Math 1530 Final Exam
December 2010

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.

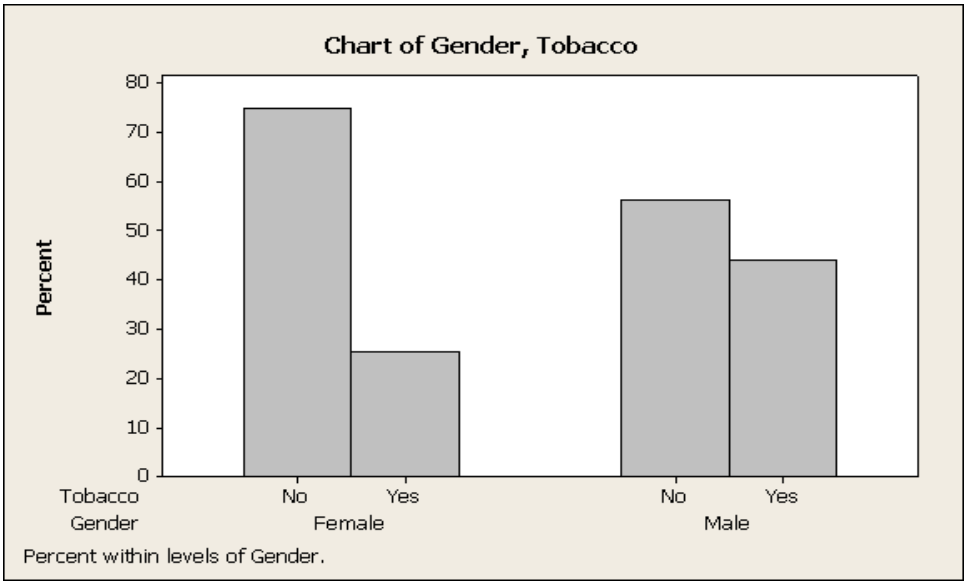
FORM A

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

   i. Are you a male or female?
   ii. How many pairs of shoes do you own?
   iii. Do you text while driving?
   iv. How many Tweets do you follow?
   v. What do you typically drink with dinner?
   vi. How many Facebook friends do you have?

   (A) i, iii, v (B) ii, iv, vi (C) i, ii, iii (D) ii only (E) All these questions.

   Use the following for the next 2 questions. The figure below represents the conditional distributions of tobacco use given sex.



2. Approximately what percent of the females responded "Yes" on tobacco use?
   (A) About 75% (B) About 45% (C) About 25% (D) 55% $\pm$ 10% (E) Unable to determine from the figure.

3. Which of the following is a true statement?

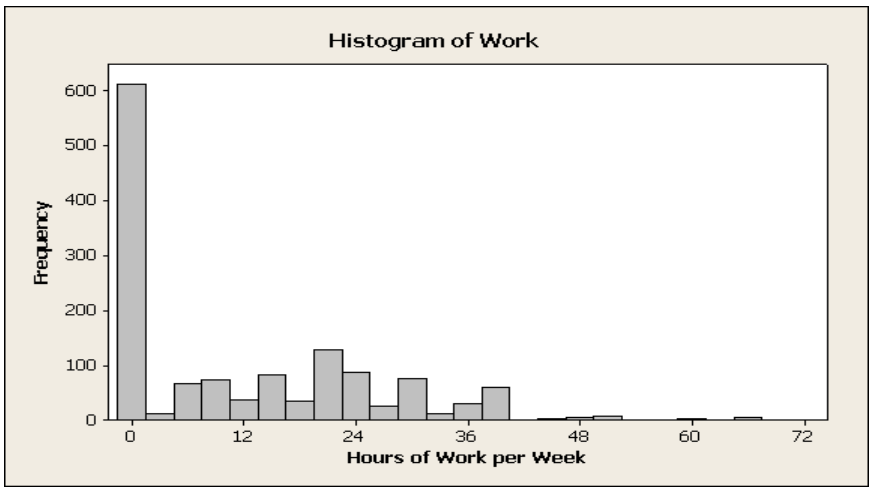
   (A) The bar graphs are bimodal.
   (B) Overall, it appears that a larger percent of males use tobacco than females do.
   (C) The bar graphs are skewed to the left.
   (D) It appears that both graphs are skewed right.
   (E) The bar graphs are skewed to the right so the mean would be larger than the median and the spread ranges from 'No' to 'Yes.'

