Name: ______________________
ID: ________________________

Stat 421, Test 3, Fall, Dec. 9, 2014; Marzban

4 + 25

Same deal as before ...

Points

1 **1.** Which of the following defining relations has undesirable consequences?
a) ABC = CDE = 1 b) ABC = BCE = 1 c) ABD = ACE = 1 d) ABD = BCDE = 1

A = E

1 **2.** Suppose a 2^{6-2} design has been completed, and now a full fold-over is under consideration. The complete plus/minus table for the combined experiment is that of which design?
a) 2^{6-1} b) 2^{6-2} c) 2^{7-1} d) 2^{7-2}

without block factor with block factor

1 **3.** Starbucks prides itself in providing uniformly high-quality espresso drinks, as measured by a single quantity called Q. In fact, the variance of Q is known to be 10 (in some units). But the company statisticians suspect there is one source of variability (and only one), namely variability across stores. So they measure the coffee quality at 3 Starbucks stores. A random effects model suggests that 60% of the variability is due to "error," i.e. NOT due to Store. Which statement is TRUE? The variability in Q can be reduced
a) to 6, if store variability is eliminated. b) to 4, if store variability is eliminated.
c) but the reduced value cannot be determined from information given.

1 **4.** Suppose you want to know if the gas-mileage of your car varies with the station from which you buy your gas, and on the type of gas you buy. So, you collect data from a sample of stations, and the three types of gas all stations provide. A mixed-effects model is developed, and it is found that the p-value for gas Type is below α, but the p-value for Station is not. Which conclusion(s) is/are correct? Gas-mileage varies across
a) three gas Types. b) three gas Stations. or c) all gas Types. d) all gas Stations.

0.5 penalty.

5. We have a problem involving three binary factors, labeled A, B, C. For the following situations. Write the most appropriate model (w/o interactions), **clearly** labeling the indices **and** their values.

~ 2 **a)** We have n=4 replicates, and have performed all ($n\, 2^k$) runs under homogeneous conditions (e.g., on one day).

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijkl}$$

1 $i, j, k = 1 - 2$

$l = 1 - n$

0.5 Is there any confounding? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
0.5 Is there Aliasing? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.

~ 2 **b)** We have n=4 replicates, but have performed 2^k runs on each of 4 days.

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \epsilon_{ijkl}$$

$i, j, k = 1 - 2$

$l = 1 - 4$

Is there any confounding? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
Is there Aliasing? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.

~ 2 c) We have $n = 1$, and have performed all 2^k runs on 1 day.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}$$
$$i, j, k = 1-2$$

Is there any confounding? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
Is there Aliasing? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.

~ 2 d) With $n = 1$, we perform half of the 2^k runs on one day, and the other half on a second day.

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \epsilon_{ijkl}$$
$$i, j, k = 1-2$$
$$l = 1-2$$

Is there any confounding? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
Is there Aliasing? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
Also ok, if fold over

~ 2 e) With no replication, we perform only half of the 2^k runs on one day. For this case, do not write the model equation; just answer the two questions.
Is there any confounding? Yes/No If fold over
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
Is there Aliasing? Yes/No
Reason: -No blocks. -Not a fractional design. -Each block is complete. -Each block is incomplete.
- This is a fractional design.

6. Consider a 2^3 design in 4 blocks. The principal block contains the run a (in $(1), a, b, c$ notation).
~ ~~3~~ 2.5 a) Construct the blocks.

8 runs in blocks ⇒ 2 runs per block

((1), a)
(b, ab)
(c, ac)
(bc, abc)

Factorial Effect

	I	A	B	AB	C	AC	BC	ABC	AB = D	AC = E
(1)	+	−	$\boxed{−}$	+	$\boxed{−}$	+	$\boxed{−}$	−	+	−
a	+	+	$\boxed{−}$	−	$\boxed{−}$	−	$\boxed{−}$	+	−	−
b	+	−	$\boxed{+}$	−	$\boxed{−}$	+	$\boxed{−}$	+	−	+
ab	+	+	$\boxed{+}$	+	$\boxed{−}$	−	$\boxed{−}$	−	+	−
c	+	−	$\boxed{−}$	+	$\boxed{+}$	−	$\boxed{−}$	+	+	−
ac	+	+	$\boxed{−}$	−	$\boxed{+}$	+	$\boxed{−}$	−	−	+
bc	+	−	$\boxed{+}$	−	$\boxed{+}$	−	$\boxed{−}$	−	−	−
abc	+	+	$\boxed{+}$	+	$\boxed{+}$	+	$\boxed{−}$	+	+	+

for problem 7b

~ ~~2~~ 1.5 b) Explain which effect(s) is/are confounded with Blocks. Consult the above plus/minus Table.
The runs in block ((1), a) have the same +/- value for B, C, and BC
" (b, ab) " " " "
Etc.

