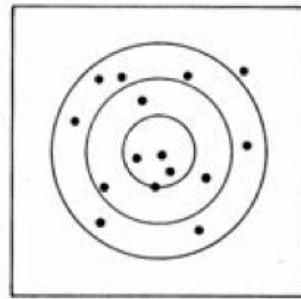
8. This experiment has
- (A) two factors, beehives present or absent. (C) matched pairs.
 (B) stratification by beehive. (D) two independent samples.
 (E) three treatments.
9. The response in this experiment is
- (A) the type of crop. (C) the presence or absence of hives.
 (B) the presence or absence of bees. (D) elephant damage.
 (E) location of trees.

10. Think of the true value of the population parameter as the bull's-eye on a target, and of the sample statistic as an arrow fired at the bull's-eye. Bias and variability describe what happens when an archer fires many arrows at the target. The figure below shows this target illustration of the two types of error. Match the figure with the following

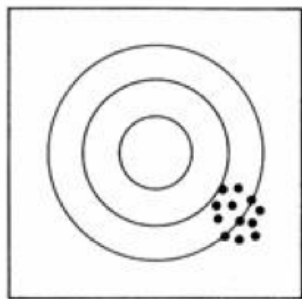
- i. High bias, low variability
- ii. Low bias, high variability
- iii. High bias, high variability
- iv. Low bias, low variability



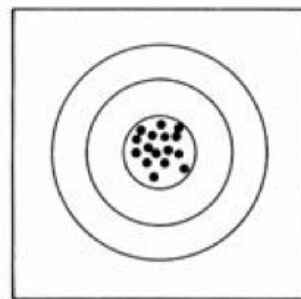
(a)



(b)



(c)



(d)

- (A) a = iii, b = ii, c = i, d = iv (C) a = i, b = iv, c = iii, d = ii
 (B) a = ii, b = iii, c = iv, d = i (D) a = iv, b = i, c = ii, d = iii
 (E) a = iv, b = ii, c = i, d = iii

11. You flip a fair coin four times and get heads every time. What are the chances of flipping a tails on the fifth time?

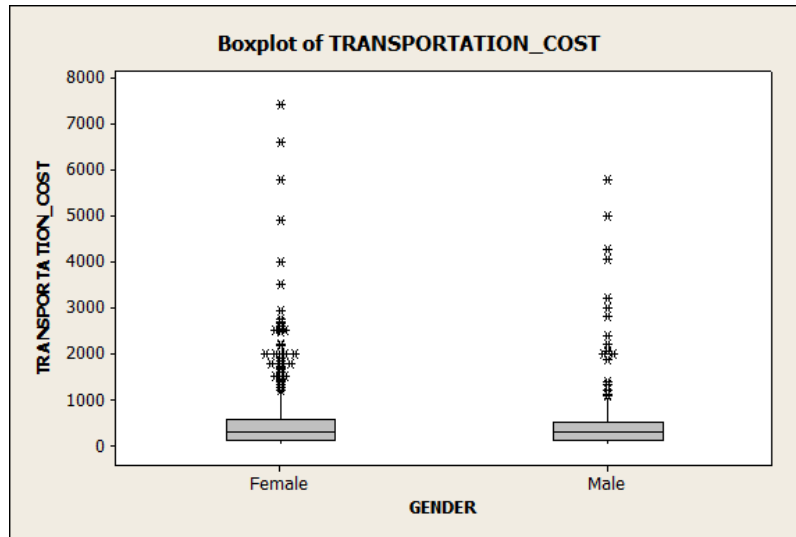
- (A) $1/2$ (B) $1/5$ (C) $4/5$ (D) 1 (E) 0

12. A pediatrician randomly selects patients she saw and records the number of ear infections that her patients have during the winter. A distribution of probabilities for number of ear infections is displayed below. What is the probability that random patient will have at least 2 ear infections?

Number of ear infections	0	1	2	3	4	5
Probability	0.05	0.16	0.25	0.33	0.15	0.06

- (A) 0.21 (B) 0.25 (C) 0.46 (D) 0.54 (E) 0.79

Use the following for the next 3 questions. The Spring 2014 MATH1530 research assignment asked students to estimate the cost of the air ticket(s), fares, or estimated cost of gas to arrive at the port of departure for their cruise. The boxplot and descriptive statistics below summarize the student responses by gender of the cost in dollars.