4. The table below represents the student responses from the class survey to the question "What do you typically drink with dinner?" Which type of graph is appropriate for these data?

Drink with Dinner	Counts
Water	618
Soda	372
Milk	93
Alcohol	20
Juice	124
Other	145

(A) histogram (B) stem plot (C) bar chart (D) scatterplot (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 "How many hours do you work per week?"




```
Descriptive Statistics: hours of work per week

Variable                    N    Mean    SE Mean  StDev  Minimum    Q1   Median     Q3    Maximum
hours of work per week   1352   11.831   0.370  13.619    0.000  0.000   6.000  20.000    65.000
```

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

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

(B) 11.831 hours $\pm$13.619 hours

(C) The data seems to be somewhat symmetric.

(D) There is an outlier and it is 65.

(E) Who: 1352 students, What: number of hours of work per week

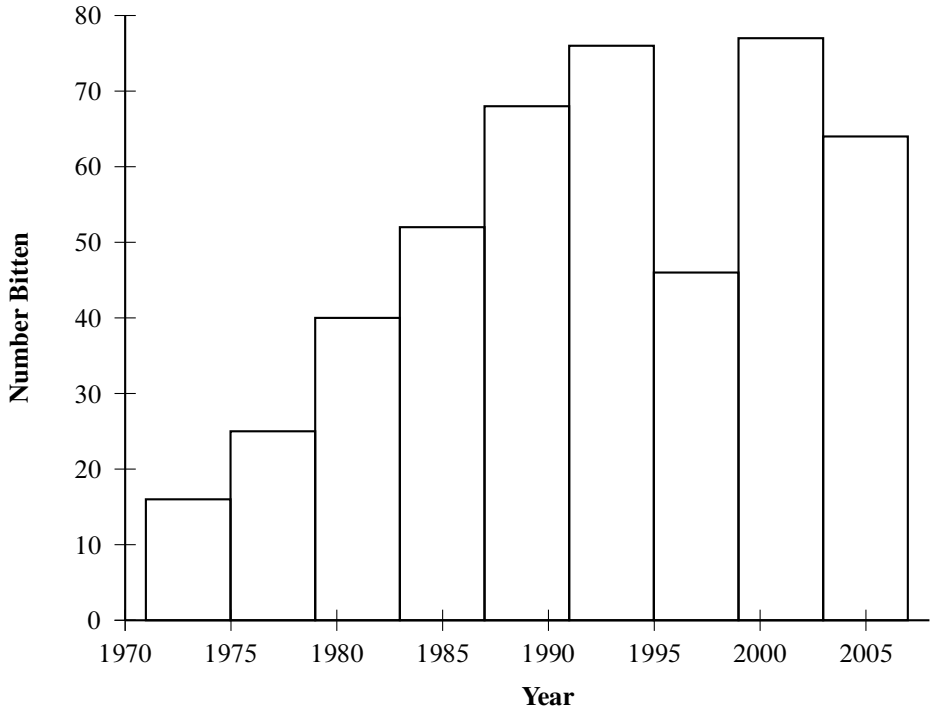
6. About what percent of the students work at most 20 hours per week?

(A) 25% (B) $z = \frac{20-11.831}{13.619} = 0.6$ (Area under Normal Curve = 28%) (C) 50% (D) 20/1352 or 1.5% (E) 75%

7. A student was given an assignment to make a histogram from data showing the number of people bitten by alligators in Florida over 36 years. Here is the data:

year	'72	'73	'74	'75	'76	'77	'78	'79	'80	'81	'82	'83
number bitten	4	3	4	5	2	14	7	2	5	10	7	9
year	'84	'85	'86	'87	'88	'89	'90	'91	'92	'93	'94	'95
number bitten	9	7	23	13	18	13	17	20	15	19	20	22
year	'96	'97	'98	'99	'00	'01	'02	'03	'04	'05	'06	'07
number bitten	13	8	9	16	23	25	17	12	13	15	18	18

The student produced the following display. Comment on this graph.