∴ The effects B and ~~C are~~ confounded with Block.
or " B " ~~BC~~ or C and BC.

∴ Each of B, C, and BC are confounded with block.

with 4 blocks, There are 3 block effects, each confounded with 1 effect.

~ 2 **7.** Consider a 2^{5-2} design with the defining relations ABD = ACE = 1.

a) Write the alias structure for the factors D and E. (You may write the entire alias structure, but all I need is the ones that say D = ... = ..., and E = ... = ...)

$BCDE = 1$

$ABCDE = A$

$AC = ABDE$

$AB = ACDE$

$ACE = 1 \Rightarrow E = AC = BDC = ABDE$

$ABD = 1 \Rightarrow D = AB = CEB = ACDE$

~ 2 **b)** Suppose it is necessary to run the experiment in two blocks, and that we would like the effect E to get confounded with Blocks. Construct the blocks (in any notation you like). Hint: You need only a 2^3 table, which you already have.

Look at the A, B, C part of table on prev. page, and generate the D and E columns from D = AB, E = AC.

$E = -1 \Rightarrow \{a, abd, cd, bc\}$

$E = +1 \Rightarrow \{de, be, ace, abcde\}$

2 blocks.

~ 1 **c)** Which other effect(s) is/are confounded with block?

Because E is confounded with block, so are AC, BCD, and ABDE. Partial

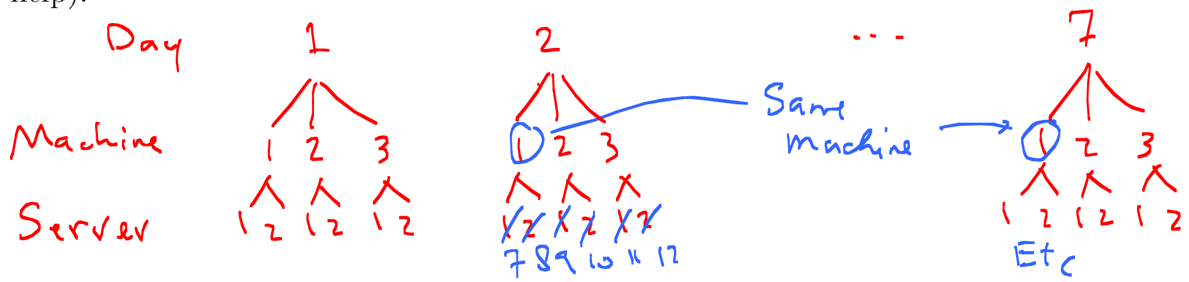
Technically, only 1 effect is conf. with block (because with 2 blocks, There is only 1 block effect), and That is E + AC + BCD + ABDE.

~ 3 **8.** Consider the random-effects model $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, where $i = 1, ...a$; $j = 1, ..., n$, and $\sigma_y^2 = \sigma_\alpha^2 + \sigma_\epsilon^2$. Compute the expected value of SS_{Total}. Hint1: The answer depends only on σ_y^2; in other words, forget the model completely! Hint2: use the formula $E[X^2] = V[X] + E^2[X]$. Hint3: $Cov[y_{ij}, y_{kl}] = \sigma_y^2$ if $i = k, j = l$, and zero otherwise.

This suggestion made sense only in The context of The quarter when This problem was assigned. In general, like I keep saying, when doing E[ ] problems, use The model as soon as possible. See The solns at The end, below.

~~3~~ 2

9. The manager of a coffee shop suspects that the quality of the coffee served (measured by a single quantity called Q) at his store varies across the three machines in the store, and the many servers who work at the store. He also suspects that the quality varies by the day of the week; for example, on Sundays quality appears to be generally worse than on Saturdays; Etc. On any day, there are two servers operating each machine, but the servers change across days. He collects data for one week. Describe the design for this experiment in terms of crossed and/or nested factors (a diagram may help).



∴ Day and Machine are crossed, but Server is nested under Machine.

FYI Day " fixed, Machine " fixed, Server " random

EXTRA CREDIT/ TIME PERMITTING / NOT REQUIRED

~~2~~ 1

10. On a sample test you saw how to use MSE tables of the type shown here to determine what is the appropriate F-ratio. Here, write down the expression for the ANOVA estimator for σ^2_τ. Note that an estimator must be a function of sample statistics, not a function of expected values.