Descriptive Statistics: TRANSPORTATION_COST

Variable	GENDER	N	N*	Mean	SE Mean	StDev	Minimum	Q1
TRANSPORTATION_COST	Female	372	0	568.7	44.2	851.7	20.0	135.0
	Male	238	0	506.1	48.9	755.0	30.0	127.5

Variable	GENDER	Median	Q3	Maximum
TRANSPORTATION_COST	Female	300.0	562.0	7430.0
	Male	300.0	500.0	5772.0

13. About what percent of the students that are female had a transportation cost of at most \$135?
- (A) $135/372 = 36\%$ (C) 25%
- (B) $z = \frac{135-568.7}{851.7} = -0.51$ (Area under Normal Curve = 31%) (D) 75%
- (E) 50%
14. Based on the above information, which of the following descriptions is false?
- (A) The highest total transportation cost in the survey was from a female student.
- (B) The data are strongly skewed right with several outliers for both genders.
- (C) The data are strongly skewed left with several outliers for both genders.
- (D) The interquartile range of the total transportation cost is larger for females.
- (E) The value of the second quartile is the same for both genders.
15. Which of the following numerical measures should be used to describe the distribution of total transportation cost for males?
- (A) $\bar{x} = 506.1, s = 755$ (C) Median = 300, $\bar{x} = 506.1$
- (B) Median = 300, $s = 755$ (D) Min = 30, $Q_1 = 127.5$, Median = 300, $Q_3 = 500$, Max = 5772
- (E) Min = 30, $Q_1 = 127.5$, $\bar{x} = 506.1$, $Q_3 = 500$, Max = 5772

16. The mean ACT of all students entering ETSU in 2012 was 22.14. The registrar plans to look at records of students entering fall 2014 to see if the mean ACT has changed. State the null and alternative hypotheses for this investigation.

- (A) $H_0 : \mu = 22.14$ vs. $H_a : \mu \neq 22.14$ (C) $H_0 : p = 22.14$ vs. $H_a : p \neq 22.14$
 (B) $H_0 : \mu = 22.14$ vs. $H_a : \mu < 22.14$ (D) $H_0 : \bar{x} = 22.14$ vs. $H_a : \bar{x} \neq 22.14$
 (E) $H_0 : \bar{x} = 22.14$ vs. $H_a : \bar{x} > 22.14$

Use the following for the next 2 questions. A marketing analyst wishes to know if males or females buy more cell phones in order for the company's advertisements to target that specific gender. After talking to the company's current sales representatives across the US, he concludes that females buy more cell phones and thus, have owned more cell phones in their lifetime compared to males. Question 7 from the survey asked students "How many cell phones have you owned?" 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 have owned more cell phones, on average, than male students?

GENDER	N	Mean	StDev	SE Mean
Female	372	5.49	4.11	0.21
Male	238	4.75	3.23	0.21

17. To perform the hypothesis test, which analysis should we conduct?

- (A) A Chi-square test. (C) The z-test.
 (B) A matched pairs t test. (D) One sample t test.
 (E) Two sample t test.

18. The appropriate analysis was performed and the P-value was 0.028. Which of the following correctly answers the research question? Use $\alpha = 0.05$.

- (A) 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 have owned more phones than male students at ETSU.
 (B) The P-value is smaller than α so we will reject the null hypothesis. There is strong evidence that on average females have owned more phones 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 have owned more phones 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 evidence to indicate that on average females have owned more phones than male students at ETSU.
 (E) Cannot draw valid conclusion since the number of male students is not equal to the number of female students.

19. A certain type of calculator battery has a mean lifetime of 100 hours and a standard deviation of $\sigma = 10$ hours. A company has developed a new battery and claims it has a longer mean life. A random sample of 2000 batteries is tested, and their sample lifetime is $\bar{x} = 100.5$ hours. A test is conducted of the hypotheses

$$H_0 : \mu = 100 \text{ vs. } H_a : \mu > 100$$

and P-value of the z-test is 0.013. The battery manufacturer says that because the evidence is strong that $\mu > 100$, you should be willing to pay a much higher price for its battery than for the old type of battery.

