Name: ______________________

ID: ________________________

5+15

Stat 421, Test 1, Fall, Oct. 17, 2018; Marzban

ONLY a half-size "cheat sheet" is allowed

Multiple choice: Circle all the correct answers; there is wrong-answer penalty

For rest, SHOW answer & work; NO CREDIT for correct answer without explanation

Points

~1 **1.** For a $N(\mu, \sigma^2)$ population, what is the variance of the sampling distribution of the y-intercept of the Normal qq-plot for samples of size n?

for Normal

obs. median ~ obs. mean

a) 0 b) μ/n c) σ^2 d) σ^2/n e) The y-int has no variance.

~1 lect 5 **2.** Let UCB denote the a 95% random Upper Confidence Bound for a population mean μ. From data, we have found $\text{UCB}_{obs} = 5$. Circle the correct statement. There is a 95% probability that

a) $\mu < 5$ b) UCB < 5 c) UCB $> \mu$ d) $\bar{y} < 5$

covers μ 95% of time.

~~~1~~ lect 7, 8 **3.** Consider a problem involving $a > 2$ treatment levels. If we are interested only in comparing the mean response at 2 specific treatment levels, then the appropriate test(s) is/are

drop!

a) 1-way ANOVA F-test for only a=2

b) 2-sample t-test with $df = n_2 + n_1 - 2$

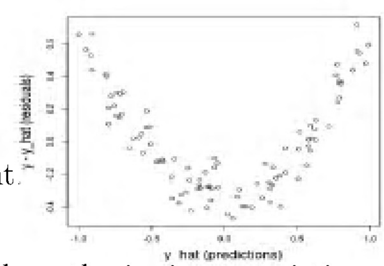
c) 2-sample t-test with df = Welch

d) 2-sample t-test with $df = an(n-1)$ lect 7 (an-a)

e) 2-sample test of a contrast with $df = an(n-1)$ lect 8

f) There exists no such test.

~1 lect 7 **4.** Consider the adjacent residual plot presented in an article. Which of the following is the most appropriate interpretation?



a) Model is adequate because the residuals are around 0.

b) There is a quadratic relation between treatment and response.

c) Model is not adequate because variance of residuals is not constant.

d) This is not a residual plot from a 1-way ANOVA model. for ANOVA, predictions are discrete.

~1 lect 8 **5.** The alternative hypothesis in the ANOVA F-test is a complex hypothesis, i.e., consisting of many simpler hypotheses. Which of the following statements is/are true?

drop → misphrased.

a) When the test is significant, one rejects all of those simpler hypotheses.

b) When the test is significant, one rejects at least one of those simpler hypotheses.

c) Contrasts allow us to test each of those simpler alternatives, with no concerns over multiple hypothesis testing (MHT). The simpler alternatives are like $\mu_1 = \mu_2 \neq \mu_3$ etc.

d) Tukey's test allows us to do all pairwise tests without any concern over MHT. These are not zero-sum contrasts.

e) None of the above.

~~~2~~ 1.5 **6.** At a given confidence level, what are the $\overline{y_{obs}}$ values such that the observed 2-sided CI for a pop mean suggests that it could be zero, but the observed LCB suggests that it cannot be zero?

$$\boxed{\begin{array}{c} \bar{Y}_{obs} - t_{\alpha/2}\frac{s}{\sqrt{n}} < 0 < \bar{Y}_{obs} + t_{\alpha/2}\frac{s}{\sqrt{n}} \\ 0 < \bar{Y}_{obs} - t_{\alpha}\frac{s}{\sqrt{n}} \end{array}}$$

$$\Downarrow$$

$$t_\alpha \frac{s}{\sqrt{n}} < \bar{Y}_{obs} < t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Moral: it's possible for a 2-sided test to fail to reject, but a 1-sided test to reject.

hw/lect3-2

~ 2 **7.** Let $\epsilon \sim f_\epsilon(t)$, $-\infty < t < \infty$. Now, suppose the r.v. ϵ is associated with some error.

a) In terms of these quantities, write an integral expression for MSE (i.e., mean squared error).

$$MSE = \int_{-\infty}^{\infty} t^2 f_\epsilon(t)\, dt$$

Mean $\int_{-\infty}^{\infty} t^2 f_\epsilon(t)\, dt$ Squared Error

~ 1 **b)** Write the integral in part a) as the expected value of something with respect to something.

