STATISTICS 2023 **NAME IN PRINT**________________________

FINAL EXAM **SIGNATURE IN INK**________________________

FALL 2010 **CWID IN INK**________________________

FILL IN THE BLANK. Write the word or phrase from the list on the right that belongs in the blank. (Each blank 2 points each)

THE WORDS OR PHRASES IN THIS LIST MAY BE USED MORE THAN ONCE OR NOT USED AT ALL.

Standard error
Test statistic
Null hypothesis
Alternative hypothesis
Parameter
Sample statistic
Confidence interval
Point estimate
Bound of error
Population
Sample
Statistical inference
Alpha
P-value
Rejection region
Critical value
Population parameter

1. The decision in a hypothesis test is whether or not to reject the null hypothesis.

2 The confidence intervals covered in this course are centered around the point estimate for the parameter being estimated.

3. A Sample statistic is a value calculated from the data that is used as a point estimate.

4. A population parameter is an unknown constant that described the population.

5. A confidence interval is a way to construct an interval estimate so that there is a certain degree of accuracy associated with the estimator.

6. A test statistic is a calculation from the data that is used to test the null hypothesis.

7. A test statistic is used to test the validity of the null hypothesis.

8. The tail area associated with the test statistic is the p - value of the hypothesis test.

9. If the null hypothesis is rejected based on specific data, then the risk of error is called the p - value of the hypothesis test.

10. In a confidence interval the bound of error is added and subtracted from the point estimate to form the interval estimator.

11. The rejection region OR critical value divides the distribution of the test statistic under a true null hypothesis into likely and unlikely values.

Z and t table questions **(2 points each)**

-1.53 ____ 12. What is z_o, such that $P(Z< z_o)=0.0630$?

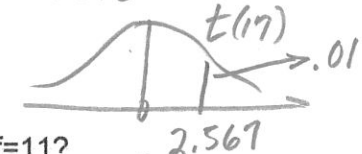
$.167$ ____ 13. What does $P(.85 < Z < 1.87)$ equal?

$.4693 - .3023 = .1670$

$.9918$ ____ 14. What does $P(Z > - 2.4)$ equal?

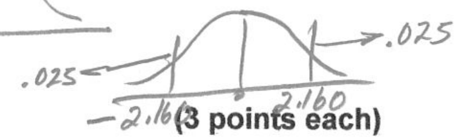
$\to .5 + .4918$

$.01$ ____ 15. What is the $P(t > 2.567)$ in the t with df=17?



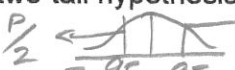
-4.025 ____ 16. What is the value of t_o if $P(t < t_o) = .001$, when df=11?

$.95$ ____ 17. What is the $P (-2.160 < t < 2.160)$, if df=13?



STATE THE ANSWER. Write the answer on the line. **(3 points each)**

$.3422$ ____ 18. What is the p-value of a two-tail hypothesis test based on a large sample if the test statistic value is 0.95?



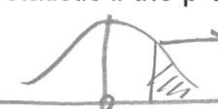
$P/2 = .5 - .3289 = .1711$

$P = 2(.1711) = .3422$

375.5 ____ 19. If a 90% confidence interval to estimate a population mean is (353, 398) what is the value of the point estimate for the population mean?

$\bar{X}$ is the center of the interval

1.39 ____ 20. What is the value of the test statistic if the p-value in a right tail hypothesis test based on a Z test statistic is equal to 0.0823?



$\to P = .5 - .4177 = .0823$

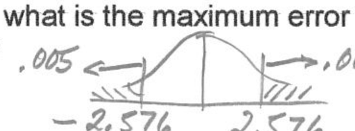
143.54 ____ 21. Consider a 98% confidence interval to estimate a population mean based on a sample of 25 observations with a sample mean of 425 and a sample standard deviation of 144. How wide is this interval? Round to two digits past the decimal.

$$W = 2B = 2\left(t_{\frac{.02}{2}(25-1)} \cdot S_{\bar{x}}\right) = 2(2.492)\frac{144}{\sqrt{25}} = 143.54$$

100 ____ 22. If a data set with fifty observations yields a sum of squares of 12,687.52 and a sum of 624 what is the value of the point estimate for the population variance?

$$S^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1} = \frac{12{,}687.52 - \frac{624^2}{50}}{49} = 100$$

$.01$ ____ 23. If the rejection region in a two-tail hypothesis test based on a large sample is above 2.576 and below -2.576 what is the maximum error rate of rejecting a true null hypothesis that this researcher will tolerate?

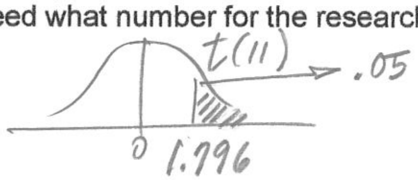


$.005 = \frac{\alpha}{2} \Rightarrow \alpha = .01$

$\mu_1 - \mu_2 > 30$ ____ 24. If the research question is, "Do the data indicate that the mean of the first population exceeds the mean of the second population by more than 30 units?" what is the appropriate alternative hypothesis?

