Name: ____________________________

ID: ________________________________

7+13

Stat 421, Test 2, Fall, Nov. 10, 2014; Marzban

ONLY a half-size "cheat sheet" is allowed

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

For questions above horizontal line, do not explain.

For questions below horizontal line,

SHOW answer & work; NO CREDIT for correct answer without explanation

Points

2 **1.** Circle the correct statements. One-way ANOVA is appropriate

a) if A=qualitative, and a scatterplot of y vs. the levels of A shows a linear relationship.

b) if A=quantitative, and a scatterplot of y vs. the levels of A shows a linear relationship.

c) if A=qualitative, and a scatterplot of y vs. the levels of A shows no ~~linear~~ relationship.

d) if A=quantitative, and a scatterplot of y vs. the levels of A shows no ~~linear~~ relationship.

0.5 points

1 **2.** Circle correct statements. A regression model for quantitative factors can be developed only

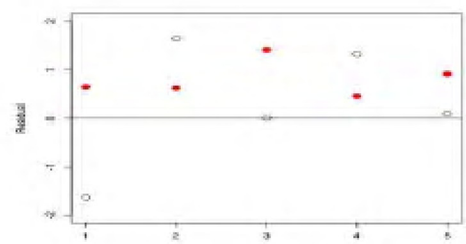
a) when there is no interaction between the factors   c) for 2^k designs

b) when there is no block factor   d) none of the above

1 4.19 **3.** An industrial engineer is conducting an experiment on eye focus time. He is interested in the effect of the distance of the object from the eye on the focus time. Four different distances are of interest. He has five subjects available for the experiment, but he suspects there may be differences among individuals. The best design for such an experiment is a

a) CRD   c) RCBD, with subject as the block

b) RCBD, with focus as the block   d) Latin Square Design



1 Lect 14 p.2 **4.** The adjacent figure shows the residual plot for a model involving a 5-level factor (A) and a binary factor (B), with the empty (filled) circles corresponding to the low (high) level of B. Based on this graph, there is evidence that the model is

a) additive   c) based on a CRD

b) includes interaction   d) based on a RCBD

The residuals are not scattered about zero.

1 hw-AB **5.** In a factorial design involving 2 factors and with $n = 1$ replication, the df for SSE is

a) 0 for additive model b) 0 for full model c) nonzero for additive model d) nonzero for full model.

1 hw **6.** In a 2^k problem, the product of the A effect and B effect is equal to the AB interaction effect if

a) it is significant   b) it is not significant   c) None of the above.

for full $df = ab(n-1) = 0$ for $n=1$

for additive $df = abn - a - b \neq 0$ $n=1$

~~~2~~ 2.5 **7.** Fill in the empty cells in the following ANOVA Table.

Source	SS	df	MS
Model	66	3	22
$contr.\ \mu_1 = \mu_2$	3	1	3
$contr.\ \mu_1 + \mu_2 = \mu_3 + \mu_4$	47	1	47
$contr.\ \mu_3 = \mu_4$	16	1	16
Error	5	15	1/3
Total	71	18	XX

66+5

$5 = 15\left(\frac{1}{3}\right)$

$3 = 15 + 18$

$66 = 3 \cdot 22$

$16 = 66 - (3 + 47)$

$71 = 66 + 5$

1 = SS of Contrasts
~~~

~~~2~~ 1 Lect 13 **8.** In a Latin Square Design involving 3 factors (A, B, C), each with 3 levels, we are interested in whether the factor A has an effect on response, treating factors B and C as nuisance factors. Provide an argument for why such a design also allows one to estimate the effect of B.

$$\begin{array}{c|ccc} & B_1 & B_2 & B_3 \\ \hline C_1 & A_1 & A_2 & A_3 \\ C_2 & A_2 & A_3 & A_1 \\ C_3 & A_3 & A_1 & A_2 \end{array} \Longrightarrow \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline C_1 & B_1 & B_2 & B_3 \\ C_2 & B_3 & B_1 & B_2 \\ C_3 & B_2 & B_3 & B_1 \end{array} \quad \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline C_1 & B_1 & B_3 & B_2 \\ C_2 & B_2 & B_1 & B_3 \\ C_3 & B_3 & B_2 & B_1 \end{array} = \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline C_1 & B_1 & B_2 & B_3 \\ C_3 & B_2 & B_3 & B_1 \\ C_2 & B_3 & B_1 & B_2 \end{array}$$

This LSD for testing the effect of A can be rearranged into this design, which is a standard LSD if ~~A_2, A_3~~ $C_2\ C_3$ are relabeled.

~~~3~~ 2 Lect 8 **9.** EXTRA CREDIT: Suppose we are working with a factorial design involving two factors, each with 3 levels. If we were to perform pairwise t-tests (each at significance level α), to see if any of the factors have any (main) effect, what is the probability of making some type I error? Show work. May be long, so keep until the end!

