Name: ________________________
ID: ___________________________
Quiz section or time: ___________________________

Stat/Math 390, Winter, Test 3, March 13, 2015; Marzban

12.5+

Same deal as test 1, ...

Points

1 — 5. **1.** Recall that for an exponential distribution with parameter λ, the expected value and the standard deviation are both $\frac{1}{\lambda}$. If x has an exponential distribution with parameter λ, the sampling distribution of sample mean has a mean and standard deviation given by

a) $(\bar{x}, \frac{s_x}{\sqrt{n}})$ b) $(\bar{x}, \frac{1}{\lambda})$ c) $(\frac{1}{\lambda}, \frac{s_x}{\sqrt{n}})$ $(\mu_x, \frac{\sigma_x}{\sqrt{n}})$ d) $(\frac{1}{\lambda}, \frac{1}{\lambda}\frac{1}{\sqrt{n}})$

1 — 5.7, 7.16 **2.** Which of the following is/are true? If the population is Normal, and sample size is large, then approximately the sampling distribution of the sample

0.5 penalty.

a) mean is Normal b) proportion is ~~Binomial~~ Normal c) standard deviation is Normal.

1 — 7.11 **3.** Suppose we are operating at 95% confidence level. An upper confidence bound has been computed for a single proportion, and it is found to be 0.2. Then (circle all the correct statements)

a) We can be 95% confident that the true proportion is between 0 and 0.05 0.5 penalty.
b) We can be 95% confident that the true proportion is between 0 and 0.2
c) There is a 95% prob that a random upper confidence bound will be greater than 0.2
d) There is a 95% prob that a random upper confidence bound will be greater than the true prop.

1 — 7.33 **4.** You want to see if the difference between two population means, $\mu_1 - \mu_2$, is at most 13. Then, the appropriate quantity regarding $\mu_1 - \mu_2$ is

a) a lower confidence bound b) a 2-sided CI c) an upper confidence bound d) none of the above.

8.1 **5.** Which of the following is/are NOT a legitimate hypothesis? 0.5 penalty

a) H: $\mu_1/\mu_2 \geq 1$ b) H: $\sigma_1/\sigma_2 \leq 1$ c) H: $\sigma = 1.1$ d) H: $\bar{x} = 45$ e) Poisson $\lambda > 1$

1 — 8.3 **6.** A manufacturer of 40-amp fuses wants to make sure that the true average amperage at which its fuses burn out is indeed 40. If the average amperage is lower than 40, purchasers will complain because the fuses will have to be replaced too frequently, whereas if the average exceeds 40, the manufacturer might be liable for damage to an electrical system due to fuse malfunction. How should the manufacturer setup the hypotheses?

a) $H_0 : \mu \leq 40$; $H_1 : \mu > 40$ c) $H_0 : \mu = 40$; $H_1 : \mu \neq 40$
b) $H_0 : \mu \geq 40$; $H_1 : \mu < 40$ d) none of the above.

1 — 8.7 **7.** Many games involve a disc (called a wheel) with some number of bins (e.g., 12, 36, ...) along its perimeter. The wheel is spun and then one waits until it stops, with one of the bins pointing to a fixed arrow. Suppose, we suspect that the wheel is uneven in that the bins do not have the same probability of stopping at the arrow. Which test can be used for testing this suspicion?

a) t-test b) chi-squared c) 1-way anova F-test d) F-test of model utility

1 — 8.6 **8.** A cement is to be used for building a wall. The strength of the cement is measured by a quantity called Average Compressive Strength (ACS). We do NOT know if higher (lower) ACS implies stronger (weaker) cement, but we do know that a Type I error will be committed if the true mean ACS is in fact lower than 13. What are the appropriate hypotheses?

a) $H_0 : \mu \geq 13,\ H_1 : \mu < 13$ b) $H_0 : \mu = 13,\ H_1 : \mu \neq 13$ c) $H_0 : \mu \leq 13,\ H_1 : \mu > 13$

Type I = (Data reject $\mu < 13$ in favor of $\mu > 13$ | $\mu < 13$) $\Rightarrow$ $H_0 : \mu < 13$