- (A) Since the P-value is extremely small the results are both statistically and of practical significance. We should buy the new battery at a much higher cost.
 (B) Even though the result is statistically significant, we cannot conclude that it is of practical significance. Hence, we should not buy the new battery.
 (C) Since the P-value is small, we should not reject H_0 . The results are not statistically significant. Hence, the new battery is much better.
 (D) Since the P-value is small, we should not reject H_0 . The results are only of practical significance not statistically significance.
 (E) Cannot draw a valid conclusion since a two-sided test was not conducted.

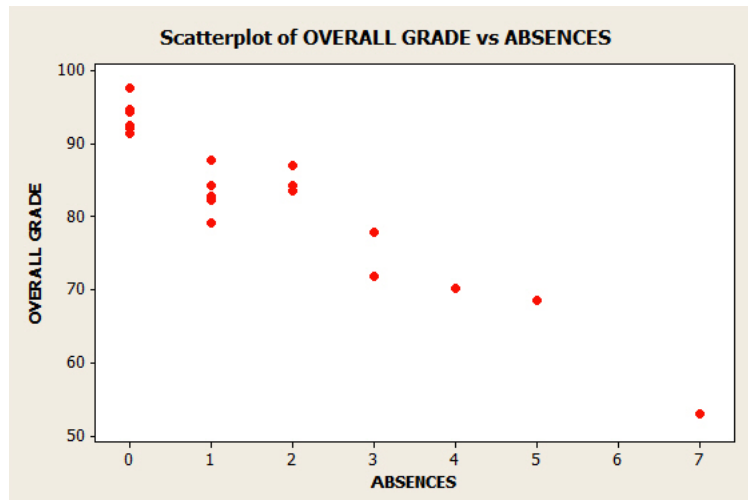
20. A headline on the website E! Science news in December 2009 stated, "Exposure to tobacco smoke home is associated with early emphysema in adulthood." Which type of study is this most likely to have been?
- (A) a controlled experiment (C) two-sample t-test
(B) linear regression (D) double-blind randomized experiment
(E) an observational study
21. A tax assessor wants to assess the mean property tax bill for all homeowners in Madison, Wisconsin. A survey ten years ago yielded a standard deviation of \$1000. How many tax records should be sampled for a 90% confidence interval to have a margin of error of \$100?
- (A) 271 (B) 270 (C) 17 (D) 16 (E) 100
22. An automobile insurer has found that repair claims have a mean of \$1876 and a standard deviation of \$820. Suppose that the next 100 claims can be regarded as a random sample from the long-run claims process. The probability that the average amount of the costs of the 100 claims is less than \$1687 is
- (A) 0.0107 (B) 0.4090 (C) 0.5910 (D) 0.6800 (E) 0.9893

Use the following for next 3 questions. In a study of human development, investigators showed two movies of different types to groups of children. Crackers were available in a bowl, and the investigators compared the number of crackers eaten by children watching both movies. One movie was shown at 8 AM (right after the children had breakfast) and the other at 11 AM (right before the children had lunch). It was found that during the movie shown at 11 AM, more crackers were eaten than during the movie shown at 8 AM. The investigators concluded that the different types of movies had different effects on appetite.

23. The results cannot be trusted because
- (A) the study was not double-blind. Neither the investigators nor the children should have been aware of which movie was being shown.
(B) the investigators were biased. They knew beforehand which movie was better.
(C) the investigators should have used several bowls, with crackers randomly placed in each.
(D) the time each movie was shown is a confounding variable.
(E) this is an observational study.
24. The response variable in this experiment is
- (A) the bowls.
(B) the time each movie was shown.
(C) the different kinds of movies.
(D) the number of crackers eaten.
(E) the children.
25. The treatment in this experiment is
- (A) the bowls.
(B) the time each movie was shown.
(C) the different kinds of movies.
(D) the number of crackers eaten.
(E) the children.

