

STATISTICS 2023

NAME IN PRINT _____

FINAL EXAM

SIGNATURE IN INK _____

FALL 2002

SS # OR ID IN INK _____

TRUE OR FALSE. Answer with a capital T or F.

(3 points each)

_____ 1. If a hypothesis test on a population mean produced a z-test statistic of 4.5 then the researcher would conclude with an alpha-value of 5% that the null hypothesis is false.

_____ 2. In a small sample test on the difference between two means, if variances are assumed equal then a pooled estimate of variance is used in the standard error for the difference between the sample means.

_____ 3. The standard deviation of a sample measures the center of the data.

_____ 4. The p-value of a hypothesis test is the error rate that must be tolerated if the null hypothesis is rejected.

_____ 5. The t and Z probability density functions are both left skewed.

_____ 6. Hypotheses are statements about values of point estimates.

_____ 7. The point estimator used to estimate a population parameter is almost always equal to the parameter being estimated.

_____ 8. The statement that is assumed true and tested for validity in the hypothesis test process is the alternative hypothesis.

_____ 9. An interval estimator for a population parameter called a confidence interval provides a set of reasonable and plausible values for the parameter.

_____ 10. If the value of the estimated linear correlation, r , is close to positive one then a researcher could conclude that the bivariate data lies close to an upward sloping line.

_____ 11. What is the numerical value of the mean of a sample of six observations, 42, 58, 29, 32, 31, and 27?

_____ 12. What is the numerical value of the standard deviation of a sample of six observations, 42, 58, 29, 32, 31, and 27? Round your answer to two digits past the decimal.

_____ 13. If a hypothesis test based on a large sample has a test statistic value of 2.28 and the researcher is trying to prove that the population mean is not equal to some stated value what is the p-value of the hypothesis test?

_____ 14. If the p-value in a right tail hypothesis test based on a z-test statistic is equal to 0.0885 what is the numerical value of the test statistic?

_____ 15. In a two-tail hypothesis test based on a small sample of 19 observations the absolute value of the test statistic must exceed what number for the researcher to reject the null hypothesis with only a 0.05 error rate?

_____ 16. If a researcher who was comparing the means of two populations was attempting to prove that the mean of population one is more than 3 units larger than the mean of population two, what is the appropriate alternative hypothesis?

_____ 17. If the mean of the sample from population one is 25.1 and the mean of the sample from population two is 28.9 what is the numerical value of the point estimate for the mean of population one minus the mean of population two?

_____ 18. If a sample of 10 observations had a sample variance of 4 and a sample of 15 observations had a sample variance of 5 then what is the numerical value of the pooled variance estimate based on these two samples?

STATE THE ANSWER. Write the answer on the line.

(3 points each)

A marketing executive is looking into running a TV commercial for a client. Their budget is limited and they want to know if it is more advantageous to run the AD in the New York or the Chicago area. They look into the mean number of hours per week people in these areas watch TV to determine where their money is better spent. Use this information to answer the questions on this page.

New York

$n_1=100$

mean=15

variance=4

Chicago

$n_2=100$

mean=23

variance=5

_____ 19. What is the numerical value of the point estimate for the mean hours of TV watched by people in Chicago?

_____ 20. What is the numerical value of the point estimate for the standard deviation of TV watched by people in New York?

_____ 21. What is the estimated standard error of the point estimate for the mean hours of TV watched by people in New York?

_____ 22. What is the numerical value of the point estimate for the difference between the average hours of TV watched in New York and the average hours of TV watched in Chicago?

_____ 23. What is the estimated standard error of the point estimate for the difference between the average hours of TV watched in these two cities not assuming equal population variances?

_____ 24. If the estimated standard error of the point estimate for the difference between the average hours of TV watched in these two cities is 0.9 hours what is the bound of error for a 95% confidence interval to estimate the difference between the mean hours of TV watched in these two cities?

_____ 25. A pooled variance estimate would probably not be used in this case due to large sample sizes, but if it were calculated what would be the numerical value of the pooled variance estimate based on these two samples?

_____ 26. The observed test statistic in this case would have to exceed what magnitude for the data to indicate at the 5% significance level that the two types of investments do not result in equal mean returns?

LINEAR REGRESSION QUESTIONS. Write the answer on the line. (3 point each)

The number of years on a store shelf is used to estimate the potency of a drug. The bivariate data recorded below are number of years on a store shelf (X) and the potency of a drug (Y). Use this data to answer the next five questions. The answers to these questions are related so double check your calculations.

| | | | | | |
|---|----|----|----|----|----|
| X | 1 | 4 | 4 | 3 | 2 |
| Y | 98 | 65 | 70 | 85 | 92 |

_____ 27. What is the numerical value of the corrected sum of squares for the y-variable based on the above data?

_____ 28. What is the least squares estimate of the slope in the linear regression equation to estimate potency of a drug from the number of years on a store shelf? Round your answer to two digits past the decimal.

_____ 29. What is the least squares estimate of the y-intercept in the linear regression equation to estimate potency of a drug from the number of years on a store shelf? Round your answer to one digit past the decimal.

_____ 30. Write the estimated regression equation to estimate the potency of a drug from the number of years on a store shelf. Use the estimates of the slope and y-intercept that you calculated in problems 28 and 29 above.

_____ 31. What is the numerical value of the estimated linear correlation between the two variables, the potency of a drug and the number of years on a store shelf? Round your answer to two digits past the decimal.

ANOTHER LINEAR REGRESSION QUESTION.

The estimated linear regression equation below uses X=years experience to estimate Y=salary of an employee. Use it to answer the remainder of the questions on this page.

$$\hat{y} = 40,500 + 526x$$

_____ 32. If an employee has 2.5 years experience, what is the least squares estimate for the salary of this employee?

_____ 33. If the estimated standard error for the estimate of the slope is 263 what is the value of the test statistic to test whether the slope is equal to zero?