Tests of main effect for A: $H_0: \mu_{A_1} = \mu_{A_2}$ $\quad H_0: \mu_{A_2} = \mu_{A_3}$ $\quad H_0: \mu_{A_1} = \mu_{A_3}$

" " " " " B: $\quad B \quad B \quad\quad B \quad B \quad\quad B \quad B$

There are 6 pairwise test.

$$\text{prob}(1 \text{ type I}) = \binom{6}{1}\alpha^1(1-\alpha)^5, \quad \text{pr}(2 \text{ Type I}) = \binom{6}{2}\alpha^2(1-\alpha)^4, \ldots$$

$$\therefore \text{prob}(\text{some Type I}) = \cancel{1-} \sum_{i=1}^{6}\binom{6}{i}\alpha^i(1-\alpha)^{6-i} = 1-(1-\alpha)^6$$

There are various levels of complexity that can be added to this solution, but this is all I was looking for. Moral: More factors ⇒ More Type I

~3 **10.** Consider the following model for a Latin Square design involving three 3-level factors: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}$. Derive the least-square equation (i.e., the Normal equation) that is obtained by differentiating SSE with respect to α_i. Make sure the answer is as simplified as possible. Show work! Hint: consider i=1, first.

restricted sum $\Longrightarrow \sum'_{ijk} = \sum$ over $\left\{\begin{array}{ccc} 111 & 122 & 133 \\ 212 & 223 & 231 \\ 313 & 321 & 332 \end{array}\right\}$

$$SSE = \sum'_{ijk} \epsilon_{ijk}^2 = \sum'_{ijk}\left(y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k\right)^2$$

$$\begin{array}{c|ccc} & B_1 & B_2 & B_3 \\ \hline A_1 & C_1 & C_2 & C_3 \\ A_2 & C_2 & C_3 & C_1 \\ A_3 & C_3 & C_1 & C_2 \end{array}$$

$$SSE = (y_{111} - \mu - \underline{\underline{\alpha_1}} - \beta_1 - \gamma_1)^2 + (y_{122} - \mu - \underline{\underline{\alpha_1}} - \beta_2 - \gamma_2)^2 + (y_{133} - \mu - \underline{\underline{\alpha_1}} - \beta_3 - \gamma_3)^2$$
$$+ (y_{212} \cdots \alpha_2 \cdots)^2 + (y_{223} \cdots \alpha_2 \cdots)^2 + (y_{231} \cdots \alpha_2 \cdots)^2$$
$$+ (y_{313} \cdots \alpha_3 \cdots)^2 + (y_{321} \cdots \alpha_3 \cdots)^2 + (y_{332} \cdots \alpha_3 \cdots)^2$$

$$\frac{\partial SSE}{\partial \alpha_1} = -2\left[(y_{111} + y_{122} + y_{133}) - 3\mu - 3\alpha_1 - \beta_\cdot - \gamma_\cdot\right.$$

Similarly $$\frac{\partial}{\partial \alpha_2} = -2\left[(y_{212} + y_{223} + y_{231}) - 3\mu - 3\alpha_2 - \beta_\cdot - \gamma_\cdot\right]$$
$$\frac{\partial}{\partial \alpha_3} = -2\left[(y_{313} + y_{321} + y_{332}) - 3\mu - 3\alpha_3 - \beta_\cdot - \gamma_\cdot\right]$$

$$\therefore \left.\frac{\partial}{\partial \alpha_i} SSE\right|_{\text{hat}} = \underline{-2\left[\sum'_{jk} y_{ijk} - 3\hat{\mu} - 3\hat{\alpha}_i - \hat{\beta}_\cdot - \hat{\gamma}_\cdot\right] = 0}$$

Moral: If you continue with this, and introduce constraints $\hat{\alpha}_\cdot = \hat{\beta}_\cdot = \hat{\gamma}_\cdot = 0$, then $\hat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}$ with restricted sums.

~2 Lect 11 **11.** Let $R(A_1, A_2, ...)$ denote the reduction in SS from including the factors $A_1, A_2, ...$ in a model. Also, let $R(B_1, B_2, ...|A_1, A_2, ...)$ denote the reduction in SS from including the factors $B_1, B_2, ...,$ given that the factors $A_1, A_2, ...$ were already included in the model. Show that $R(B|\mu, A) = R(A, B|\mu) - R(A|\mu)$

$$R(B|\mu,A) = \underline{R(\mu,A,B)} - R(\mu,A)$$
$$R(A,B|\mu) = \underline{R(\mu,A,B)} - R(\mu)$$

} Eliminate $R(\mu,A,B)$

$$\therefore R(B|\mu,A) = R(A,B|\mu) + \underbrace{R(\mu) - R(\mu,A)}_{-R(A|\mu)}$$

$$\underline{R(B|\mu,A) = R(A,B|\mu) - R(A|\mu)}$$

Moral: For any model, The reduction in Variance can be computed from The reduction in Variance of 2 other models: a full model + a reduced model.

~~2~~ Lect 14 p. 2 **12.** In a 2^2 design with n replication, for an additive model find the sum of all the predictions; write your answer in the $(1), a, b, ab$ notation.

2.5

$$\hat{Y}_{ijk} = \bar{Y}_{i..} + \bar{Y}_{.j.} - \bar{Y}_{...}$$

$$\sum_{ijk} \hat{Y}_{ijk} = \sum_{ijk} \left(\frac{1}{bn} Y_{i..} + \frac{1}{an} Y_{.j.} - \frac{1}{abn} Y_{...} \right)$$

$$= \frac{1}{\cancel{bn}} \left(\sum_i Y_{i..}\right) \cancel{\sum_{jk} 1} + \frac{1}{\cancel{an}} \left(\sum_j Y_{.j.}\right) \cancel{\sum_{ik} 1} - \frac{1}{\cancel{abn}} Y_{...} \cancel{\sum_{ijk} 1}$$

$$= Y_{...} + \cancel{Y_{...}} - \cancel{Y_{...}}$$

$$\underline{= (1) + a + b + ab}$$

[Moral: Sum of predictions in additive models is The grand sum. Same as in The full model.]

Alternatively: $\hat{Y}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$

$$\therefore \sum_{ijk} \hat{Y}_{ijk} = abn\,\hat{\mu} + bn\,\underset{\overset{\parallel}{0}}{\hat{\alpha}_.} + an\,\underset{\overset{\parallel}{0}}{\hat{\beta}_.}$$ By constraints

$$= Y_{...}$$

$$\underline{= (1) + a + b + ab.}$$