Use the following for the next 3 questions. A certain university is interested in whether the number of absences (in days) in class in a semester affects the students overall grade (in points). The scatter below shows the observations and the least-squares regression line. The correlation coefficient was found to be $r = -0.944$ and the least-squares regression line for predicting the students overall average from the number of absences is

$$\text{OVERALL GRADE} = 92.8 - 5.54 \times \text{ABSENCES}.$$



26. Which of the following best describes the relationship between OVERALL GRADE and ABSENCES?
- (A) There seems to be a strong positive linear relationship between the two variables.
 (B) There seems to be a strong negative linear relationship between the two variables.
 (C) The association is very weak.
 (D) The average OVERALL GRADE of the students is 92.8 and the average number of absences is 5.54.
 (E) The relationship is skewed right because there is more data on the left and the tail extends out to the right.
27. Which of the following is the correct interpretation of the slope of the least-squares regression line?
- (A) Since the correlation is -0.944 , OVERALL GRADE increases by about 94.4% on the average.
 (B) As OVERALL GRADE increases by 1, ABSENCES will decrease by about 92.8 on the average.
 (C) As OVERALL GRADE increases by 1, ABSENCES will decrease by about 5.54 on the average.
 (D) As ABSENCES increases by 1, OVERALL GRADE will decrease by about 92.8 on the average.
 (E) As ABSENCES increases by 1, OVERALL GRADE will decrease by about 5.54 on the average.
28. Use the least-squares regression line to predict the OVERALL GRADE if a student has 3 ABSENCES.
- (A) $92.8 + 3 = 95.3$ (C) $92.8 - 3 = 89.8$
 (B) $92.8 - 5.54(3) = 76.18$ (D) Any value between 40 and 60.
 (E) Any value between 80 and 100.
29. A high correlation between two variables does not always mean that changes in one cause changes in the other. The best way to get good evidence that cause-and-effect is present is to
- (A) select a simple random sample from the population of interest.
 (B) carry out a randomized comparative experiment.
 (C) arrange the data in a two-way table.
 (D) make a scatterplot and look for a strong association.
 (E) make a histogram and look for outliers.

30. An experiment compares two brands of automobile tires. Each of a number of cars is equipped with one tire of each brand on a rear wheel (the order is randomized from car to car) and tread wear is measured periodically. This is called a
- (A) simple random sample. (C) block design.
 (B) completely randomized design. (D) matched pairs design.
 (E) observational study.

Use the following for the next 3 questions. A researcher wishes to estimate the proportion of X-ray machines that malfunction. A random sample of 40 machines is taken and 16 of the machines malfunction. Suppose that the proportion of X-ray machines that malfunction follows a normal distribution.

31. A 90% confidence interval for p is
 (A) 0.4 ± 1.645 (B) 16 ± 1.645 (C) 0.4 ± 0.127 (D) 0.4 ± 40 (E) 16 ± 0.127
32. Which of the following is the best interpretation of the above calculated confidence interval?
- (A) There is a 90% probability that the true proportion of X-ray machines that malfunction is within the calculated interval.
 (B) The proportion of X-ray machines that malfunction is within the 90% confidence interval.
 (C) 90% of all X-ray machines had the proportion of malfunctions within calculated confidence interval.
 (D) With 90% confidence, all the sample proportions will be within the calculated confidence interval.
 (E) With 90% confidence, the true proportion of X-ray machines that malfunction is within the calculated interval.
33. Suppose a random sample of 100 X-ray machines rather than 40 was measured. Which of the following statements is true?
- (A) The margin of error for our 90% confidence interval would stay the same, since the level of confidence has not changed.
 (B) The margin of error for our 90% confidence interval would increase.
 (C) The margin of error for our 90% confidence interval would decrease.
 (D) The standard error would increase.
 (E) The margin of error for our 90% confidence interval could decrease or increase depending on the value of the sample proportion of X-ray machines that malfunctioned.

Use the following for the next 3 questions. A certain major league baseball team has not drawn as many fans as one would expect even though they have had a good winning percentage the past few years. Management randomly interviews 400 fans that come to the games. One issue that management would like to know is whether there are differences in support for the team among various age groups. The distribution of counts is shown below in the table. Choose a fan at random from this group.

