Name: ______________________

ID: ________________________

Quiz section or time: ________________________

9 + 14.5

Stat/Math 390, Winter, Test 3, March 15, 2013; Marzban

CLOSED everything. ONLY a half-size "cheat sheet" is allowed. Check front page and back page.

Multiple-choice: mark answers on these pages. DO NOT EXPLAIN. There is wrong-answer penalty.

The rest: SHOW answer & WORK on these pages; NO CREDIT FOR CORRECT ANSWER WITHOUT EXPLANATION

Points

2 **1.** When computing a CI for a single mean, which of the following is/are true?

0.5 points each

a) If population is normal, and sample size is large, then z- or t-intervals may be used.
b) If population is normal, and sample size is small, then only t-intervals may be used.
c) If population is not normal, and sample size is large then z- or t-intervals may be used.
d) If population is not normal, and sample size is small, then neither z- nor t-intervals may be used.

Technically, if pop is not normal, then we should not use t at all!

1 7.8 **2.** A 2-sample test of means is equivalent to a 1-sample test of the difference between means,
a) only for paired data
c) only for a 2-sided test
b) for paired or unpaired data
d) only for a 1-sided test

1 **3.** The test $H_0 : \mu < 1 \,; H_1 : \mu > 1$ is equivalent to
a) $H_0 : \mu \leq 1 \,; H_1 : \mu > 1$
c) $H_0 : \mu = 1 \,; H_1 : \mu \neq 1$
b) $H_0 : \mu \geq 1 \,; H_1 : \mu < 1$
d) none of the above.

1 8.2 **4.** 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 most appropriate H_0, H_1 pair is
a) $H_0 : \mu \leq 100 \,; H_1 : \mu > 100$ b) $H_0 : \mu \geq 100 \,; H_1 : \mu < 100$ c) $H_0 : \mu = 100 \,; H_1 : \mu \neq 100$

1 **5.** If a population consists of a single categorical variable, with more than 3 levels, then which test(s) is/are possible?

1 point, with 1 point penalty.

a) t-test b) chi-squared test c) ANOVA F-test d) Regression F-test

1 **6.** If a population consists of a single categorical variable (with 2 or more levels) and a continuous variable, then which test(s) is/are possible?

0.5 point each, but with 0.5 point penalty.

a) t-test b) chi-squared test c) ANOVA F-test d) Regression F-test

1 **7.** In regression, the notion of a 1-sided interval for the response, y, is appropriate
a) only for CIs b) only for PIs c) for CIs and PIs d) neither CIs nor PIs

1 **8.** In regression, as the point x at which a prediction is made, moves away from the sample mean of x, how do the CI and PI for y change?
a) only CI gets wider b) only PI gets wider c) both CI and PI get wider d) neither gets wider

~ 2 7.54 **9.** To provide effective treatment for a desease, it is important to detect it as early as possible. For a group of 100 people, the age at onset of some symptom (e.g., a skin mold) is recorded, and the detection method is then applied annualy. The age at which the detection method detects the deasease is then recorded. What is the most appropriate confidence interval/bound to see if the data provide evidence that the detection method detects the deasease within 3 years of the onset of symptoms? Define the pop parameter(s), and specify your answer completely.

μ_1 = mean age at onset

μ_2 = " " " detection

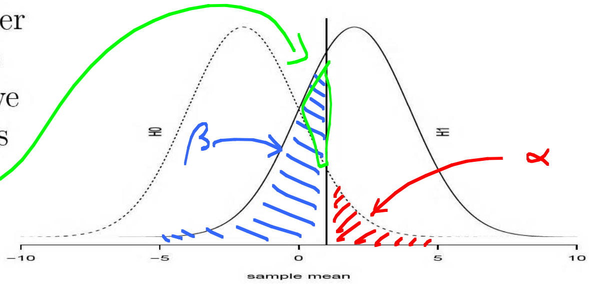
upper conf. bound for $\mu_2 - \mu_1$ for paired data

4 elements!

~ 2 Sample Test **10.** The figure shows the sampling distribution of $\bar{x}$ under the null hypothesis (H_0) and the alternative (H_1). If the sample mean is is to the right of the vertical line, then we reject H_0 in favor of H_1. On this figure shade the regions that correspond to α and β, i.e., the prob of Type I and Type II errors.

