## STAT 642 - SOLUTIONS TO FINAL EXAMINATION - May 5, 2014

**Problem I.** (20 points) For the following experiment, provide the requested information:

An entomologist is studying the effect of three types of insecticides (T1, T2, T3) on the yield of soybean crops. Five varieties of soybean (V1, V2, V3, V4, V5) were randomly selected from the 97 most widely used soybean varieties. Four large fields were randomly selected for the study. Within each field, five similar 1-acre areas were identified and a single variety of soybeans was randomly assigned to each of these areas resulting in each field having all five varieties present. Each of the 20 areas was then divided into 6 regions with 2 regions randomly assigned to each of the three types of insecticide. The total yield of soybeans was obtained for each of the 120 regions. A measure of soil permeability and soil pH was measured for each region prior to planting the soybeans. The entomologist's major goal for the experiment was to determine the interrelationships between varieties of soybeans and types of insecticides on the yield of soybeans.

1. Type of Randomization:
   - **RCBD with a Split-Plot Treatment Assignment; Blocks = Fields**
2. Type of Treatment Structure:
   - **Whole Plot Factor-Variety of Soybean; Split Plot Factor-Type of Insecticide;**

     $5 \times 3$ **factorial: Variety of Soybean is Crossed with Amount of Insecticide**
3. Identify each of the Factors as being Fixed or Random:
   - **Field - 4 Random levels; Insecticide - 3 Fixed levels; Variety - 5 Random levels**
4. Describe the Experimental Units and Measurement Units:
   - **Whole Plot EU for Variety is Area in Field;**

     **Split-Plot EU for Insecticide is Region in an Area;**

     **MU is Region in an Area**
5. Describe the Measurement Process:
   - **Response Variable is yield of soybeans; Covariate is soil permeability and soil pH;**

     **There is no SubSampling, no Repeated Measures**

**Problem II.** (35 points)

1. Complete the following AOV table by entering values for degrees of freedom and expected mean squares.

| SOURCE | DF | **MS** | **Expected Mean Square (EMS)** |
|---|---|---|---|
| T | 3 | 3.79 | $\sigma_e^2 + 2\sigma_{T*R(A)}^2 + 60Q_T$ |
| A | 2 | 13.27 | $\sigma_e^2 + 2\sigma_{T*R(A)}^2 + 8\sigma_{R(A)}^2 + 80Q_A$ |
| T*A | 6 | 1.78 | $\sigma_e^2 + 2\sigma_{T*R(A)}^2 + 20Q_{T*A}$ |
| R(A) | 27 | 2.58 | $\sigma_e^2 + 2\sigma_{T*R(A)}^2 + 8\sigma_{R(A)}^2$ |
| T*R(A) | 81 | 1.06 | $\sigma_e^2 + 2\sigma_{T*R(A)}^2$ |
| ERROR | 120 | 0.91 | $\sigma_e^2$ |

2. The model for the bitterness rating: $B_{ijkm}$, i=Age; j=Cheese Type; k=Rater; m=Container

- $B_{ijkm} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + c_{k(i)} + d_{jk(i)} + e_{ijkm}; \quad i = 1, \cdots, 4; \quad j = 1, \cdots, 3; \quad k = 1, \cdots, 10; \quad m = 1, 2$ requires the following conditions:

  $\tau_4 = 0, \quad \beta_3 = 0, \quad (\tau\beta)_{i3} = (\tau\beta)_{4j} = 0, \quad c_{k(i)}\ iid\ N(0, \sigma^2_{R(A)}, \quad d_{jk(i)})\ iid\ N(0, \sigma^2_{T*R(A)}), \quad e_{ijkm}\ iid\ N(0, \sigma^2_e)$ and $c_{k(i)}, \quad d_{jk(i)}), \quad e_{ijkm}$ are mutually independent.

3. At the $\alpha = 0.05$ level, evaluate the effect of the Rater's Age on the average bitterness rating of cheese. Note that the numbers given in the AOV table are the Mean Squares (MS) NOT the Sum of Squares (SS).

- First test for a Cheese Type by Age Interaction: $H_o : Q_{T*A} = 0$ vs $H_1 : Q_{T*A} \neq 0$.

  Using the test statistic: $F_{T*A} = \frac{MS_{T*A}}{MS_{T*R(A)}} = \frac{2.78}{1.06} = 2.62 > 2.21 = F_{.05,6,81}$

  Thus, conclude there is significant evidence of a Cheese Type by Age of Rater Interaction. Thus, the level of difference in the mean bitterness ratings across the three age groups depends on which Type of Cheese is being rated. Therefore, a test for a main effect due to Age Groups would be meaningless because it would be averaged over the age of the raters. An interaction between Age of Raters and Type of Cheese may result in the following situation: the younger group of raters find there is no difference in the four types of cheese, the middle age group may find Cheese Type 1 more bitter than Cheese Type 2, whereas the older group of raters may find that Cheese Type 1 is much less bitter than Cheese Type 2. Thus, an overall conclusion about the difference in mean ratings for the three age groups, ignoring type of cheese, would be meaningless. The meaningless test for a main effect of Age is as follows:

  The test for Main Effect of Age: $F_A = \frac{MS_A}{MS_{R(A)}} = \frac{13.27}{2.58} = 5.14 > 3.35 = F_{.05,2,27}$

4. Estimate the standard error of the estimated difference in the average ratings between age groups 20-29 and 50-70.

- Using the numeric values of the MS's given above and your EMS's from the AOV table,

$$
\begin{aligned}
Var(\bar{Y}_{1...} - \bar{Y}_{3...}) &= Var[\bar{c}_{.(1)} - \bar{c}_{.(3)}] + Var[\bar{d}_{..(1)} - \bar{d}_{..(3)}] + Var[\bar{e}_{1...} - \bar{e}_{3...}] \\
&= \frac{2\sigma^2_{R(A)}}{10} + \frac{2\sigma^2_{T*R(A)}}{(4)(10)} + \frac{2\sigma^2_e}{(10)(4)(2)} = \frac{2}{80}[8\sigma^2_{R(A)} + 2\sigma^2_{T*R(A)} + \sigma^2_e] = \frac{2}{80}[EM_{R(A)}] \Rightarrow
\end{aligned}
$$

$$\widehat{SE}(\bar{Y}_{1...} - \bar{Y}_{3...}) = \sqrt{\tfrac{1}{40} MS_{R(A)}} = \sqrt{2.58/40} = .2540$$

**Problem III.** (9 points) The quality control engineer selects six factors, A, B, C, D, E, F to investigate with respect to their impact on the level of an ingredient in a product. It was decided to use 2 levels (L or H) of each of these factors. The study design was a fractional factorial with 16 runs using the generators

$I_1 = BCD = +$ and $I_2 = ACE = +$ to select the 16 treatments to be used in the study.

1. For each of the following treatments, check YES if the treatment will appear in the experiment, otherwise check NO.

   i. $(A, B, C, D, E, F) = (H, L, H, H, L, H)$

   - $BCD = (-1)(+1)(+1) = -1 \quad \Rightarrow$ **No**

   ii. $(A, B, C, D, E, F) = (H, L, H, L, H, H)$

   - $BCD = (-1)(+1)(-1) = +1; \quad ACE = (+1)(+1)(+1) = +1; \quad \Rightarrow$ **Yes**

2. What is the resolution of this design? Justify your answer.

- The implicit generator is (BCD)*(ACE) = ABDE. Thus, the length of the shortest generator is 3 which implies that the Resolution is III.

3. What effects which must be assumed to be negligible in order to determine whether there is an interaction between Factors A and B using the data from this experiment.

- The effects which are confounded with A*B are

  AB*BCD = ACD;  AB*ACE = BCE;  AB*ABDE = DE.

  Thus, the following interactions must be assumed to be negligible in order to estimate A*B :

  A*C*D, B*C*E, D*E

## Problem IV. (36 points) **Place** one of (A, B, C, D, E) corresponding to the **BEST** answer

1. E
2. C because R.E. $= \frac{9}{11} + \left(\frac{2}{11}\right)\left(\frac{34.2}{11.4}\right) = \frac{15}{11} \quad \Rightarrow \quad r = \left(\frac{15}{11}\right) 9 = 12.4$
3. E because $\alpha_{PC} = 1 - (1 - .05)^{1/25} = .00205$
4. D
5. C
6. C
7. A because $F_1 * F_2$ non-significant implies that the linear trend should be examined over the levels of $F_1$ averaged over the levels of $F_2$ : $L = -3\mu_{1.} - \mu_{2.} + \mu_{3.} + 3\mu_{4.}$
8. B
9. D because $t = (3)(2) = 6, \quad r = 10, \quad D = 19, \quad \hat{\sigma}_e^2 = 100, \quad \alpha = .05 \quad \Rightarrow \quad \nu_1 = 5, \quad \nu_2 = 54$

   $\phi = \sqrt{\frac{(10)(19)^2}{(2)(6)(100)}} = 1.734$, From Table IX on page 608, the Power is approximately 0.8966.
10. D because of the carryover effect contaminates the responses in all the time periods after the first time period.
11. C because differing slopes imply that the size of the differences in the treatment means depends on the value of the covariate.
12. D because $r = 4, \quad \tau = \frac{\sigma_A^2}{\sigma_e^2} = 1.5625, \quad \alpha = .01 \quad \Rightarrow \quad \nu_1 = t - 1, \quad \nu_2 = 3t$

    $\lambda = \sqrt{1 + 4\tau} = 2.69$, From Table X on pages 616-620, the Power at t=7 is approximately 0.78 and at t=9 is approximately .88

```
FINAL EXAM SCORES      N  = 62           Distribution of Grades for STAT 642

   3   9                                       A - 24
   4   1                                       B - 28
   5   66                                      C -  9
   6   2279                                    D -  0
   7   01244557                                F -  1
   8   0000111122223344555666789
   9   000012233455666678899
  10   0
```