		Age of Fan				Total
		Under 25	25 to 40	40 to 54	Over 55	
Games Attended per Year	1 to 2	40	50	25	20	135
	3 to 5	60	20	50	60	190
	Over 5	20	30	10	15	75
Total		120	100	85	95	400

34. The probability that a fan attends over 5 games and is under 25 is
 (A) 0.05 (B) 0.0563 (C) 0.30 (D) 0.1875 (E) 0.4375
35. The conditional probability that the fan was under 25, given that the fan attended over 5 games, is about
 (A) 0.4875 (B) 0.30 (C) 0.2667 (D) 0.1875 (E) 0.1667
36. The conditional probability that the fan attended over 5 games, given that fan was under 25, is about
 (A) 0.4875 (B) 0.30 (C) 0.2667 (D) 0.1875 (E) 0.1667

37. How much does an average cruise cost? Yahoo reports that cruises on the major cruise lines Carnival, Royal Caribbean, Norwegian, Celebrity and Princess will cost about \$100 to \$125 PER DAY, per person. Seventy students reported on the Math 1530 survey that they would be traveling alone. Assume these students represent a SRS of all students that would take a cruise alone. The analysis below is based on the cost of the trip divided by the number of days of the trip.

One-Sample T: cost/days

Test of $\mu = 125$ vs > 125

Variable	N	Mean	StDev	SE Mean	95% Lower Bound	T	P
cost/days	70	234.1	177.3	21.2	198.7	5.15	0.000

Stem-and-Leaf Display: cost/days

Stem-and-leaf of cost/days N = 70

Leaf Unit = 10

```

1   0  3
15  0  56666677888999
29  1  00000112222444
(12) 1  557777888889
29  2  0000011
22  2  57
20  3  0011
16  3  569
13  4  0124
9   4  69
7   5  0
6   5  9
5   6  2
4   6  68
2   7  11

```

Is there significant evidence at the 5% level that the mean cruise cost per day is greater than the reported \$125 per day?

- (A) No. There is weak evidence ($P = 0$) that, on average, the mean cruise cost per day is greater than \$125 per day.
 (B) There is strong evidence ($P = 0$) that, on average, the mean cruise cost per day is greater than \$125 per day.
 (C) No. We fail to reject H_0 since the P is close to 0.
 (D) The probability that H_0 is true is 1 since the P -value is 0.
 (E) The data are not statistically significant since the P -value is small.
38. Data on the blood cholesterol levels of 16 rats (milligrams per deciliter) give $\bar{x} = 85$ and $s = 12$. A 99% t-confidence interval for the mean blood cholesterol of rats is 72.7 to 97.3. Which of the following would cause the most worry about the validity of this confidence interval?
- (A) There is a clear outlier in the data.
 (B) A stemplot of the data shows a mild right-skew.
 (C) You do not know the population standard deviation σ .
 (D) A random sample of only 16 observations were used.
 (E) There is no concern about the validity of this confidence interval.

39. The council of a municipality wanted to determine the opinion of residents about the installation of the traffic cameras installed at the corner of a dangerous intersection. They use the telephone book to randomly select 5000 residential addresses of the city and mail a survey to the selected sample. After two weeks the completed surveys were returned and the results of the survey were revealed. Of the 1500 returned surveys they found 75% were against the traffic cameras stating it was an invasion of privacy. The city council also reported the margin of error was $\pm 3\%$. How should we interpret these results?
- (A) The numbers don't add up since it appears that only 1125 citizens were against the traffic cameras. So it appears that the rest of the population is for the traffic cameras.
 - (B) We are 95% confident that the percentage of all the citizens that oppose the traffic cameras lies between 72% and 78%.
 - (C) The margin of error tells us that all sources of error (sampling variability, nonresponse and undercoverage) are covered and it appears that a strong majority of the citizens oppose the traffic cameras.
 - (D) The sample size from the phone book should have been larger to decrease the margin of error.
 - (E) The results of the survey should not be trusted since it suffers from undercoverage and nonresponse.
40. An engineer designs an improved automobile engine for a certain make of automobile. The previous design had an average highway gas mileage of 30 mpg (miles per gallon). The new engine has an average highway gas mileage of 30.2 mpg, using a sample of 3600 cars of the model with the new engine. Although the difference is quite small, the effect is statistically significant. The explanation is
- (A) that new designs typically have less variability than standard designs, and so small differences can appear to be statistically significant.
 - (B) that the mean of 30.2 is large compared to the gas mileage of most cars.
 - (C) that the mean of 30.2 is small compared to the gas mileage of most cars.
 - (D) that the sample size is very large.
 - (E) All of the above.
41. The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures the motivation, attitudes, and study habits of college students. Scores range from 0 to 200 and follow (approximately) a Normal distribution with mean 115 and standard deviation $s = 25$. You suspect that incoming freshman have a mean m , which is different from 115, since they are often excited yet anxious about entering college. To verify your suspicion, you test the hypotheses

