

6. The z value associated with a two-sided 90% confidence interval is ____.
- A) 1.28 B) 1.645 C) 1.96 D) 2.33
7. A study involves 673 randomly selected deaths, with 34 of them caused by accidents. A 98% confidence interval for the percentage of all deaths that are caused by accidents is:
- A) [0.034, 0.067] B) [0.037, 0.065] C) [0.038, 0.064] D) [0.031, 0.070]
8. Suppose a researcher is interested in understanding the variation in the price of store brand milk. A random sample of 36 grocery stores selected from a population and the mean price of store brand milk is calculated. The sample mean is \$3.13 with a standard deviation of \$0.23. Construct a 99% confidence interval to estimate the population mean.
- A) [\$3.03, \$3.23] B) [\$3.12, \$3.14] C) [\$2.90, \$3.36] D) [\$3.05, \$3.21]
9. A researcher wishes to estimate the proportion of fish in a certain lake that is inedible due to pollution of the lake. How large a sample should be tested in order to be 99% confident that the true proportion of inedible fish is estimated to within 6%?
- A) 188 B) 461 C) 376 D) 267
10. In a survey of 1,000 television viewers, 40% said they watch network news programs. For a 90% confidence level, the margin of error for this estimate is 2.5%. If we want to be 99% confident, how will the margin of error change?
- A) Since more confidence requires a narrow interval, the margin of error will be larger.
 B) Since more confidence requires a narrow interval, the margin of error will be smaller.
 C) Since more confidence requires a wider interval, the margin of error will be larger.
 D) Since more confidence requires a wider interval, the margin of error will be smaller.
11. Whenever the population standard deviation is unknown and the population has a normal or normal distribution, which distribution is used to construct a confidence interval for the population mean?
- A) z distribution. B) standard distribution C) F distribution D) t distribution
12. A random variable X follows a normal distribution with standard deviation 8. We take a random sample of 25 individuals from the population and we calculate a confidence interval for μ to be (36.7136, 43.2864). What is the confidence level of this interval?
- A) 90% B) 95% C) 96% D) 99%

The next **two** questions (23 and 26) refer to the following:

A researcher wants to conduct a before/after study on 11 subjects to determine if a new cholesterol medication results in higher HDL cholesterol readings. Scores are obtained on the subjects both before and after taking the medication. After subtracting the after scores from the before scores, the average difference is computed to be -2.40 with a sample standard deviation of 1.21. Assume that the differences are normally distributed in the population. The degrees of freedom for this test are _____.

23. The Null hypothesis to determine whether the new cholesterol medication results in higher HDL cholesterol readings are:

- A) $H_0 : \mu_D \leq 0$ B) $H_0 : \mu_D > 0$ C) $H_0 : \mu_D \neq 0$ D) $H_0 : \mu_D \geq 0$

24. The degrees of freedom for this test are _____.

- A) 11 B) 10 C) 9 D) 20

25. The observed t value (t_{calc}) for this test is _____.

- A) -21.82 B) -6.58 C) -2.40 D) -2.33

26. The table t value (t_{crit}) for this test is _____.

- A) -1.812 B) -2.228 C) -2.086 D) -2.262

27. When testing the difference between two population means, the variances are pooled when

- A) the population standard deviations are known and equal
B) the population standard deviations are assumed unequal and unknown
C) the population standard deviations are assumed equal but unknown
D) None of the above

The next **two** questions (28 and 29) refer to the following:

A two-sided significance test for two population proportions is to be performed. The hypotheses for the test are $H_0 : p_1 - p_2 = 0$ vs. $H_A : p_1 - p_2 \neq 0$. Given that $n_1 = 200$; $p_1 = 0.10$; $n_2 = 200$ $p_2 = 0.08$

28. The value of the test statistic is:

- A) $T = -0.70$ B) $T = 0.70$ C) $Z = 0.70$ D) $Z = -0.70$

29. The P-value for the significance test is:

- A) 0.2119 B) 0.4238 C) 0.484 D) 0.242

The next **four** questions (30 and 33) refer to the following:

Medical Wonders is a specialized interior design company focused on healing artwork. The CEO claims that artwork has healing effects for patients staying in a hospital, as measured by reduced length of stay. The CEO is interested in determining the effect of three different pieces of healing artwork on children. She chooses three paintings (a horse photo, a bright abstract, and a muted beach scene) and randomly assigns six hospital rooms to each painting. Analysis of her data yielded the following partial ANOVA table. Complete the partial ANOVA table and use it to answer question 30 to 33.