(A) The distribution of the number bitten looks bimodal with a center of 1989 and a spread of 36 years.

(B) The distribution is skewed left.

(C) The distribution is skewed right with a center of 1989 and a spread of 36 years.

(D) This is not correct since the years are not centered on each of the bars.

(E) This is not correct. The horizontal axis should split the number bitten into classes. The vertical axis should show the number of years in each class.

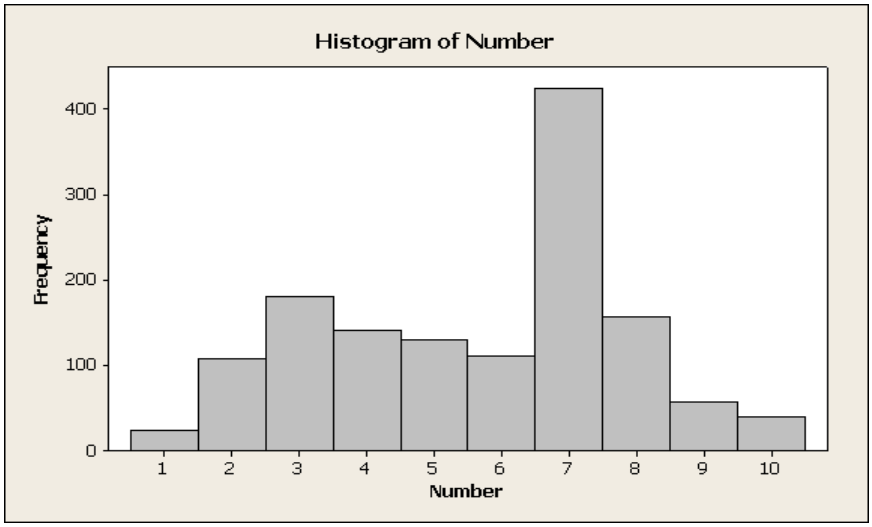
8. The following data is based on a sample of 5 students who responded to the survey questions "What time did you go to bed last night?" (To Sleep) and "What time did you get up this morning?" (Wake Up) Note: The data are coded as follows: -2 for 10 pm, -1 for 11 pm, 0 for midnight, 1.5 for 1:30 am, etc.

To Sleep	1	-3	-1	3	2
Wake Up	6.75	6.5	5	11	7.5

What is the median total number of hours of sleep for these 5 students?

(A) 6.75 (B) 1 (C) 7.5 (D) 6 (E) 7.75

Use the following for the next 2 questions. From the Fall 2010 class survey students were asked to "pick a random number between 1 and 10." The histogram below illustrates the 1368 responses from this survey question.



9. Approximately what percent of the students chose the number 2?
   (A) 100 (B) 80% (C) 110 (D) $2/10 = .2$ or 2% (E) 8%