$$H_0 : \mu = 115, H_a : \mu \neq 115$$

- You give the SSHA to 25 incoming freshman and find their mean score. Based on this, you reject H_0 at significance level $\alpha = 0.01$. Which of the following would be most helpful in assessing the practical significance of your results?
- (A) Test the hypotheses again, this time using significance level $\alpha = 0.001$.
 - (B) Construct a 99% interval for μ to see the magnitude of the difference between 115 and your sample results.
 - (C) Report the P-value of your test.
 - (D) Take another sample and retest just to make sure the results are not due to chance.
 - (E) None of the above.
42. In assessing the validity of any test of hypotheses, it is good practice to
- (A) determine what assumptions the researchers made.
 - (B) determine exactly how the study was conducted.
 - (C) examine the probability model that serves as a basis for the test by using exploratory data analysis on the data.
 - (D) All of the above.
 - (E) None of the Above.

43. Will consuming a power bar with special carbohydrates before running improve a runner's time on a 5K run? The running club randomly chose eight people who are part of the running club to perform a test to see if this power bar could really help improve someone's running speed. Each person participated in a 5K run around the same track and each individual was timed. A week later, each individual consumed the power bar before running the same 5K run in similar conditions. What does this sample suggest about the power bar improving a runner's speed? Only one of the following four Minitab outputs is correct.

Output 1:
Paired T-Test

	N	Mean	StDev	SE Mean
Difference	8	-0.38	3.07	1.08

T-Test of mean difference = 0 (vs > 0): T-Value = -0.35 P-Value = 0.630

Output 2:
Paired T-Test

	N	Mean	StDev	SE Mean
Difference	8	-0.38	3.07	1.08

T-Test of mean difference = 0 (vs < 0): T-Value = -0.35 P-Value = 0.370

Output 3:
Two-Sample T-Test

	N	Mean	StDev	SE Mean
With Powerbar	8	25.25	3.33	1.2
Without Powerbar	8	25.63	4.27	1.5

Difference = μ (With Powerbar) - μ (Without Powerbar)
Estimate for difference: -0.38

T-Test of difference = 0 (vs <): T-Value = -0.20 P-Value = 0.424 DF = 13

Output 4:
Two-Sample T-Test

	N	Mean	StDev	SE Mean
With Powerbar	8	25.25	3.33	1.2
Without Powerbar	8	25.63	4.27	1.5

Difference = μ (With Powerbar) - μ (Without Powerbar)
Estimate for difference: -0.38

T-Test of difference = 0 (vs >): T-Value = -0.20 P-Value = 0.576 DF = 13

Which of the following options answers the research question at $\alpha = 0.10$?

- (A) Output 1; There is a not significant evidence to suggest that consuming the power bar before running will improve a runner's speed.
- (B) Output 2; There is a not significant evidence to suggest that consuming the power bar before running will improve a runner's speed.
- (C) Output 3; There is a not significant evidence to suggest that consuming the power bar before running will improve a runner's speed.
- (D) Output 4; There is a not significant evidence to suggest that consuming the power bar before running will improve a runner's speed.
- (E) All outputs provide a P-value greater than α . Hence, we can use any of them to address the research question.