Source of Variation	SS	df	MS	F _{calc}
Treatment		2	16738.1	
Error		15		
Total	60022.37			

30. The CEO's null hypothesis is _____.
- A) $\mu_1 = \mu_2 = \mu_3$ B) $\mu_1 \neq \mu_2 \neq \mu_3$ C) $\mu_1 \geq \mu_2 \geq \mu_3$ D) $\mu_1 \leq \mu_2 \leq \mu_3$
31. The value of the test statistic (F_{calc}) for this test is:
- A) 13.68 B) 5.39 C) 19.45 D) 9.46
32. Using $\alpha = 0.05$, the critical F value is _____.
- A) 13.68 B) 19.43 C) 3.59 D) 3.68
33. What is the appropriate decision for this test?
- A) Reject H_0 : All means are different.
 B) Reject H_0 : At least one mean is different
 C) Do not reject H_0 : The mean are equal
 D) Do not reject H_0 : Can not conclude.
34. What type of data do you need for a chi-square test of independence?
- A) interval B) categorical C) parametric D) scale
35. The degree of freedom for a chi-square test statistic when testing for independence in a contingency table with 4 rows and 3 columns would be:
- A) 9 B) 6 C) 7 D) 5

36. In a chi-square test of independence between gender and opinion on tax cut, the null hypothesis is rejected. The proper conclusion is that:

- A) Gender and opinion are independent of each other
- B) Gender and opinion are related to each other
- C) Knowing a person's gender gives you no clue to his or her opinion on tax cut
- D) None of the above

The next **three** questions (37 and 40) refer to the following:

A variable contains five categories. It is expected that data are uniformly distributed across these five categories. To test this, a sample of observed data is gathered on this variable resulting in frequencies of 27, 30, 29, 21, and 24.

37. Using $\alpha = .01$, the degrees of freedom for this test are _____.

- A) 1 B) 2 C) 3 D) 4

38. Using $\alpha = .01$, the critical value of chi-square is _____.

- A) 7.78 B) 15.09 C) 9.24 D) 13.28

39. Using $\alpha = .01$, the observed value of chi-square is _____.

- A) 17.81 B) 13.28 C) 9.82 D) 2.09

40. Using $\alpha = .01$, the appropriate decision is _____.

- A) reject the null hypothesis that the observed distribution is uniform
- B) reject the null hypothesis that the observed distribution is not uniform
- C) do not reject the null hypothesis that the observed distribution is uniform
- D) do not reject the null hypothesis that the observed distribution is not uniform

41. The sampling distribution for a goodness of fit test is:

- A) the Poisson distribution
- B) the normal distribution
- C) the t distribution
- D) the chi-square distribution

42. Contingency tables should not be used with expected cell frequencies _____.

- A) less than the number of rows
- B) less than the number of columns
- C) less than 5
- D) less than 30