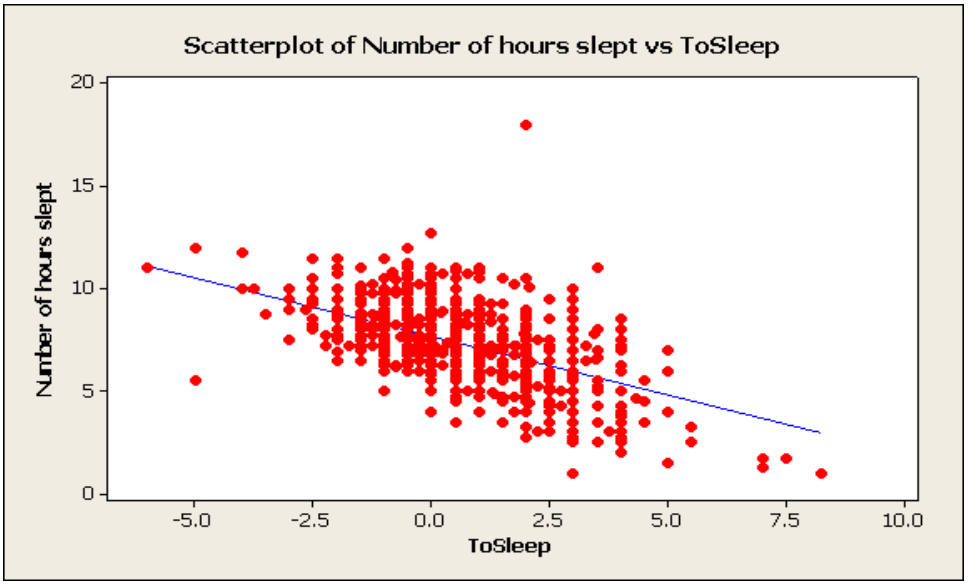
10. Are we good randomizers?
   (A) The distribution is somewhat symmetric about the number 5.
   (B) The distribution is skewed right with a center of about 5.
   (C) It doesn't appear that we are good randomizers. If we were good randomizers we should see a more uniform (flat) distribution across the 10 numbers. We see that the number 7 is picked the most!
   (D) It appears that we are good randomizers since all numbers were chosen.
   (E) The distribution should be more bell-shaped with a center of 5 and standard deviation of 2 to satisfy the 68-95-99.7 percent rule.

11. Scores on the SAT verbal test in recent years follow approximately the Normal distribution with mean $\mu = 504$ and standard deviation $\sigma = 111$. About what percent of these students scored between 282 and 726?
   (A) 68%. (B) 95%. (C) 99.7%. (D) 90%. (E) 97.5.%.

12. The *Information Please Almanac,* (1991, p. 809) reported the divorce rates in the United States for various years from 1960 to 1986. For these same years the *World Almanac and Book of Facts,* reports the percentage of those admitted to state prisons because of drug offenses. A strong positive correlation ($r = 0.67$) was found between the divorce rate and the percentage of criminals admitted for drug offenses. Based on this correlation, advocates of traditional family values may argue that increased divorce rates have resulted in more drug offenses. Is this a sensible conclusion?
   (A) Yes, as the number of drug offenses goes up this is causing the divorce rate to increase.
   (B) Yes, as the divorce rate goes up this causing people to use drugs to cope with their failed relationship.
   (C) No, since the correlation coefficient r is positive we know that as one variable increases the other one goes down.
   (D) No, it is incorrect to compute r here since one variable is a rate and the other variable is a percentage. If both variables were recorded as percentages than we would get a different r.
   (E) Not necessarily. A closer look at this data found both variables to have a strong positive correlation with time. In other words, time is a lurking variable.

Use the following for the next 3 questions. Questions 8 & 9 from the survey asked students about bed time and wake up time. (The data are coded as follows: -2 for 10 p.m., -1 for 11 p.m., 0 for midnight, 1.5 for 1:30 a.m., etc.) This raises the question: "Is the length of sleep for a student related to the time that he or she goes to bed?" In other words, can we predict the number of hours of sleep for students based on the time they went to sleep? The observations and least-squares regression line appear in the scatterplot. The correlation between the two variables is $r = -0.523$ and the least-squares regression line for predicting the *Number of hours sleep* of a student from the time that they went *To Sleep* is

$$\texttt{Number of hours slept} = 7.66 - 0.573 \times \texttt{To Sleep}.$$



13. Which of the following best describes the relationship between time to bed and number of hours slept?

    (A) There seems to be a positive linear relationship between time to bed and number of hours slept with possible outliers.

    (B) The association is very weak since the points on the scatterplot do not all fall on a line.

    (C) There seems to be a negative linear relationship between time to bed and number of hours slept with possible outliers.

    (D) We see that low values of bed time go with low values of sleep time. High values of bed time go with high values of sleep time.

    (E) Since $r = -0.523$ we know that for every hour later a student goes to bed we will know that a student gets out of bed earlier.

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

    (A) Since the correlation is -0.523, sleep decreases by about 27.4% on the average.

    (B) Sleep increases by about 7.66 hours on the average.

    (C) A student loses about a half hour of sleep for every hour later that he or she goes to bed.

    (D) As the hours of sleep increases, to bed decreases by about 30 minutes on the average.