1.5 **9.** The notion of a 1-sided CI, or a 1-sided hypothesis test, exists for which of the following quantities? Circle all correct answers.

a) y-intercept b) regression coefficients c) predictions 0.5 penalty

2 **10.** In regression, circle all of the quantities that can be computed. 0.5 penalty

a) CI for α b) CI for β_i c) CI for $y(x)$ d) PI for α e) PI for β_i f) PI for $y(x)$

1 **11.** 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 7.27 **12.** A 99% CI for the **ratio** between 2 proportions is found to be (0.1, 5). What can one say? regarding equality of π_1 and π_2

C.I. for π_1/π_2 : $[0.1, 5]$ $\Rightarrow$ 1.0 is in the CI.

$\therefore$ π_1 may or may not be equal to π_2. We cannot tell.

$\sim$ ~~2~~ 2.5 **13.** Suppose we drop phones from increasing heights and record the height at which the panel breaks. A friend claims that the three big makers (Apple, Samsung, Ziaomi) all use the same panel manufacturer, and so there is no difference between them in this respect. a) Set up the most appropriate H_0, H_1 in terms of **well-defined** parameters, and b) state the correct statistical test.

a) μ_i = true mean height from which the i^{th} brand breaks.

$H_0: \mu_1 = \mu_2 = \mu_3$ b) 1-Way ANOVA F-Test.

H_1: At least 2 μ's are different

$\sim$ ~~5~~ 5.5 8.38 **14.** Consider the data given in the adjacent table on the energy intake (x_1), and energy expenditure (x_2), for 4 people selected at random. It is important that $(\mu_{\text{intake}} - \mu_{\text{expenditure}})$ exceeds 0.5, otherwise the people will starve to death. Based on this data, is there cause for concern? **a)** Set-up the correct hypotheses in terms of CLEARLY defined parameters, **b)** Compute a p-value, and **c)** answer the question asked (at $\alpha = 0.05$). NOTE: This problem does not require lengthy calculations; pay attention to the design of the experiment.

person	x_1	x_2	d
1	3	2	1
2	8	3	5
3	4	3	1
4	2	1	1

a) 1.5 "Bad error": (Reject $\mu_{in} - \mu_{ex} < 0.5$ in favor of $\mu_{in} - \mu_{ex} > 0.5$ | $\underbrace{\mu_{in} - \mu_{ex} < 0.5}_{H_0}$)

$\therefore H_0: \mu_{in} - \mu_{ex} \leq \boxed{0.5}$ $(\mu_1 - \mu_2)$

$H_1: \mu_{in} - \mu_{ex} > 0.5$

b) 3 The data are paired. $\Rightarrow$ "make a new column"

$\bar{d} = \frac{1}{4}(1+5+1+1) = 2$, $s_d^2 = \frac{1}{3}[1^2 + 3^2 + 1^2 + 1^2] = 4 \Rightarrow s_d = 2$

$\therefore t_{obs} = \frac{2 - \boxed{0.5}}{2/\sqrt{4}} = 1.5$ $df = n-1 = 3$

$\therefore$ p-value $= pr(t > 1.5) \doteq 0.115$

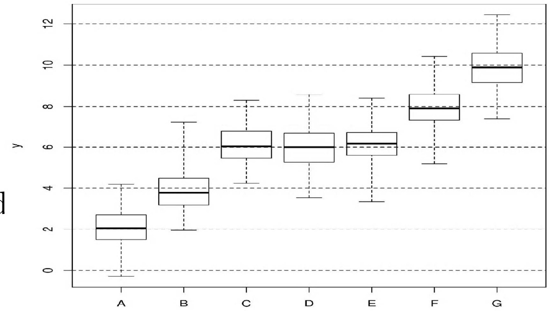
c) 1 At $\alpha = .05$, cannot reject H_0 in favor of H_1. There is cause for concern.

part b) $\bar{\bar{Y}}$, $\bar{Y}_i$, n_i, formula, s_i^2, df (0.5 each)

−0.5 if no p-value.