The Right-hand side is, by defn., $E_\epsilon[\epsilon^2]$.

Moral: $MSE = E_\epsilon[\epsilon^2] = V_\epsilon[\epsilon] + (E_\epsilon[\epsilon])^2 = \text{Variance} + (\text{Bias})^2$

lect7,8

~ 2 **8.** For the σ_ϵ^2 parameter of our models we built the following CI that involves SSE:

$$SSE \times \left(\frac{1}{\chi^2_{\alpha/2,an-a}}, \frac{1}{\chi^2_{1-\alpha/2,an-a}}\right)$$

We also showed that when $H_0 : \Gamma = 0$ is true, then $SSC/\sigma_\epsilon^2 \sim \chi_1^2$, where $SSC = n\hat{\Gamma}^2/|\vec{c}|^2$. Build a CI for σ_ϵ^2 that involves SSC.

(It will be interesting to compare This C.I with the one with SSE.)

$$pr\left(\chi^2_{1-\frac{\alpha}{2},1} < \frac{SSC}{\sigma_\epsilon^2} < \chi^2_{\frac{\alpha}{2},1}\right) = 1-\alpha$$

$$\frac{SSC}{\chi^2_{1-\frac{\alpha}{2},1}} > \sigma_\epsilon^2 > \frac{SSC}{\chi^2_{\alpha/2,1}} \Rightarrow \text{C.I. for } \sigma_\epsilon^2 : SSC\left(\frac{1}{\chi^2_{\alpha/2,1}}, \frac{1}{\chi^2_{1-\frac{\alpha}{2},1}}\right)$$

9. Suppose we want to use SSC/MSE as a test statistic for testing some parameter. Recall that we need to know $E[SSC]$ and $E[MSE]$.

hw/lect6-8, test1-16

~ 2 **a)** Find $E[SSC]$. You may use (without proof) the following result from lectures: $V[\hat{\Gamma}] = \frac{\sigma_\epsilon^2}{n}|\vec{c}|^2$.

$$E[SSC] = E\left[n\frac{\hat{\Gamma}^2}{|\vec{c}|^2}\right] = \frac{n}{|\vec{c}|^2}E[\hat{\Gamma}^2] = \frac{n}{|\vec{c}|^2}\left(\underbrace{V[\hat{\Gamma}]}_{=\frac{\sigma_\epsilon^2|\vec{c}|^2}{n}} + \underbrace{E^2[\hat{\Gamma}]}_{=(E[\hat{\Gamma}])^2 = \Gamma^2}\right) = \sigma_\epsilon^2 + \frac{n}{|\vec{c}|^2}\Gamma^2$$

lect6, p.5

$\sim \cancel{2}$ 1.5 **b)** I'm going to tell you that $E[MSE] = \sigma_\epsilon^2$. What is the null hypothesis that SS_c/MSE can be used for testing. Rephrased: SSC/MSE is a test statistic for what H_0?