    (E) For every hour later that a student goes to bed, the number of hours increases by 7.66 minutes on the average.

15. Use the least-squares regression line to predict the number of hours slept when a student goes to bed at midnight.

    (A) .784 hours.

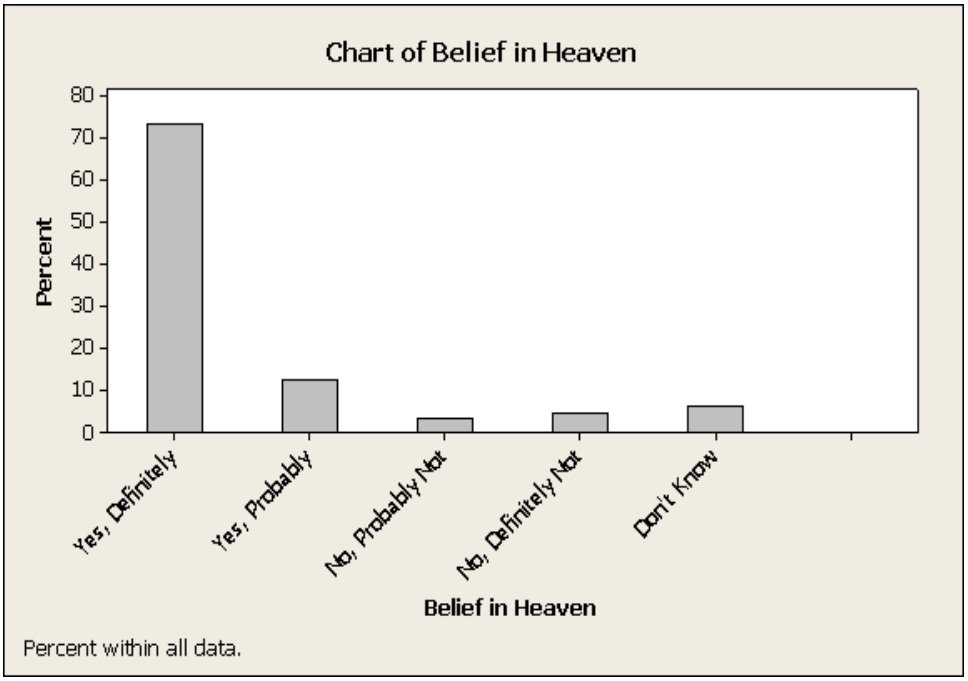
    (B) 7.66 hours.

    (C) 7.1 hours.

    (D) 10 hours.

    (E) Close to 0.

16. In Spring 2010, Math 1530 students at ETSU took a survey in which they were asked to choose one of the listed responses to the question "Do you believe in heaven?"



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

(A) The bar graph is skewed to the right.

(B) The bar graph is skewed to the left.

(C) It appears that a majority of students responding to the survey have a strong belief in heaven.

(D) The center of the graph is approximately "no, probably not."

(E) Exactly 73 students said they "yes, definitely" believe in heaven.

Use the following to answer the next 3 questions about gender and tattoos. The class survey asked for gender type and "Do you have any tattoos?" The distribution of counts is shown below in the table. Choose a student at random from this group.

Tabulated statistics: Gender, Tattoo

Rows: Gender Columns: Tattoo

	Yes	No	All
Female	195	569	764
Male	134	408	542
All	977	329	1306

Cell Contents: Count

17. The probability that the student is a female and has a tattoo is

(A) .1996 (B) .1493 (C) .2552 (D) .5850 (E) .7481

18. The conditional probability that the student was a female, given that the student has a tattoo, is about

(A) .1493 (B) .5850 (C) .1996 (D) .7481 (E) .2552

19. The conditional probability that the student has a tattoo, given that the student was a female, is about

(A) .2552 (B) .1996 (C) .7481 (D) .1493 (E) .5850

20. Which of these describes a simple random sample?

    (A) A high school marketing club needed to select 6 members from the club to go on a study trip to New York. Each club member put their name on a slip of paper, and all the slips of paper were put into a hat. The club sponsor then thoroughly stirred up the slips and picked out 6 names.

