Use the following information for questions 1, 2, 3, and 4

The mean height of American women in their early twenties is about 64.5 inches and the standard deviation is about 2.5 inches. The mean height of men the same age is about 68.5 inches, with standard deviation about 2.7 inches. The correlation (r) between heights of husbands and wives is about 0.5.

1. You are interested in predicting the height of a young man from the height of his wife. What is the slope of the regression line you would use in this case?
   a. 32.78  b. 0.463  c. 33.67  d. 0.540

2. A young woman is 67 inches tall. Compute the predicted height of her husband using the least squares regression equation appropriate for this case.
   a. 64.73  b. 69.85  c. 68.50  d. 70.25

3. If you are interested in predicting the height of a young woman from the height of her husband, what is the slope of the regression line you would use in this case?
   a. 32.78  b. 0.463  c. 33.67  d. 0.540

4. A young man is 69 inches tall. Compute the predicted height of his wife using the least squares regression equation appropriate for this case.
   a. 64.73  b. 69.85  c. 68.50  d. 70.25

5. What is the best characterization of points A, B, C, and D in the following scatterplot
a. A and C are influential observations, B and D are outliers
b. D and C are influential observations, A and B are outliers
c. A and D are outliers, B and C are influential observations
d. A and B are outliers, C and D are influential observations.

6. According to government data, 20% of employed women have never been married. If 150 employed women are selected at random, what is the approximate probability that at least 30 have never been married?
   a. 0.1876  b. 0.1643  c. 0.3519  d. 0.5000

7. We would like to know the true mean μ of all possible values of some quantitative variable X in a particular population. If all the values of X in this population are normally distributed, then the distribution of the sample mean \( \bar{X} \) is:
   a. not normally distributed    b. always normally distributed
c. normally distributed if the sample size n is 30 or larger
d. normally distributed if n is less than 30

8. We would like to know the true mean μ of all possible values of some quantitative variable X in a particular population. If all the values of X in this population are not normally distributed, then the distribution of the sample mean \( \bar{X} \) is:
   a. never normally distributed    b. always normally distributed
c. approximately normally distributed if the sample size n is 30 or larger
d. approximately normally distributed if n is less than 30

9. You want to compute a 95% confidence interval for a population mean. Assume that the population standard deviation is known to be 10 and the sample size is 50. The value of \( z^* \) to be used in this calculation is
   a. 1.645  b. 2.009  c. 1.960  d. .8289  e. .8352

10. A sample of 25 seniors from a large metropolitan area school district had a mean Math SAT score of 450. Suppose we know that the standard deviation of the population of Math SAT scores for seniors in the district is 100. Assume the population of Math SAT scores for seniors in the district is approximately normally distributed. What is a 90% confidence interval for μ, the true mean Math SAT score of all seniors in the district?
    a. 450 ± 7.84  b. 450 ± 32.9  c. 450 ± 39.2  d. 450 ± 1.96
11. A sample of 25 seniors from a large metropolitan area school district had a mean Math SAT score of 450. Suppose we know that the standard deviation of the population of Math SAT scores for seniors in the district is 100. Assume the population of Math SAT scores for seniors in the district is approximately normally distributed. A 90% confidence interval for the mean Math SAT score \( \mu \) for the population of seniors is used. Which of the following would produce a confidence interval with a smaller margin of error?

a. using a confidence level of 95%  
b. using a confidence level of 99%  
c. using a sample of 100 seniors  
d. using a sample of only 10 seniors

12. A sample of 100 postal employees found that the average time these employees had worked for the postal service was 8 years. Assume that we know that the standard deviation of the population of times postal service employees have spent with the service is 5 years. A 95% confidence interval for the true mean time \( \mu \) postal service employees have spent with the postal service is what?

a. 8 ± 0.98  
b. 5 ± 1.32  
c. 5 ± 1.57  
d. 8 ± 0.82

13. A medical researcher is curious if the mean weight \( \mu \) of all college students has changed since 1980. Suppose in 1980 the mean weight was 150 lb. The researcher plans to examine the weights of a sample of college students taken in 1995 to see if the mean weight of all college students has changed in the last 15 years. What is the alternative hypothesis the researcher wishes to test?

a. Ha : \( \mu = 150 \) lb.  
b. Ha : \( \mu \neq 150 \) lb.  
c. Ha : \( \mu > 150 \) lb.  
d. Ha : \( \mu < 150 \) lb.