1.5 if β excluded

0 = both wrong, 1 = only 1 right, 2 = both right



~~~3~~ 2.5 8.44 **11.** We draw x-y axes on a table and drop 16 beads from a fixed point above the origin onto the table. The number of beads ending up in each quadrant is 1, 7, 6, and 2, respectively. Does this data suggest that the table is not level? State the hypotheses regarding CLEARLY DEFINED POPULATION PARAMETERS, compute a p-value, and state your conclusion regarding the levelness of the table ($\alpha = 0.1$).

π_i = prop. of beads in i^{th} quadrant.

Level Table $\rightarrow$ $H_0: \pi_1 = \frac{1}{4},\ \pi_2 = \frac{1}{4},\ \pi_3 = \frac{1}{4},\ \pi_4 = \frac{1}{4}$

H_1: At least 1 π_i is not $1/4$.

Exp count under H_0: $\frac{1}{4}\,16 = 4$ for all i

obs. count : 1, 7, 6, 2

$$\chi^2_{obs} = \frac{(4-1)^2}{4} + \frac{(4-7)^2}{4} + \frac{(4-6)^2}{4} + \frac{(4-2)^2}{4}$$

$$\chi^2_{obs} = \frac{1}{4}(9+9+4+4) = \frac{26}{4} = \frac{13}{2} = 6.5$$

From Table VII,

for $df = k-1 = 4-1 = 3$, $\chi^2_{obs} = 6.5$

p-value = 0.09.

At $\alpha = 0.1$, reject H_0 in favor of H_1

I.e. There is evidence from data that Table is not level.

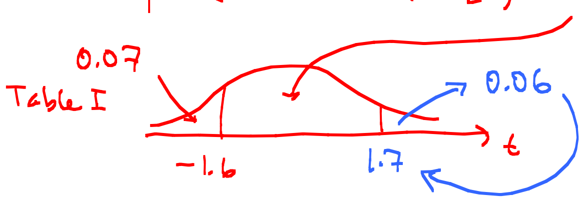
7.45 as PI **12.** In a simple regression problem, an 87% PI for y^* is desired, BUT of the form $(\hat{y} - t_2 X, \hat{y} + t_1 X)$, where the multipliers t_1 and t_2 are different.

~~~2~~ 2.5 **a)** In terms of THESE SYMBOLS, write down the appropriate "self-evident fact."

$$\text{PI} \Rightarrow (\hat{y} - t_2 X < y^* < \hat{y} + t_1 X) \Rightarrow \text{Prob}\left(-t_1 < \frac{\hat{y} - y^*}{X} < t_2\right) = 0.87$$

~~~2~~ 1.5 **b)** If the sample size is 12 (i.e., $df = 12 - 2 = 10$), and $t_1 = 1.6$, what value must t_2 have?

$\text{prob}(-1.6 < t < t_2) = 0.87$



$\therefore\ t_2 = 1.7$

0 0.5 1 1.5 points

if got 0.07

more than 0.07

c) Moreover, suppose in our analysis the sample mean of x is 1, and the standard deviation of the

~ 2 errors is 1 as well. What is the necessary value of X, if the prediction is made at $x = 1$?

$\bar{x} = 1,\ S_e = 1,\ x = 1.$

The interval is a PI. $\Rightarrow$ $X = S_{pred.err.} = \sqrt{S_{\hat{y}}^2 + S_e^2} = S_e\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}} + 1}$

$$= 1\sqrt{\frac{1}{12} + \frac{(1-1)^2}{\cdots} + 1}$$

$$X = \sqrt{\frac{1}{12} + 1}$$
~~~

$\sim \cancel{3}$
2

13. In simple linear regression, if all the x_i values in the sample are multiplied by a constant, c, how does the F-ratio for a test of model utility change? SHOW WORK. Hint: The F-ratio is a constant times $SS_{\text{explained}}/SS_{\text{unexplained}}$

$$F = \frac{SS_{exp}}{SS_{unexp}} = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (\hat{y}_i - y_i)^2} \quad \text{where } \hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \text{ , and}$$

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{\sum \overbrace{(x_i - \bar{x})}^{c} (y_i - \bar{y})}{\sum \underbrace{(x_i - \bar{x})^2}_{c^2}} \longrightarrow \frac{1}{c}\hat{\beta}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\,\bar{x} \longrightarrow \hat{\alpha}$$

$\hat{\beta} \to \frac{1}{c}$, $\bar{x} \to c$

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \longrightarrow y_i$$

$$\therefore \quad F \longrightarrow \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (\hat{y}_i - y_i)^2} = F.$$

1-way ANOVA $\rightarrow$ 0.5 points

$\hat{\beta}$, no c $\rightarrow$ 1