    (B) A statistics student went out and measured the arm lengths of several family members.

    (C) The website 'excitepoll.com' posted a survey asking what people think of the growing trend toward restrictive dress codes in American high schools. Respondents were invited to log in and click their answers.

    (D) A marketing class designs two videos advertising a Mercedes sports car. They test the videos by asking fellow students to view both (in random order) and say which makes them more likely to buy the car.

    (E) A study of dating among college students wanted a sample of 200 of the 9000 single male students on campus. The sample consisted of every 45th name from a list of the 9000 students.

21. Mike Huckabee's 2012 Republican presidential nomination campaign needs to know the opinions of American voters. To that end, the campaign workers plan to randomly select a group of 1000 likely voters and have them answer a carefully designed survey. Who are the individuals in this survey?

    (A) The survey.

    (B) The proportion of likely voters that plan on voting for Mike Huckabee.

    (C) Mike Huckabee's campaign staff.

    (D) The 1000 likely voters.

    (E) The opinions of the voters.

22. A group of disabled women ages 65 and over were tracked for several years. Within that group, those who had a vitamin B12 deficiency were twice as likely to suffer severe depression as those who did not. This study is an example of

    (A) an observational study.

    (B) a matched pairs experiment.

    (C) a completely randomized experiment.

    (D) a stratified random sample.

    (E) an uncontrolled study.

Use the following for the next 3 questions. Space Exploration Aptitude Tests (S.E.A.T.) for college students are known to have a roughly Normal distribution with standard deviation $\sigma = 12$ but unknown mean μ. Here are the S.E.A.T. scores for a simple random sample of college students:

35 45 25 49 39 30 25 30 29

23. Find the mean S.E.A.T. scores for these 9 students.

    (A) 34.1 (B) 30.0 (C) 25.6 (D) 35.0 (E) 31.0

24. Use this sample to calculate a 95% confidence interval for μ.

    (A) $(27.5, 40.7)$ (B) $(23.8, 44.4)$ (C) $(31.4, 36.8)$ (D) $(32.0, 36.0)$ (E) $(26.3, 41.9)$

25. Which of these is a valid interpretation of the above confidence interval?

    (A) There is a 95% probability that the true μ for SEAT scores of college students is within the 95% confidence interval.

    (B) Based on the sample data, we can be 95% confident the mean S.E.A.T. score μ for all college students will be within the calculated confidence interval.

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

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

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

Use the following for the next 3 questions. A survey in 2008 found that nearly half of all drivers aged 18-24 send text messages while they drive. Question 12 of our class survey asked "Do you text while driving?" The class survey revealed that 750 of the 1234 students who responded to this question said "Yes." Suppose we are interested in the proportion of all college students that text while driving a car.

26. Does our class survey give evidence to conclude that more than half of all college students text while driving? The hypotheses for a test to answer this question are

(A) $H_0 : p = .5, H_a : p < .5$ (C) $H_0 : \hat{p} = 750, H_a : \hat{p} > 750$

(B) $H_0 : p = .5, H_a : p > .5$ (D) $H_0 : \hat{p} = .61, H_a : \hat{p} > .61$

(E) $H_0 : \mu = 1234,\ H_a : \mu > 1234$

27. Find the p-value for the previous test.

(A) approximately 1 (B) greater than .05 (C) between .01 and .05. (D) approximately 0 (E) Unable to be determined.

28. Interpret your results relative to the survey conducted in 2008 and comment on any assumptions that are needed for your conclusions to be accurate.

(A) There is very strong evidence that more than half of college students text while driving. It is important that this student survey represents all college students. An SRS of college students is needed to make a strong conclusion.

(B) There is very weak evidence that more than half of college students text while driving.

(C) Do not reject H_0 since the $p-$value is too large. The assumptions should be okay since the sample size is large ($n = 1234$).

(D) Reject H_0 and conclude that the percentage of all college students that text while driving is 61%. It is important that the data follow the Normal distribution and there are no outliers.

(E) Reject H_0. There is no need to worry about the assumptions since the sample size is large ($n = 1234$).