14. A company advertises that by using its program, high school students can increase their Verbal SAT scores by at least 50 points. You are skeptical and decide to test the company’s claim. What are the null and alternative hypotheses for your test?

a. Ho: \( \mu = 50 \)  
b. Ho: \( \mu = 50 \)  
c. Ho: \( \mu > 50 \)  
d. Ho: \( \mu = 50 \)

Ha: \( \mu < 50 \)  
Ha: \( \mu > 50 \)  
Ha: \( \mu < 50 \)  
Ha: \( \mu \neq 50 \)

15. Which of the following would be strong evidence against the null hypothesis \( H_0 \) in a hypothesis test?

a. a very large P-value  
b. a very small P-value  
c. a large sample size  
d. a small sample size
16. In a test of hypotheses, we say that the data are statistically significant at level $\alpha$ (alpha) if:
   a. $\alpha$ is very small
   b. the P-value is larger than $\alpha$
   c. the P-value is smaller than $\alpha$
   d. the data indicate that an important and meaningful effect has been detected

17. A significance test gives a P-value of .04. From this we can:
   a. reject $H_o$ with $\alpha = .01$
   b. reject $H_o$ with $\alpha = .05$
   c. say that the probability that $H_o$ is false is .04
   d. say that the probability that $H_o$ is true is .04

**Use the following information for questions 18, 19, & 20**

“Jumbo” bags of a certain brand of potato chips are supposed to contain 32 ounces (2 lb) of chips. There is some variation from bag to bag because the filling machinery is not perfectly precise. The distribution of contents is normal with mean $\mu$ and standard deviation $\sigma = 1$ ounce. An inspector who suspects that the manufacturer is under filling the bags measures the contents of 4 bags. The results (in oz) are:

31, 33, 30, & 30.

18. Is this convincing evidence that the mean contents of “jumbo” bags is less than the advertised 32 ounces? Using $\bar{X}$ as the test statistic, what is the P-value?

   a. 31     b. 0.0228     c. 0.1587     d. 0.0456     e. 0.9772

19. At what level of significance do the above data provide enough evidence to reject the null hypothesis?

   a. 1%     b. 2%     c. 5%     d. 0.05%     e. none of these

20. In the above data, what are the hypotheses being tested?

   a. $H_o : \mu = 32$ and $H_a : \mu \neq 32$     b. $H_o : \mu = 32$ and $H_a : \mu > 32$
   c. $H_o : \mu = 32$ and $H_a : \mu < 32$     d. $H_o : \mu \neq 32$ and $H_a : \mu = 32$
21. I wish to test the hypotheses $H_0: \mu = \mu_0$ and $H_a: \mu \neq \mu_0$ based on an SRS of size $n$ from a normal population with unknown mean $\mu$ and known standard deviation $\sigma$. I should reject $H_0$ at significance level $\alpha = .05$ if the value of the $z$-statistic $z$ satisfies what?
   a. $z \neq 1.96$
   b. $z > 1.645$ or $z < -1.645$
   c. $z < -1.96$ or $z > 1.96$
   d. $z \neq 1.645$

22. A particular model car advertises that it gets 30 miles per gallon in city driving. A consumer group wishes to test if this is true or if the gas mileage is something different than 30 miles per gallon. They purchase a sample of 4 cars from this model, and find the gas mileage of each after 15,000 miles of city driving. The mean gas mileage for these 4 cars is found to be 28.8 miles per gallon. Assume the distribution of city gas mileage for the population of cars is normal with mean $\mu$ and standard deviation $\sigma = 2$ miles per gallon. What is the $P$-value of the test statistic for this test?
   a. 0.1151
   b. 0.2302
   c. 10%
   d. 5%

23. In a test of hypothesis, the null hypothesis is that the population mean is equal to 60 and the alternative hypothesis is that the population mean is not equal to 60. A sample of size 36 from this population produced a sample mean of 63. If we can assume that the population standard deviation equals 6.3, then the $p$-value for this test is approximately:
   a. 0.0042
   b. 0.0347
   c. 0.0952
   d. 0.021

24. In a test of hypothesis, the null hypothesis is that the population mean is equal to 54 and the alternative hypothesis is that the population mean is greater than 54. A sample of 24 elements selected from this population produced a sample mean $\bar{x} = 61$ and a sample standard deviation $s = 6.3$. The significance level $\alpha = 2.5\%$. The critical value $t^*$ for the $t$-statistic is:
   a. 2.500
   b. -2.093
   c. 2.069
   d. 2.064