?.??

~5 **15.** Seven battery brands, labeled A-G, are selected. The battery lifetime, y, of 10 batteries of each brand is measured. The resulting data is summarized in the adjacent graph. The sample variance of each of the boxplots is approximately 1. Is there any evidence that the 7 brands have different lifetimes? **a)** CLEARLY state the hypohteses, **b)** compute a p-value (range), and **c)** state the conclusion "In English." ($\alpha = 0.05$) Note: when reading numbers off this graph, approximate them to the nearest integer.



$\bar{Y}$: 10, 8, 6, 6, 6, 4, 2

a) (1) $H_0: \mu_A = \mu_B = \cdots = \mu_G$, H_1: At least 2 μ's are different.

μ_i = mean lifetime of brand i $\quad \bar{\bar{y}} = \sum_i \frac{n_i}{n} \bar{Y}_i = \frac{10}{70}(10+8+3(6)+4+2) = 6$

b) (3) $SS_{between} = \sum_i^7 n_i(\bar{Y}_i - \bar{\bar{Y}})^2 = 10\left[(2-6)^2 + (4-6)^2 + 3(6-6)^2 + (8-6)^2 + (10-6)^2\right] =$

$= 10[16 + 4 + 0 + 4 + 16] = \underline{400}$

$SS_{within} = \sum^7 (n_i - 1) s_i^2 = 7(9)(1) = 63$ $\qquad$ df = (6, 63) Table $\overline{\underline{VIII}}$

$F_{obs} = \frac{400/(7-1)}{63/(70-7)} = \frac{400/6}{63/63} = \frac{400}{6} = \frac{200}{3} \simeq 67 \Rightarrow \underline{p < .001}$

c) (1) There is evidence from data that at least 2 μ's are different.

16. A simple regression model based on data with $n = 16, \bar{x} = 10, s_x = 1/\sqrt{8}$, has led to $y = 2 + 3x$, with $s_e = 4$. We are operating at 95% confidence level.

~~~2~~ 2.5 **a)** Estimate the true average change in y associated with a 1 unit increase in x, and do so in a way that conveys information about precision and reliability. You do not need to do the arithmetic.

C.I. β : $\hat{\beta} \pm t^* \frac{S_e}{\sqrt{S_{xx}}} = 3 \pm 2.15 \frac{4}{\sqrt{16-1}\,\frac{1}{\sqrt{8}}} = 3 \pm 2.15\,(4)\sqrt{\frac{8}{15}}$

.95 .025 t 2.15 $\quad$ df = n − 2 = 14 $\quad S_{xx} = (n-1)s_x^2$

$\hat{\beta}$ 0.5 + 0.5
t^* 0.5 + 0.5
S_{xx} 0.5

~~~2~~ 1 **b)** We make a single prediction at $x = 11$. What is the numerical value of $s_{\hat{y}}$? For numerical ease, assume $n - 1 \sim n = 16$.

$S_{\hat{y}} = S_e\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}} = S_e\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)S_x^2}} = 4\sqrt{\frac{1}{16} + \frac{(11-10)^2}{15(\frac{1}{8})}}$

$\simeq 4\frac{1}{\sqrt{16}}\sqrt{1+8} = \sqrt{9} = 3$

~~~3~~ 2 **c)** Suppose $s_{\hat{y}} = 3$. What is the value of T_0 such that Prob (prediction error $> T_0$) $= 0.01$? Note: you are NOT asked to compute a prediction interval.

prob(pred. err. > T_0) = .01 $\quad$ standardize

prob($\frac{\text{pred err}}{S_{\text{pred err}}} > \frac{T_0}{\sqrt{S_{\hat{y}}^2 + S_e^2}}$) = .01

prob($t > \underbrace{\frac{T_0}{\sqrt{9+16}}}_{2.6}$) = .01 $\Rightarrow$ $\underline{T_0 = 2.6\,(5)}$

2.6 ← df = n − 2 = 14 , Table $\overline{\underline{VI}}$
~~~