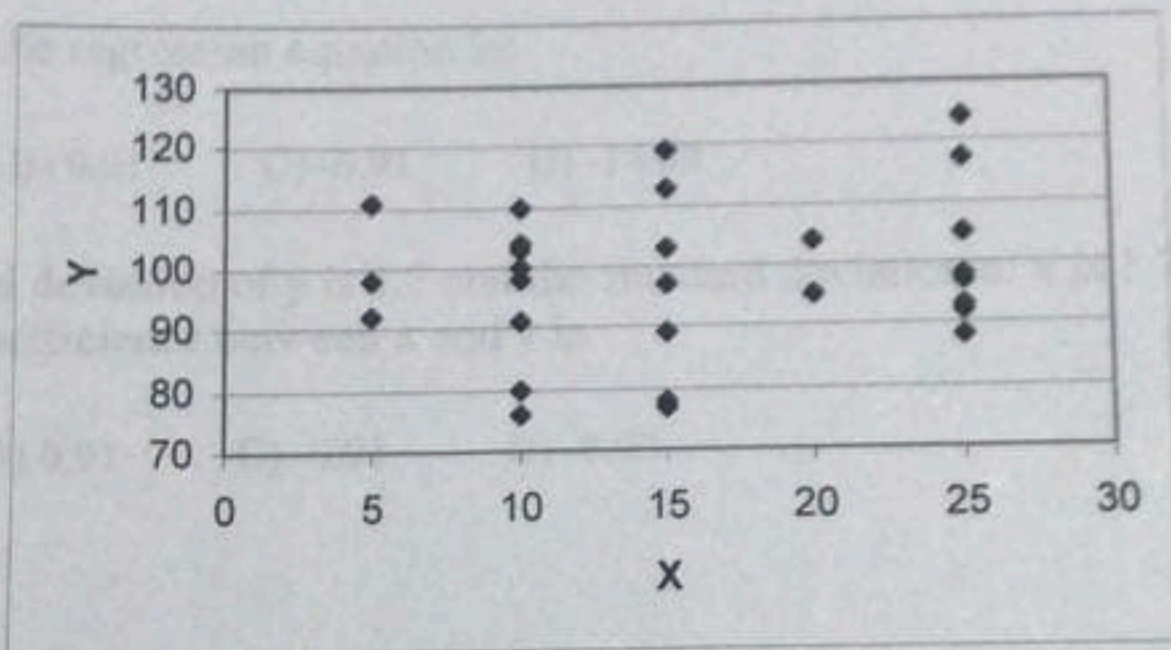
43. In a simple linear regression (where y is the dependent variable and x is the independent variable) if the y intercept is positive, then:

- A) there is a positive correlation between x and y
- B) there is a negative correlation between x and y
- C) there is no correlation between x and y
- D) None of the above

44. The value of a correlation is reported by a researcher to be $r = -0.5$. Which of the following statements is correct?:

- A) The x -variable explains 25% of the variability in the y -variable.
- B) The x -variable explains -25% of the variability in the y -variable.
- C) The x -variable explains 50% of the variability in the y -variable.
- D) The x -variable explains -50% of the variability in the y -variable.

45. According to the following graph, X and Y have _____.



- A) strong negative correlation.
- B) virtually no correlation.
- C) strong positive correlation.
- D) moderate negative correlation

46. If there is a perfect negative correlation between two quantitative variables, then

- A) $r = 0$
- B) $r = -1$
- C) $r = +1$
- D) $r \leq 0.5$

The next **four** questions (47 and 50) refer to the following:

A clinical psychologist wants to determine the regression equation of the number of weeks spent in a therapy hospital on number of seizures per week. He estimates the regression equation to be $\hat{y} = 14.09 - 0.91x$. This is based on a sample size of 50 patients. Using this information answer the following four questions:

47. The dependent variable is:

- A) number of seizures per week
- B) number of weeks spent in a therapy hospital
- C) therapy hospital
- D) none of the

48. How many seizures per week would you predict for a person who has been at the therapy hospital for ten weeks?

- A) 4.49
- B) 4.99
- C) 14.09
- D) 13.91

49. The slope of the regression equation is:

- A) 14.09
- B) 0.91
- C) -0.91
- D) -14.09

50. If the standard deviation of y is 1.7 and the standard deviation of x is 1.74, then the correlation coefficient r between x and y is

- A) 0.93
- B) 0.91
- C) -0.91
- D) -0.93

Part 2 – Long Answer (Show your work)

[30 marks]

Question # 1 [10 marks]

Chapter 10 - 1 Comparing the means of two independent populations with σ_1 and σ_2 unknown

The amount of a certain element in blood is different for men and women. A sample of 75 male blood donors were tested and it was found that their blood contained a concentration mean of 28 parts per million with a standard deviation of 14.1 ppm of this element. A smaller sample of 50 female blood donors whose blood was also tested found that their blood contained a concentration mean of 33 ppm with a standard deviation of 9.5 ppm. Does this data support the claim that the population means of concentrations of the element are the same for men and women? Population standard deviations are unknown but assumed to be unequal. Use $\alpha = 0.01$ and 44 degrees of freedom.

