10 + 12.5

Name: __________________________

ID: __________________________

Quiz section or time: __________________________

Stat/Math 390, Summer, Test 3, Aug. 18, 2016; Marzban

Same deal as test 1, ...

Points

7.54, Spring 14/10

1 **1.** Reflectivity measurements are made of both sides of 61 coins. The company wants to see if the true difference in reflectivity ($\mu_1 - \mu_2$) between the two sides exceeds 7. The company should compute

a) A 2-sample, lower confidence bound for $\mu_1 - \mu_2$
b) A 2-sample, 2-sided CI for $\mu_1 - \mu_2$
c) A 2-sample, upper confidence bound for $\mu_1 - \mu_2$
d) a for paired data
e) b for paired data
f) c for paired data

rosella page

1 **2.** Is it possible to have a data set for which a test of $H_0 : \mu \le \mu_0; H_1 : \mu > \mu_0$ leads to "can't tell," and a test of $H_0 : \mu \ge \mu_0; H_1 : \mu < \mu_0$ on the same data and at the same α also leads to "can't tell?"
a) Yes
b) No.

e.g. when one Test has p-value ~0.5
The other will have 1 - 0.5 ~ 0.5

many

1 **3.** Which statement(s) is/are correct?
a) p-value is the probability that H_0 is true
b) α is the probability that H_0 is true
c) β is the probability that H_1 is true
d) power is the probability that H_1 is true.
e) None of the above.

Lect23

1 **4.** The 1-way ANOVA F-test is appropriate in dealing with
a) Two or more continuous variables
b) Two or more categorical variables
c) One continuous and one categorical variable
d) None of the above.

Spring 14/12

5. Suppose we are interested in whether the mean of y varies across 4 different groups. For the special situation where the sample size is constant (and known) across the 4 groups, for computing the F ratio in 1-Way ANOVA we also need to know the

1 a) sample variance of all of the y observations
b) sample variance of the y observations within each group ← SS within +0.5
c) sample variance of the sample means of each group ← SS between −0.5
d) the sample variances of the sample variances

Lect25

6. In a regression setting, as sample size increases, circle all of the correct statements:
1 a) The width of CI for the true mean of y, at a given x, approaches zero.
b) The width of PI for a single y, at a given x, approaches zero. −0.5
c) The CI becomes wider than PI.
d) The center of the PI becomes larger than the center of the CI.

Lect27

1 **7.** In multiple regression, suppose you have the equation for the fit in the population. When you make a prediction using that equation,
a) the estimation error is always zero
b) the prediction error is always zero
c) both a and b
d) none of the above

est. err. = $\hat{y}(x) - y(x)$
Same

8.6, Summer 14/4

1 **8.** Joe is expected to manufacture a concrete whose mean breaking point μ is at least $13KN/m^2$. Although larger values of μ are desirable, increasing μ is an expensive process. But smaller values of μ can have disasterous consequences. Joe should set up the problem so that $\alpha = prob($ __$\mu > 13$__ $|$ __$\mu < 13$__ $)$. Fill in the blanks with statements regarding μ.

Technically, "Data reject $\mu < 13$ in favor of" hw-power 0.25 penalty.

9 It may not seem like it, but the calculation of the power of a (say, 1-sided) test for a population mean depends on a lot of quantities. Which of the quantities in the adjacent table are necessary for computing

A. The significance level
B. Pop mean if H_0 is true
C. Pop mean if H_1 is true
D. Pop (or sample) variance
E. Sample size
F. Sample mean

1 a) The rejection region? List the capital letters here: 0.25 each ABDE

1 b) The power? No need to repeat the quantities listed in part a: C (1) 0.5 penalty

~2 Ch7.8 **10** Suppose you suspect that the direction in which hair is combed has an effect on how well one does on a test. (If you don't have hair, don't worry about it - everything will be OK.) You observe a 100 students with their hair combed to the right, and another 100 students with their hair combed to the left. You give all the students the same test and count how many in each group scored above 80%. What CI/test is most appropriate? CLEARLY DEFINE THE POPULATION PARAMETERS, AND either state the name of the CI/test, or write a formula.

