STATISTICS 2023 **NAME IN PRINT**_____

FINAL EXAM **SIGNATURE IN INK**_____

SPRING 2008 **CWID IN INK**_____

TRUE OR FALSE. Answer with a capital T or F. **(3 points each)**

T 1. If a hypothesis test on a population mean produced a z-test statistic of 8.5 then the researcher could conclude with very low error rate that the null hypothesis is false and that the data supports the alternative hypothesis.

F 2. In a small sample test on the difference between two means, if the population variances are not assumed equal, then a pooled variance estimator is used to estimate the variances of both populations in the calculation of the standard error for the difference between the sample means.

T 3. The magnitude of the mistakes or standard errors for point estimates increase when the sample size decreases.

F 4. Hypotheses are hypothetical statements about values of point estimates.

F 5. A statistical point estimator used to estimate a population parameter is almost always equal to the parameter being estimated.

F 6. An interval estimator for a population parameter, called a confidence interval, provides a set of values for the parameter that would be rejected in a two tail hypothesis test with the same significant level as the confidence interval.

FILL IN THE BLANK **(3 points each)**

Choose the word or phrase from the following list to write on each blank below. The words or phrases can be used more than once. Please write clearly. The words and phrases to choose from are:

confidence interval	point estimate	p-value	null hypothesis
alternative hypothesis	standard error	population parameters	
bound of error	sample variance	test statistic.	

7. Sample statistics are used as point estimates to estimate population parameters.

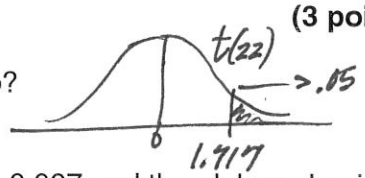
8. The probability of an error when a null hypothesis is rejected based on specific data is called the p-value of the hypothesis test.

9. The general form of a confidence interval covered in this course has as a center value the point estimate for the parameter being estimated.

10. The width of a confidence interval is twice the value of the bound of error.

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

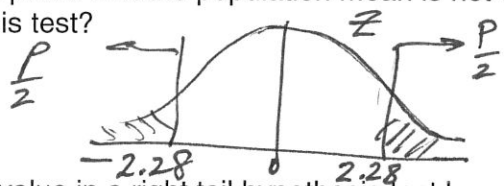
1.717 11. What is the value of $t_{0.05, 22}$ equal to?



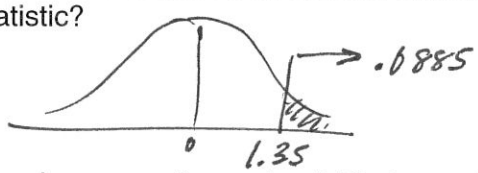
col → row

.007 12. If the p-value in a hypothesis test is 0.007 and the alpha value is 0.03, what is the probability of observing data at least as extreme as the data observed in the sample assuming that the null hypothesis is true?

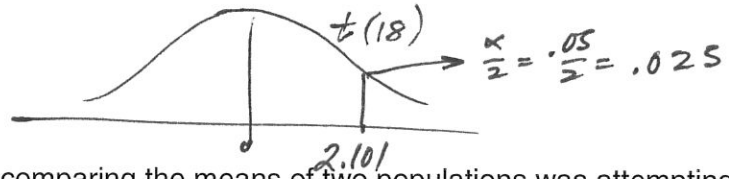
.0226 13. If a hypothesis test based on a large sample has a test statistic value of 2.28 and the researcher is trying to prove that the population mean is not equal to some stated value what is the p-value of the hypothesis test?



1.35 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?



2.101 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?



$M_1 - M_2 < -3$ 16. If a researcher comparing the means of two populations was attempting to prove that the mean of population two is more than 3 units larger than the mean of population one, what is the appropriate alternative hypothesis? Write the parameter as mean one minus mean two.

$M_2 - M_1 > 3 \Rightarrow M_1 - M_2 < -3$

-3.8 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?

$M_1 - M_2 = \bar{X}_1 - \bar{X}_2 = 25.1 - 28.9 = -3.8$

36 18. If two samples of the same size produced variances estimates of 28 and 44, what is the value of the pooled variance estimate that would result from these two samples?

If $n_1 = n_2$ then $s_p^2 = \frac{s_1^2 + s_2^2}{2}$

$168,900. 19. The selling price of houses, y, measured in 1000's of dollars was estimated from the square footage of house, x, measured in 100's of square feet. The estimated regression equation that describes the relationship between these variables is 56.2 + 4.9 x. What is the estimated selling price a house with 2,300 square feet? State your answer in dollars.

$\hat{y} = (56.2 + 4.9(23))\,1000 = \$168,900.$

STATE THE ANSWER. Write 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. A sample of eight 48-month fixed-rate auto loans had the following loan rates:

10.29% 9.75% 9.50% 9.99% 9.75% 9.99% 11.40% 10.00%

while a sample of five 48-month variable-rate auto loans had loan rates as follows:

9.59% 8.75% 8.99% 8.50% 9.00%

Hypothesis Test: **Independent Groups**

(t-test, pooled variance)

Fixed	Variable
10.0838	
	0.4046
8	5

df

difference (Fixed - Variable)

pooled variance

0.29862 standard error of difference

0 hypothesized difference

3.7 t

0.0032 p-value (two-tailed)

8.966 20. State the mean of the sample of variable rate loan percentages. State your answer with three digits past the decimal. $\bar{X}_2 = \frac{\Sigma X_2}{n_2}$

.581 21. State the estimate for the standard deviation for the population of fixed rate loan percentages. State your answer with three digits past the decimal. $S_1 = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}{n-1}}$

1.1178 22. State the estimate of the difference between the mean of the population of fixed rate loan percentages and the mean of the population of variable rate loan percentages. State your answer with four digits past the decimal. $\mu_F - \mu_V = \bar{X}_F - \bar{X}_V$

$\mu_1 - \mu_2 \neq 0$ 23. What is the alternative hypothesis if the question is whether the data supports the conclusion that there is a difference in the mean of the population of loan percentages for fixed rate loans and the mean of the population of loan percentages for variable rate loans?

3.743 24. 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 three digits past the decimal. $t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{1.1178}{.29862}$

2.201 25. What is the value of the t multiplier that would be used in the bound of error 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? State your answer with three digits past the decimal. $t_{\frac{.05}{2}(n_1+n_2-2)} = t_{.025(11)}$

2.069 26. What is the value of the test statistic that would be used to test the hypothesis that the difference between the mean of the population of fixed rate loan percentages and the mean of the population of variable rate loan percentages equals 0.5?

$$t = \frac{\bar{X}_1 - \bar{X}_2 - .5}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{1.1178 - .5}{.29862} = 2.069.$$

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	3	4	5	2
Y	98	75	60	35	92

See example in spiral following Ch 18 for calculation details.

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

$$SS_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$$

−15.8 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? State your answer with one digit past the decimal.

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_x}$$

119.4 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? State your answer with one digit past the decimal.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$\hat{y} = 119.4 - 15.8x$ 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.

−.98 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.

$$r = \frac{SS_{xy}}{\sqrt{SS_x \cdot SS_y}}$$

ANOTHER LINEAR REGRESSION QUESTION. (3 points each)

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.

y-hat = 38,200 + 1,315x

$41,487.5 32. If an employee has 2.5 years experience, what is the least squares estimate for the salary of this employee?

$$\hat{y}_{x=2.5} = 38,200 + 1,315(2.5) =$$

5 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?

$$t = \frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} = \frac{1,315 - 0}{263} = 5.$$