Model Term	Factor	Expected Mean Squares
τ_i	A, main effect	$\sigma^2 + cn\sigma^2_{\tau\beta} + bn\sigma^2_{\tau\gamma} + n\sigma^2_{\tau\beta\gamma} + bcn\sigma^2_\tau$
β_j	B, main effect	$\sigma^2 + cn\sigma^2_{\tau\beta} + an\sigma^2_{\beta\gamma} + n\sigma^2_{\tau\beta\gamma} + acn\sigma^2_\beta$
γ_k	C, main effect	$\sigma^2 + bn\sigma^2_{\tau\gamma} + an\sigma^2_{\beta\gamma} + n\sigma^2_{\tau\beta\gamma} + abn\sigma^2_\gamma$
$(\tau\beta)_{ij}$	AB, two-factor interaction	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta}$
$(\tau\gamma)_{ik}$	AC, two-factor interaction	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + bn\sigma^2_{\tau\gamma}$
$(\beta\gamma)_{jk}$	BC, two-factor interaction	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + an\sigma^2_{\beta\gamma}$
$(\tau\beta\gamma)_{ijk}$	ABC, three-factor interaction	$\sigma^2 + n\sigma^2_{\tau\beta\gamma}$
ϵ_{ijkl}	Error	σ^2

$E[MS_A + MS_{ABC}] = \underset{\text{I}}{\cdots} + bcn\,\sigma_\tau^2$

$E[MS_{AB} + MC_{AC}] = \cdots$

$\therefore\ \hat{\sigma}_\tau^2 = \frac{1}{bcn}\left[MS_A + MS_{ABC} - MS_{AB} - MS_{AC}\right]$

No E[ ].

Extra Space:

Here is the "best" soln: use the model!

$$E[SS_T] = E\Big[\sum_{ij}(y_{ij}-\bar{y}_{..})^2\Big] = E\Big[\sum_{ij}(\cancel{\mu}+\tau_i+\epsilon_{ij}-\cancel{\mu}-\bar{\tau}_{.}-\bar{\epsilon}_{..})^2\Big]$$

$$= \sum_{ij} E\Big[\big((\tau_i-\bar{\tau}_{.})+(\epsilon_{ij}-\bar{\epsilon}_{..})\big)^2\Big]$$

$$= \sum_{ij}\Big\{V\big[(\tau_i-\bar{\tau}_{.})+(\epsilon_{ij}-\bar{\epsilon}_{..})\big] + \overbrace{E^2\big[\tau_i-\bar{\tau}_{.}+\epsilon_{ij}-\bar{\epsilon}_{..}\big]}^{0\,=\,E[\tau_i]\,=\,E[\epsilon_{ij}]}\Big\}$$

$$= \sum_{ij}\Big\{V[\tau_i-\bar{\tau}_{.}] + V[\epsilon_{ij}-\bar{\epsilon}_{..}] + 2\,\cancel{Cov[\tau,\epsilon]}\Big\} \rightarrow 0 \text{ because } \tau_i \text{ and } \epsilon_{ij} \text{ are indep.}$$

$$V[\tau_i-\bar{\tau}_{.}] = \underbrace{V[\tau_i]}_{\sigma_\tau^2} + \underbrace{V[\bar{\tau}_{.}]}_{\sigma_\tau^2/a} - 2\underbrace{Cov(\tau_i,\bar{\tau}_{.})}_{\frac{1}{a}\sum_j Cov[\tau_i,\tau_j] \,=\, \frac{1}{a}V[\tau_i] \,=\, \frac{\sigma_\tau^2}{a}}$$

$$V[\epsilon_{ij}-\bar{\epsilon}_{..}] = \underbrace{V[\epsilon_{ij}]}_{\sigma_\epsilon^2} + \underbrace{V[\bar{\epsilon}_{..}]}_{\sigma_\epsilon^2/an} - 2\underbrace{Cov[\epsilon_{ij},\bar{\epsilon}_{..}]}_{\frac{1}{an}\sum_{kl} Cov[\epsilon_{ij},\epsilon_{kl}] \,=\, \frac{1}{an}\sigma_\epsilon^2}$$

$$= an\Big\{\sigma_\tau^2\Big(1+\frac{1}{a}-\frac{2}{a}\Big) + \sigma_\epsilon^2\Big(1+\frac{1}{an}-\frac{2}{an}\Big)\Big\} = \underline{(an-1)\sigma_\epsilon^2 + n(a-1)\sigma_\tau^2}$$

That's good enough, but if you like you can use $\sigma_Y^2 = \sigma_\epsilon^2 + \sigma_\tau^2$ to replace either σ_ϵ^2 or σ_τ^2 in terms of σ_Y^2. E.g.

$$= \begin{cases} (an-1)\sigma_\epsilon^2 + n(a-1)[\sigma_Y^2-\sigma_\epsilon^2] = n(a-1)\sigma_Y^2 + (n-1)\sigma_\epsilon^2 \\ (an-1)[\sigma_Y^2-\sigma_\tau^2] + n(a-1)\sigma_\tau^2 = (an-1)\sigma_Y^2 - (n-1)\sigma_\tau^2 \end{cases}$$

Also, because $s_Y^2 = \frac{1}{an-1}SST$, it follows that

$$E[s_Y^2] = \sigma_Y^2 - \frac{n-1}{an-1}\sigma_\tau^2$$

ie. s_Y^2 is <u>not</u> an unbiased estimator of σ_Y^2.