$SSC \longrightarrow E[SSC] = \sigma_\epsilon^2 + \frac{n}{|\vec{c}|^2}\Gamma^2$

$MSE \longrightarrow E[MSE] = \sigma_\epsilon^2$

$\Rightarrow H_0: \Gamma^2 = 0$

lect6, p.5

~ 1 **c)** Is the test a lower-tailed, upper-tailed, or 2-sided? Explain

Under $H_1: \Gamma^2 \neq 0$, $\frac{E[SSC]}{E[MSE]} > 1 \Rightarrow$ upper-tailed test.

hw/lect6-4, hw/lect7-1

~ 2 **10.** Consider the model $y_i = \mu + \epsilon_i$, $i = 1, \cdots n$, with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$. Find $E[\sum_i (y_i - \overline{y})^2]$. You may use (without proof) $E[\overline{\epsilon}] = 0$ and $V[\overline{\epsilon}] = \sigma_\epsilon^2/n$. $\overline{y} = \mu + \overline{\epsilon}$

$$E\left[\sum_i (y_i - \overline{y})^2\right] = \sum_i E[(y_i - \overline{y})^2] = \sum_i E[(\cancel{\mu} + \epsilon_i - \cancel{\mu} - \overline{\epsilon})^2] = \sum_i E[(\epsilon_i - \overline{\epsilon})^2]$$

$$= \sum_i \left(V[\epsilon_i - \overline{\epsilon}] + E^2[\epsilon_i - \overline{\epsilon}]\right) = n\,\sigma_\epsilon^2\left(1 + \frac{1}{n} - \frac{2}{n}\right) = \sigma_\epsilon^2(n-1)$$

$$V[\epsilon_i - \overline{\epsilon}] = \underbrace{V[\epsilon_i]}_{=\sigma_\epsilon^2} + \underbrace{V[\overline{\epsilon}]}_{=\frac{\sigma_\epsilon^2}{n}} - 2\,\underbrace{Cov[\epsilon_i, \overline{\epsilon}]}_{=\frac{1}{n}\sum_j Cov[\epsilon_i, \epsilon_j] = \frac{1}{n}V[\epsilon_i] = \frac{1}{n}\sigma_\epsilon^2}$$

$$E^2[\epsilon_i - \overline{\epsilon}] = (\underbrace{E[\epsilon_i]}_{=0} - \underbrace{E[\overline{\epsilon}]}_{=0})^2$$

(Note: iid is used in $V[\overline{\epsilon}] = \frac{\sigma_\epsilon^2}{n}$)

~2 hw-lect8-2 **11.** You are reading an article involving a 4-level factor and data from 4 (n) replications. The authors do a 2-sided test of $H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4$ and report $t^2 = 48$. They also perform a 2-sided test of $H_0 : \mu_1 = \mu_2$ and report $t^2 = 6$. Now, you know they should not be doing too many tests, so you want to perform a single ANOVA F-test. You manage to find $MSE = 1/3$, and from one of their graphs you manage to extract the sample means at the 3rd and 4th treatment levels: 1 and 0, respectively. Find their F. Hint: $t^2 = SSC/MSE$.

$H_0: \mu_1+\mu_2=\mu_3+\mu_4 \rightarrow \vec{C}_1=(1,1,-1,-1) \Rightarrow t^2 = \frac{SSC_1}{MSE} = 48 \Rightarrow SSC_1 = 48(MSE) = 16$

$H_0: \mu_1-\mu_2 \rightarrow \vec{C}_2 = (1,-1,0,0) \Rightarrow t^2 = \frac{SSC_2}{MSE} = 6 \Rightarrow SSC_2 = 6(MSE) = 2.$

$SS_{Tr} = SSC_1 + SSC_2 + SSC_3$ if we can find $\vec{C}_3$ orthogonal to $\vec{C}_1$ and $\vec{C}_2$.

Easy to show: $C_3 = (0,0,1,-1) \Rightarrow SSC_3 = n\frac{\hat{\gamma}^2}{|\vec{C}_3|^2} = 4\frac{(\bar{Y}_3-\bar{Y}_4)^2}{2} = 2(1-0)^2 = 2$

$\therefore SS_{Tr} = 16+2+2 = 20$

$F = \frac{SS_{Tr}/(4-1)}{MSE} = \frac{20/3}{1/3} = 20.$

Moral: orthogonal contrasts are wonderful! Get used to them, because we will use them a lot in what's to come.