29. A local grocery store received many customer complaints about the quantity of chips in 16-ounce bags of a particular brand. Wanting to assure its customers they were getting their money's worth, the store decided to test the hypotheses concerning the true average weight (in ounces) of chips in such a bag in the next shipment received from the supplier. If there is evidence in favor of the alternative hypothesis, the shipment will be refused and a complaint registered with the supplier. Bags of chips were selected from the next shipment and the weight of each selected bag was measured. The researcher for the supermarket chain stated that the data were statistically significant. What does "statistically significant" mean in this context?

(A) Perhaps the results are attributable to some confounding variable (e.g., color of the bag).

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

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

(D) The researcher found no evidence to suggest that the mean weight of the bag of chips is less than 16 ounces.

(E) The researcher found evidence to suggest that the true average weight is under 16 ounces.

Use the following for the next 4 questions. A group of college students believes that herbal tea has remarkable restorative powers. To test their belief, they visit a local nursing home, visiting with and talking to the residents while serving them herbal tea. After several weeks of this, the nursing home staff reports that the residents are healthier and more cheerful.

30. The subjects in this experiment are

(A) the college students. (C) the nursing home staff.

(B) the nursing home residents. (D) the brands of herbal tea.

(E) the overall health of the residents.

31. The factor (explanatory variable) in this experiment is

(A) the nursing home staff. (C) the nursing home residents.

(B) the health and cheerfulness of the nursing home residents. (D) the age of the students.

(E) the herbal tea.

32. The response variable in this experiment is

    (A) the college students.
    (B) the brands of herbal tea.
    (C) the several weeks of visits.
    (D) the herbal tea.
    (E) health and cheerfulness of the nursing home residents.

33. A Math 1530 student is skeptical about the significant improvement in the health of the nursing home residents due to the herbal tea because

    (A) the observed effect was so large that it would rarely occur by chance.
    (B) the response of the nursing home staff may not have any practical importance.
    (C) the experiment needs to be double-blinded.
    (D) the improvement might be due to the visits by the college students (lurking variable) and not by the herbal tea.
    (E) more than one brand of tea is needed.

34. Vigorous exercise helps people live several years longer (on average). Whether mild activities like slow walking extend life is not clear. Suppose that the added life expectancy from regular slow walking is just 2 months. A statistical test is more likely to find a significant increase in mean life if

    (A) it is based on a very small sample.
    (B) The size of the sample doesn't have any effect on the significance of the test.
    (C) the $P-$value is large.
    (D) it is based on a very large sample.
    (E) the mean of 2 months is different from 0.

35. The most important condition for sound conclusions from statistical inference is usually

    (A) that the distribution is exactly Normal.
    (B) that the data can be thought of as a random sample from the population of interest.
    (C) that the data contain no outliers.
    (D) that the sample size needs to be large.
    (E) that the margin of error in a confidence interval accounts for chance variation.

36. *What is significance good for?* Which of the following questions does a test of significance answer?

    (A) Is the observed effect due to chance?
    (B) Is the sample or experiment properly designed?
    (C) Is the observed effect important?
    (D) Is the margin of error large?
    (E) What is the chance that the null hypothesis is true?

37. The amount of money college students spend each semester on textbooks is normally distributed with a mean of \$195 and a standard deviation of \$20. Suppose you take a random sample of 100 college students from this population and calculate the mean amount $\bar{X}$ spent on textbooks in a semester. There would be a 68% chance that $\bar{X}$ would be between:

    (A) \$191 and \$199. (B) \$175 and \$215. (C) \$155 and \$235. (D) \$193 and \$197. (E) $\$195 \pm (.95)(20/\sqrt{100})$

38. "Do you feel confident or not confident that the food available at most grocery stores is safe to eat?" When a Gallup Poll asked this question, 87% of the sample said they were confident. Gallup announced the poll's margin of error for 95% confidence as ± 3 percentage points. Which of the following sources of error are included in this margin of error?

    (A) There is chance variation in the random selection of telephone numbers.
    (B) Gallup dialed landline telephone numbers at random and so missed all people without landline phones, including people whose only phone is a cell phone.
    (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 the grocery store food.
    (E) There will always be mistakes in conducting polls and the margin error tells us the percent of all these errors to expect.