25. In a test of hypothesis, the null hypothesis is that the population mean is equal to 90 and the alternative hypothesis is that the population mean is not equal to 90. A sample of 16 elements selected from this population produced a mean of 86.75 and a standard deviation of 12.54. The significance level is 2%. The critical values $t^*$ for the $t$-statistic are:
   a. -2.602 and 2.602
   b. -2.483 and 2.583
   c. -2.131 and 2.131
   d. -1.341 and 1.341
26. A random sample of 25 tourists who visited Hawaii last summer spent an average of $1420 on this trip with a standard deviation of $285. Assume that the money spent by all tourists who visit Hawaii has an approximate normal distribution. The 95% confidence interval for the mean money spent by all tourists who visit Hawaii is:

a. $1281 to $1559  
   b. $1372 to $1468  
   c. $1302 to $1538  
   d. $1347 to $1493

27. The value of t for 19 degrees of freedom and .01 area in the right tail is:

a. 2.539  
   b. 2.861  
   c. 1.328  
   d. -2.539

Use the following information for questions 28, 29, & 30

The national mortgage rate for 30 year fixed rate mortgages is 9.2%. A Realtor in a large Midwestern city believes that local mortgage rates are lower than the national average. Let \( \mu \) represent the mean 30 year fixed mortgage rate for all lending institutions in the city and assume the distribution of rates in this population of lending institutions is approximately normal. The Realtor takes a survey of 25 local lending institutions and finds the sample mean 30 year fixed mortgage rate is 9.0% with standard deviation \( s=0.25\% \). Are these data evidence that the local rates are below the national rate? To determine this we test the hypotheses \( H_0: \mu = 9.2\% \) vs. \( H_a: \mu < 9.2\% \) using the one-sample t test.

28. What are the appropriate degrees of freedom for this test?

a. 25  
   b. 24  
   c. 30  
   d. 29

29. Using the above data, what is the P-value for the one-sample t test?

a. between .10 and .05  
   b. larger than .10  
   c. less than .01  
   d. between .05 and .01

30. Using the above data, a 95% confidence interval for \( \mu \) is what?

a. 9.0% \( \pm \) .10%  
   b. 9.0% \( \pm \) .05%  
   c. 9.0% \( \pm \) .52%  
   d. 9.0% \( \pm \) .25%
Use the following information for questions 31, 32, & 33

A researcher wished to compare the effect of two stepping heights (low and high) on heart rate in a step-aerobics workout. A collection of 50 adult volunteers was randomly divided into two groups of 25 subjects each. Group 1 did a standard step-aerobics workout at the low height. The mean heart rate at the end of the workout for the subjects in group 1 was $\bar{x}_1 = 90$ beats per minute with a standard deviation $s_1 = 9$ beats per minute. Group 2 did the same workout but at the high step height. The mean heart rate at the end of the workout for the subjects in group 2 was $\bar{x}_2 = 92$ beats per minute with a standard deviation $s_2 = 12$ beats per minute. Assume the two groups are independent and the data are approximately normal. Let $\mu_1$ and $\mu_2$ represent the true mean heart rates using the low or high step heights, respectively.

31. What is a 95% confidence interval for the difference $\mu_1 - \mu_2$? (Use the conservative value for the degrees of freedom.)
   
   a. $-5 \pm 6.19$ beats per minute  
   b. $5 \pm 6.19$ beats per minute  
   c. $-5 \pm 5.13$ beats per minute  
   d. $5 \pm 5.13$ beats per minute

32. Using the above data, suppose the researcher had wished to test the hypotheses: $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 < \mu_2$. What is the P-value for the test?
   
   a. larger than .10  
   b. between .10 and .05  
   c. between .05 and .01  
   d. below .01

33. At what level of significance do the above data provide enough evidence to reject the null hypothesis
   
   a. lower than 1%  
   b. 1%  
   c. 5%  
   d. 10%

34. In a sample of 500 items produced by a machine, 7% are found to be defective. The 95% confidence interval for the proportion of defective items $p$ in all items produced by this machine is:
   
   a. 0.057 to 0.083  
   b. 0.061 to 0.079  
   c. 0.032 to 0.108  
   d. 0.048 to 0.092
Use the following information for questions 35, 36, & 37

A newspaper conducted a statewide survey concerning the 1994 race for state senator. The newspaper took a random sample of 1200 registered voters and found that 620 would vote for the Republican candidate. Let \( p \) represent the proportion of registered voters in the state that would vote for the Republican candidate.