$H_a: \mu_1 - \mu_2 > 30$

1.796 ____ 25. In a right-tail hypothesis test based on a small sample of 12 observations the value of the test statistic must exceed what number for the researcher to reject the null hypothesis with only a 0.05 error rate?



State the answer on the line. **(3 points each)**

A loan officer compares the interest rates for 48-month fixed-rate auto loans and 48-month variable-rate auto loans. Two independent, random samples of auto loan rates are selected. Use the following two samples to answer the questions on this page. Assume equal population variances for this page of questions.

Fixed-rate: 8.9% 7.8% 7.6% 9.1% 6.8% 8.2% 7.6%

Variable-rate: 6.2% 4.5% 5.5% 4.4% 6.1% 6.9%

5.6 26. State the point estimate for the mean of the population of variable-rate loan percentage. State your answer with one digit past the decimal.

$$\bar{X}_V = \frac{\sum X_V}{6} = 5.6$$

.8 27. State the point estimate for the standard deviation for the population of fixed-rate loan percentage. Round your answer to one digit past the decimal.

$$S_F = .80208$$

.99 28. State the point estimate for the variance for the population of the variable-rate loan percentage. Round your answer to two digits past the decimal.

$$S_V^2 = .992$$

2.4 29. State the point estimate for the difference between the mean of the population of fixed-rate loan percentage and the mean of the population of variable-rate loan percentage. State your answer with one digit past the decimal.

$$\bar{X}_F - \bar{X}_V = 8.0 - 5.6 = 2.4$$

.8 30. What is the pooled variance estimate that would result from these two samples? Round your answer to one digit past the decimal.

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{6(.643\bar{3}) + 5(.992)}{11} = .8018$$

11 31. How many degree of freedom are associated with the pooled variance estimate from these two samples?

$$n_1 + n_2 - 2 = 7 + 6 - 2 = 11$$

4.8 32. If the estimated standard error of the difference between the sample means is .5 then what is the value of the test statistic to test whether the mean of the population of fixed-rate loan percentages is equal to the mean of the population of variable-rate loan percentages? State your answer with one digit past the decimal.

$$t = \frac{\bar{X}_1 - \bar{X}_2 - 0}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{8.0 - 5.6 - 0}{.5} = 4.8$$

Doesn't indicate how to round.

1.1005 33. What is the value of the bound of error that would be used to construct a 95% confidence interval to estimate the difference between the mean of the population of fixed-rate loan percentages and the mean of the population of variable-rate loan percentages if the estimated standard error of the difference between the sample means is .5? Assume equal population variances.

$$B = t_{\frac{\alpha}{2}(n_1+n_2-2)} \cdot S_{\bar{X}_1-\bar{X}_2} = t_{\frac{.05}{2}(11)} \cdot S_{\bar{X}_1-\bar{X}_2} = 2.201(.5) =$$

$$1.1005$$

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

The average hourly temperature affects the weekly fuel consumption of natural gas in a small city. The following data are the average hourly temperature measured in Fahrenheit degrees (X) and the weekly fuel consumption of natural gas measured in millions of cubic feet (Y) for eight randomly chosen weeks for a small city in the mid-west United States. Use this data to answer the following questions.

X	28.0	28.0	32.5	39.0	45.9	57.8	58.1	62.5
Y	12.6	11.9	12.2	11.8	9.4	9.1	8.0	7.5

$\Sigma x = 351.8$, $\Sigma y = 82.5$, $\Sigma x^2 = 16{,}874.76$, $\Sigma y^2 = 879.87$, $\Sigma xy = 3{,}433.69$

$SS_x = 1{,}404.355$, $SS_y = 29.08875$, $SS_{xy} = -194.2475$

3,433.69 34. What is the sum of the cross product for the daily average temperature and fuel consumption data provided above? State your answer with two digits past the decimal.

29.08875 35. What is the numerical value of the corrected sum of squares for the y-variable based on the daily average temperature and fuel consumption data provided above? State your answer with five digits past the decimal.

$$SS_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n} = 879.87 - \frac{(82.5)^2}{8} = 29.08875$$

−.1383 36. What is the least squares estimate of the slope in the linear regression equation to estimate the weekly fuel consumption based on the average daily temperature? Round your answer to four digits past the decimal.

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_x} = \frac{-194.2475}{1{,}404.355} = -.1383179$$

16.4 37. What is the least squares estimate of the y-intercept in the linear regression equation to estimate the weekly fuel consumption based on the average daily temperature? Round your answer to one digit past the decimal.

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \frac{82.5}{8} - (-.1383)\frac{351.8}{8} = 16.3942$$

9.8 38. What is the least squares estimate of the weekly fuel consumption based for an average daily temperature of 48F? Round your answer to one digit past the decimal.

$$\hat{y}_{x=48} = 16.4 - .1383(48) = 9.7616$$

−.96 39. What is the numerical value of the estimated linear correlation between the two variables? Round your answer to two digits past the decimal.

$$r = \frac{SS_{xy}}{\sqrt{SS_x \cdot SS_y}} = \frac{-194.2475}{\sqrt{1{,}404.355(29.08875)}} =$$