Use the following for the next 2 questions. The Census Bureau provides estimates of numbers of people in the United States classified in various ways. We will look at college students. The following table gives us data to examine the relation between age and full-time or part-time status. The Minitab output shows the two-way table and related information. The counts in the table are expressed as thousands of U.S. college students.

Tabulated statistics: Age, Status
Using frequencies in Counts

Rows: Age Columns: Status

	Full-time	Part-time	All
15 – 19	3388	389	3777
	2556	1221	
20 – 24	5238	1164	6402
	4333	2069	
25 – 34	1703	1699	3402
	2302	1100	
35 and over	762	2045	2807
	1900	907	
All	11091	5297	16388

Cell Contents: Count
Expected Count

Pearson Chi-Square = 4013.424, DF = 3, P-Value = 0.000

39. The null hypothesis for the chi-square test for this two-way table is

(A) The observed counts will be different from the expected counts.
(B) The number of full-time college students is the same as the number of part-time college students.
(C) There is no relationship between the age groups and college status.
(D) The number of students in each of the age groups is the same.
(E) There is some relationship between and the age groups and college status.

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

(A) There doesn't appear to be a relationship between the age groups and college status.
(B) The number of full-time and part-time college students are different.
(C) The number of students in each of the age groups is different.
(D) There does appear to be a relationship between the age groups and college status.
(E) Do not reject the null hypothesis.

Use the following for the next 2 questions. What is normal Internet use? Some previous studies found that college students averaged 8 hours or so per week on the Internet. With more people owning cell phones we think that Internet use has decreased maybe because of texting/twittering has increased. We want to test this hypothesis with some sample data from Math 1530 students. A random sample of 25 Math 1530 students from the Fall 2009 semester were asked how many hours per week they spent on the Internet. Their responses had a sample mean of 2.368 hours with standard deviation 1.313 hours. What does this sample suggest about college students mean Internet use? Only one of the following four Minitab outputs is correct.

```
Output 1:
One-Sample T

Test of mu = 8 vs not = 8

 N   Mean  StDev  SE Mean      95% CI           T      P
25  2.368  1.313    0.263  (1.826, 2.910)  -21.45  0.000
-------------------------
Output 2:
One-Sample T
Test of mu = 8 vs < 8
                            95% Upper
 N   Mean  StDev  SE Mean      Bound       T      P
25  2.368  1.313    0.263      2.817  -21.45  0.000
-------------------------
Output 3:
One-Sample T
Test of mu = 8 vs > 8
                            99% Lower
 N   Mean  StDev  SE Mean      Bound       T      P
25  2.368  1.313    0.263      1.714  -21.45  1.000
-------------------------
Output 4:
One-Sample Z

Test of mu = 8 vs < 8
The assumed standard deviation = 1.313

                    95% Upper
 N   Mean  SE Mean      Bound       Z      P
25  2.368    0.263      2.800  -21.45  0.000
------------------------
```

41. State the appropriate null and alternate hypotheses.

    (A) $H_0 : \mu = 8$ $H_a : \mu \neq 8$

    (B) $H_0 : \mu = 8$ $H_a : \mu < 8$

    (C) $H_0 : \mu = 8$ $H_a : \mu > 8$

    (D) $H_0 : \mu = 2.368$ $H_a : \mu < 2.368$

    (E) The 95% confidence interval is $(1.826, 2.910)$.

42. What can we conclude from the hypothesis test results?

    (A) The sample proves that the null hypothesis is false. College students have always spent a mean of 8 hours per week on the Internet.

    (B) The P-value is close, so the hypothesis test is ambiguous. A larger sample must be used to reach any valid conclusion.

    (C) Without a significance level, it is impossible to draw any conclusion from any hypothesis test.

    (D) The hypothesis test result proves that the null hypothesis is true. It is certain that all college students spend at least 8 hours per week on the Internet.

    (E) The very low P-value is strong evidence against the null hypothesis. The mean weekly hours of Internet use among college students has dropped down below the historical value of 8.