Name: ______________________________

ID: ______________________________

Quiz section or time: ______________________________

8 +

Stat/Math 390, Winter, Test 3, March 9, 2012; Marzban

Points

1 7.33 **1.** We want to see if there is evidence that the difference between two means exceeds 2.1. Which is the appropriate CI to compute?

a) 1-sample, 2-sided b) 2-sample, 2-sided c) 2-sample, upper-bound d) 2-sample, lower-bound

1 7.53 Lab **2.** The t-interval assumes that the population is Normal. Which is the best way of assessing the validity of that assumption when a paired 2-sample CI-interval (or test) is computed?

a) Examine qqplot of sample differences. c) Fit a regression line to the 2 samples.

b) Examine scatterplot of the 2 samples. d) It's impossible to check a population assumption.

1 8.2 **3.** To decide whether the pipe welds in a nuclear power plant meet specifications, a random sample of welds is to be selected and the strength of each weld (force required to break the weld) determined. Suppose a population mean strength of $100 lb/in^2$ is the dividing line between welds meeting specification or not doing so. The appropriate H_1 is (Hint: Think of Type I and II errors, and which one is the "bad error."

Type I = (reject Ho when Ho = True) = Bad error ✓. $H_0 : \mu < 100$

a) $H_1 : \mu < 100$ b) $H_1 : \mu \neq 100$ c) $H_1 : \mu > 100$

1 669, 8.62a, 1.56 **4.** In which of the following situations it is misleading/wrong to look at comparative boxplots?

a) Comparing two independent samples. c) Comparing two paired samples.

b) Comparing three independent samples. d) None of the above.

1 8.75 **5.** For each sport, the following data are a sample of games selected from all games played during a season. For each sport, a "Leader" is defined to be a team which is winning midway through a season of many games. We want to know if the 4 sports appear to be identical with respect to the proportion of games won by the Leader. The appropriate test is

	Basketball	Baseball	Hockey	Football
Leader wins	150	86	65	72
Leader loses	39	6	15	21

a) chi-squared test of homogeneity of 4 populations with respect to 2 categories

b) chi-squared test of homogeneity of 2 populations with respect to 4 categories

c) chi-squared test of specific values of four proportions in 1 population.

d) An F-test of the equality of 4 population means.

1 9.12 **6.** In 1-WAY ANOVA for testing k population means, $\mu_i (i = 1, k)$, which of the following is true?

a) The test can be done with k sample means (i.e., without the sample variances) .

b) The sample sizes for the k samples must be equal.

c) The null hypothesis is $H_0 : \mu_1 = \mu_{01}, \mu_2 = \mu_{02}, ...,$ where μ_{0i} are the null parameters.

d) None of the above.

1 11.3b7 **7.** In a regression problem it has been previously believed that when a specific predictor changes by one unit, then the associated true average of y would change by at most 2 units. To test this belief, the most appropriate quantity to compute is

a) CI for true mean y b) PI for an observed y c) CI for β d) F-ratio.

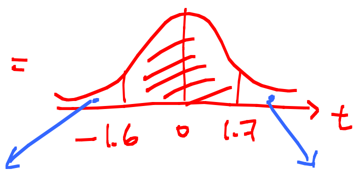
1 11.7 **8.** A 95% CI for $y(x)$ will cover individual observed values of y _______ than 95% of the time. Hint: think of both the CI and the PI.

a) less often b) equally often c) more often

95% PI for y(x) covers individual obs values of y 95% of time. PI > CI.

7.45 **9.** There are situations in which an asymmetric CI is called for. What is the confidence level associated with the t-interval $(\overline{x} - 1.6\frac{s}{\sqrt{11}},\ \overline{x} + 1.7\frac{s}{\sqrt{11}})$ for a single population mean?

Conf. level = prob (a < t < b)

= 

$= 1 - (.06 + .07) = 1 - .13 = 0.87$

$= \boxed{87\%}$

0.07   Table VI   0.06

df = 11 − 1 = 10

~~2~~,3 2 9.14 **10.** Three groups of rats consisting of 10 rats per group, are placed on three different diets. Some months later, the weight of each rat is measured. We want to know if the diet had an effect on the average weight. The sample average weights are shown below. Assuming the total sum of squares is 506, compute an appropriate p-value. Hint/calculator: $486/27 = 18$.

Diet	Starch	Glucose	Maltose
Sample mean	3.0	4.0	5.0

↳ 1-way ANOVA F-test, $k = 3$

same #'s as sample-test

$n_i = 10$

$n = 30$

$$\bar{\bar{y}} = \sum_{i=1}^{3} \frac{n_i}{n} \bar{y}_i = \frac{1}{3}(3+4+5) = 4$$

$$SS_{between} = \sum_{i=1}^{3} n_i(\bar{y}_i - \bar{\bar{y}})^2 = 10(3-4)^2 + 10(4-4)^2 + 10(5-4)^2 = 20$$

$$SS_{within} = SST - SS_{between} = 506 - 20 = 486$$

$$F_{obs} = \frac{SS_{between}/(3-1)}{SS_{within}/(30-3)} = \frac{20/2}{486/27} = \frac{10}{18} = \frac{5}{9} \simeq .55 \Rightarrow \underset{df=(2,27)}{\text{Table VII}} \Rightarrow \boxed{\text{p-value} > 0.1}$$

~~2~~,3 3 11.24 **11.** In a simple regression problem, SHOW how the std dev of the OLS estimate of the intercept, $s_{\hat{\alpha}}$, is related to the std dev of the OLS prediction, $s_{\hat{y}}$, where $s_{\hat{y}} = s_e\sqrt{\frac{1}{n} + \frac{(x-\overline{x})^2}{(n-1)s_x^2}}$.

$$\hat{y}(x) = \hat{\alpha} + \hat{\beta}x \Rightarrow \text{at } x = 0,\ \hat{y}(x=0) = \hat{\alpha}$$

$$\therefore S_{\hat{\alpha}} = S_{\hat{y}(x=0)}$$

$$= S_e\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{(n-1)S_x^2}}$$

~~2~~,3 2 11.? **12.** In a simple regression problem where $n = 9, s_{\hat{y}} = 3, s_e = 4$, what is the value of T_0 such that Prob(prediction error $> T_0) = 0.35$?

$$0.35 = \text{Prob}(\text{pred. err.} > T_0) = \text{Prob}\left(\underbrace{\frac{\text{pred. err.}}{S_{\text{pred. err}}}}_{t} > \underbrace{\frac{T_0}{\sqrt{S_{\hat{y}}^2 + S_e^2}}}_{T_0/\sqrt{9+16}}\right)$$

$$0.35 = \text{prob}\left(t > \underbrace{\frac{T_0}{5}}_{}\right)$$

$$\underset{df = 9-2 = 7}{\text{Table VI}} \longrightarrow = 0.4 \Rightarrow T_0 = 5(0.4) \Rightarrow \boxed{T_0 = 2}$$