2-sample, 2-sided test/CI of $\pi_1 - \pi_2$ where

π_1 = pop. prop. of students with > 80% grade among right-combed students

π_2 = " > 80% " " Left-combed students

Formula: $(P_1 - P_2) \pm z^* \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$

~~~3~~ 2 Winter13 /11 **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? Compute a p-value, and state your conclusion regarding the levelness of the table (i.e., in English). Use $\alpha = 0.05$.

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

H_1: At least one of These is wrong

expected counts under H_0: $n\pi_0 = \frac{1}{4}(16) = 4$; obs. counts: 1, 7, 6, 2

$$\chi^2_{obs} = \frac{(1-4)^2}{4} + \frac{(7-4)^2}{4} + \frac{(6-4)^2}{4} + \frac{(2-4)^2}{4} = \frac{9}{4} + \frac{9}{4} + \frac{4}{4} + \frac{4}{4} = 6.5$$

p-value = $pr(\chi^2 > 6.5) \sim 0.09 \Longrightarrow \therefore$ p-value > α

Table VII df = 4-1 = 3

$\therefore$ Cannot reject H_0 in favor of H_1

$\therefore$ There is No evidence That Table is Not level.

~2 Clicker lect23 **12** What is the numerical value of the F-ratio $= \frac{R^2/k}{(1-R^2)/(n-(k+1))}$ in the multiple regression test of model utility, if all $\hat{\beta}_i$ are zero. **Show work.** Hints: Look at the numerator of R^2 only; also, use the fact that if all $\hat{\beta}_i$ are zero, then $\hat{\alpha} = \bar{y}$.

$R^2 = \frac{SS_{exp}}{SS_T} = \frac{1}{SST} \sum_i (\hat{Y}_i - \bar{Y})^2$

If all $\hat{\beta}_i = 0$, Then $\hat{Y}_i = \hat{\alpha} + 0 = \bar{Y}$

$\Rightarrow R^2 = 0 \Rightarrow \boxed{F = 0}$.
~~~

~3 Ch.9 **13.** An article reports the 95% CI for the true mean of x on hot days to be $6 \pm 1.96\frac{4}{\sqrt{16}}$. Another article reports the 95% CI for the true mean of x on cold days to be $1 \pm 1.96\frac{8}{\sqrt{16}}$. Of course, you can always visually compare these CIs and come to a conclusion, but here perform a 1-WAY ANOVA F-test to see if there is a difference between the true mean of x for cold and hot days. Write a numerical expression for the F-ratio, but don't waste time on arithmetic; no need for the p-value, or a conclusion.

From 1st C.I. : $\bar{Y}_1 = 6,\ S_1 = 4,\ n_1 = 16$
" 2nd " : $\bar{Y}_2 = 1,\ S_2 = 8,\ n_2 = 16$
$\Rightarrow$ $n = n_1 + n_2 = 32$, $\bar{\bar{Y}} = \frac{1}{2}(\bar{Y}_1 + \bar{Y}_2) = \frac{7}{2}$

$$SS_{between} = \sum_{i=1}^{2} n_i(\bar{Y}_i - \bar{\bar{Y}})^2 = 16(6-\tfrac{7}{2})^2 + 16(1-\tfrac{7}{2})^2 = 16\left[\left(\tfrac{5}{2}\right)^2 + \left(\tfrac{5}{2}\right)^2\right] = 8(25)$$

$$SS_{within} = \sum (n_i - 1) S_i^2 = 15(4)^2 + 15(8)^2 = 15(16+64) = 15(80)$$

$$F = \frac{SS_{bet}/(k-1)}{SS_{within}/(n-k)} = \frac{8(25)/(2-1)}{15(80)/(32-2)} = \boxed{5}.$$

11.18+... **14.** A simple regression involving $n = 10,000$ gives the line $y = 2 + 3x$, with $s_e = 1$. At $x = \bar{x}$,

~~2~~ 2.5 a) What is the probability that the estimation error is less than 0.1?

$$? = \text{prob}(\text{est. err.} < 0.1) = \text{pr}\left(\frac{\text{est.err}}{S_{\text{est.err}}} < \frac{0.1}{S_{\text{est.err}}}\right)$$

$$= \text{pr}\left(t < \frac{0.1}{0.01}\right)$$

$$S_{\text{est.err}} = S_{\hat{y}} = S_e\sqrt{\frac{1}{n} + \frac{\cancel{(x-\bar{x})^2}}{S_{xx}}} = \frac{1}{100} = 0.01$$

$\to 0$ at $x=\bar{x}$

$$= \text{pr}(t < 10)$$

$$\simeq \boxed{1}.$$

~~2~~ 1 b) What is the probability that the prediction error is less than 0.1?

$$? = \text{pr}(\text{pred. err} < 0.1) = \text{pr}\left(\frac{\text{pred err.}}{S_{\text{pred.err.}}} < \frac{0.1}{S_{\text{pred.err}}}\right)$$

$$= \text{pr}\left(t < \frac{0.1}{1}\right)$$

$$S_{\text{pred.err}} = \sqrt{S_{\hat{y}}^2 + S_e^2} = \sqrt{(0.01)^2 + 1} \simeq 1$$

$$= \boxed{0.54}$$

Table I is good enough, because $n = 10,000$ = large.