35. Using the above data, a 90\% confidence interval for \( p \) is what?
   a. \( 0.517 \pm 0.024 \)  b. \( 0.517 \pm 0.014 \)
   c. \( 0.517 \pm 0.249 \)  d. \( 0.517 \pm 0.028 \)

36. Using the above data, suppose you wished to see if the Republican candidate had a “clear” majority. To do this you test the hypotheses: \( H_0: p = 0.50 \) vs. \( H_a: p > 0.50 \)
What is the P-value of your test?
   a. 0.88  b. 0.95  c. 0.05  d. 0.12

37. At what level of significance do these data provide enough evidence to support that the Republican candidate has a “clear” majority?
   a. 5\%  b. 10\%  c. 15\%  d. 1\%  e. none of these

Use the following information for questions 38, 39, & 40

An SRS of 100 flights of airline X showed that 64 were on time. An SRS of 100 flights of airline Y showed that 80 were on time. Let \( p_1 \) be the proportion of all flights that are on time for airline X and \( p_2 \) be the proportion of all flights that are on time for airline Y.

38. What is a 95\% confidence interval for the difference \( p_1 - p_2 \)?
   a. \(-0.16 \pm 0.122 \)  b. \(-0.16 \pm 0.062 \)
   c. \(0.16 \pm 0.062 \)  d. \(-0.16 \pm 0.103 \)
39. Using the above data, is there evidence of difference in the on-time rate for the two airlines? To determine this, you test the hypotheses: $H_0 : p_1 = p_2$ vs. $H_a : p_1 \neq p_2$. What is the P-value of your test?

   a. between .05 and .01        b. between .10 and .05
   c. below .001                d. between .01 and .001

40. At what level of significance do the above data provide enough evidence to reject the null hypothesis?

   a. 10%      b. 5%      c. 2%      d. all of the above levels

Questions 41 - 47 are from chapter 9

Use the following information for questions 41, 42, & 43

A study was performed to examine the personal goals of children in grades 4, 5, and 6 from schools in Georgia. The students received a questionnaire regarding achieving personal goals. They were asked what they would most like to do at school: make good grades, be good at sports, or be popular. Here are the results:

<table>
<thead>
<tr>
<th>Make good grades</th>
<th>Be popular</th>
<th>Be Good in sports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>96</td>
<td>32</td>
</tr>
<tr>
<td>Girls</td>
<td>295</td>
<td>45</td>
</tr>
</tbody>
</table>

41. Using the above data, suppose we wish to test the null hypothesis that there are no differences among the proportion of boys and the proportion of girls choosing each of the three personal goals. Under the null hypothesis, what is the expected number of boys that would select “make good grades”?

   a. 196      b. 96      c. 144      d. 222

42. Using the above data, the appropriate degrees of freedom for the chi-square statistic is which of the following?

   a. 3        b. 2        c. 4        d. 1
43. What is an approximate value of the chi-square statistic computed from the above data?

   a. 90   b. 25   c. 12   d. 43

44. At $\alpha = 0.01$, what is the critical value ($\chi^*$) of the chi-square statistic for the above data?

   a. 89.9   b. 16.81   c. 9.21   d. 13.82

45. Using the above data, which of the following is the best conclusion?

   a. There is no evidence of any relation between gender and personal goals children set for themselves.
   b. There appears to be evidence of an association between gender and personal goals children set for themselves.
   c. Personal goals children set for themselves is the result of parental guidance, the test is inconclusive.
   d. Because girls are over represented in the sample the results cannot be trusted. A new study is needed. The number of boys and girls in the sample should be approximately equal.

46. A sample in which the population is first divided into groups of similar individuals and then a separate simple random sample is selected from each group and combined to form the full sample is called what?

   a. a probability sample   b. convenience sample
   c. a simple random sample   d. a stratified random sample

47. When are the results of an experiment said to be statistically significant?

   a. They are important to statisticians, regardless of their importance to the investigators.
   b. The observed effect is too large to attribute plausibly to chance.
   c. They support the findings of previous, similar studies.
   d. Both researchers and statisticians agree the results are meaningful and important.
<table>
<thead>
<tr>
<th>Answers</th>
</tr>
</thead>
</table>