Question # 2 [10 marks]

Sam Hill, Director of Media Research, is analyzing subscribers to the Travel World magazine. He wonders whether subscriptions are influenced by the highest degree earned by the head of household. His staff prepared the following contingency table from a random sample of 300 households.

Chapter 11

		Head of Household Classification		
		Associate	Bachelor	Master/ PhD
Subscribers	Yes	10	90	60
	No	60	60	20

Test if subscriptions and head of household classification are dependent at 5% level of significance.

Question # 3 [10 marks]

Chapter 12

A specialist in hospital administration stated that the number of full-time employees (FTEs) in a hospital can be estimated by counting the number of beds in the hospital (a common measure of hospital size). We decide to develop a regression model in an attempt to predict the number of FTEs of a hospital by the number of beds. The following table summarizes a portion of the regression results:

Regression Statistics	
Multiple R	0.9415062511
R Square	0.88643402
Adjusted R Square	0.875077422
Standard Error	15.64907915
Observations	12

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	30.91246961	13.25422291	2.332273	0.041888	1.380220594	60.44471862
Number of Beds	2.231503994	0.252579783	8.834848	4.89E-06	1.668721166	2.794286823

a) Name the independent and dependent variables [2 mark]

b) What proportion of the variation in the dependent variable is explained by the independent variable [2 marks]

c) Determine the least squares regression equation and use it to predict FTEs if the specialist expects 95 beds in the hospital. [3 marks]

d) Can you conclude that there is a positive linear relationship between FTEs and the number of beds in the hospital? Use $\alpha = .05$ to test the appropriate hypotheses. [5 marks]

19. An art collector has bought a painting for \$15,000 and is trying to decide whether to buy insurance for the painting for \$2,000. In making this decision, he is essentially testing the hypothesis

H_0 : The painting will not be stolen vs. H_1 : The painting will be stolen

Based on the his knowledge and experience he determines a P-value of 0.15. Using a 10% level of significance, we conclude that:

- A) the probability that the painting will be stolen is 0.15.
- B) the probability that the painting will not be stolen is 0.15.
- C) there is sufficient evidence that the painting will be stolen.
- D) there is sufficient evidence that the painting will not be stolen.

The next **two** questions (20 and 21) refer to the following:

In the 2011 Canadian federal elections, the Liberal Party received 18.9% of all votes cast. We would like to conduct a hypothesis test to determine whether there is evidence that the party's support has increased since the election. In a random sample of 400 voters taken, 82 of them say they support the party.

20. What are the hypotheses for the appropriate test?

- A) $H_0: p \leq 0.189$ vs. $H_A: p > 0.189$
- B) $H_0: p < 0.189$ vs. $H_A: p \geq 0.189$
- C) $H_0: p \leq 0.21$ vs. $H_A: p > 0.21$
- D) $H_0: p < 0.21$ vs. $H_A: p \geq 0.21$

21. The P-value of the appropriate test is:

- A) 0.1038
- B) 0.1251
- C) 0.1443
- D) 0.1635

22. A university wants to increase its retention rate of 4% for graduating students from the previous year. After implementing several new programs during the last two years, the university re-evaluated its retention rate using a random sample of 352 students and found the retention rate at 5%. Test if the retention rate did increase using $\alpha = 0.05$.

- A) $H_0: p \leq 0.04$; $H_A: p > 0.04$; $z_{calc} = 0.96$; P-value = 0.1685.
- B) $H_0: p \geq 0.04$; $H_A: p < 0.04$; $z_{calc} = 1.07$; P-value = 0.8577.
- C) $H_0: p \leq 0.04$; $H_A: p > 0.04$; $z_{calc} = -1.07$; P-value = 0.8577.
- D) $H_0: p \leq 0.04$; $H_A: p > 0.04$; $z_{calc} = -1.07$; P-value = 0.1423.

13. A **Type I** error occurs when we _____.

- A) Reject the null hypothesis when it is actually true
- B) Reject the null hypothesis when it is actually false
- C) Do not reject the null hypothesis when it is actually true
- D) Do not reject the null hypothesis when it is actually false

14. If the chosen significant level is 5%, then there is a.....

- A) 10% probability of rejecting a false null hypothesis.
- B) 10% probability of rejecting a true null hypothesis.
- C) 5% probability of rejecting a false null hypothesis.
- D) 5% probability of rejecting a true null hypothesis.

15. If the p-value for a hypothesis test is 0.0271, which would be a correct decision?

- A) Do not reject the null hypothesis when alpha is 0.10.
- B) Do not reject the null hypothesis when alpha is 0.05.
- C) Reject the null hypothesis when alpha is 0.025.
- D) Reject the null hypothesis when alpha is 0.10.

The next **two** questions (16 and 18) refer to the following:

Carbon monoxide (CO) is a colourless and odourless gas. Even at low levels of exposure, CO can cause serious health problems. A home is considered safe if the mean CO concentration is 5.2 parts per million (ppm) or lower. We take a random sample of 10 readings in various locations in a home and find the mean CO concentration to be 5.5 ppm with a standard deviation of 0.4 ppm. We would like to conduct a hypothesis test to determine whether there is significant evidence that the home is **unsafe**.

16. What are the hypotheses for the appropriate test?

- A) $H_0: \mu < 5.5$ vs. $H_A: \mu \geq 5.5$
- B) $H_0: \mu \geq 5.5$ vs. $H_A: \mu \leq 5.5$
- C) $H_0: \mu < 5.2$ vs. $H_A: \mu \geq 5.2$
- D) $H_0: \mu \geq 5.2$ vs. $H_A: \mu \leq 5.2$

17. The test statistic of the test is:

- A) $Z = 2.372$
- B) $T = 2.372$
- C) $T = -2.372$
- D) $Z = -2.372$

18. With $\alpha = 0.05$, the critical value of the test is:

- A) $Z = 1.833$
- B) $T = 1.833$
- C) $T = -1.833$
- D) $Z = -1.833$

Part 1 – Multiple Choice

[50 marks]

For each of the following questions, circle the **best** answer or response. [1 mark each].

1. A sample proportion is the best point estimate of...
A) the population standard deviation B) the population median
C) the population mean D) the population proportion
2. The Director of Marketing Research, is evaluating consumer acceptance of a new Her staff reports that 17% of a random sample of 200 households prefers the new package to all other package designs. If She concludes that 17% of all households prefer the new package, she is using _____.
A) a point estimate
B) a range estimate
C) a statistical parameter
D) an interval estimate
3. A 95% confidence interval for the mean time μ spent in the shower (in minutes) by adult males is calculated to be 6 ± 3 . Which of the following provides a correct interpretation of the confidence interval?
A) The probability that μ lies in this interval is 0.95
B) In 95% of all samples of the same size, the sample mean will be between 3 and 9.
C) 95% of all adult males spend between 3 and 9 minutes in the shower
D) In repeated sampling, 95% of similarly constructed intervals will contain μ .
4. To construct a confidence interval we first choose.....
A) an interval of interest
B) a confidence level
C) a confidence coefficient
D) minimum and maximum values
5. We would like to construct a confidence interval for the mean μ of some population. Which of the following combinations of confidence level and sample size will produce the narrowest interval?
A) 90% confidence, $n = 20$ B) 90% confidence, $n = 50$
C) 96% confidence, $n = 20$ D) 96% confidence, $n = 50$