Review for Final Exam

Exam Details
Date: Thursday, May 5, 2011
Time: 8:00-10:00 AM
Location: ME 156
Content: Old Material (appx. 35-40%) – Chapters 1-3, 6, 7, 9, 15-18 (KNNL)

New Material (appx. 60-65%) - Chapters 19-25 (KNNL) – Lectures 26-35

The exam will test conceptual ideas. You will need to be able to do simple calculations and you will need to be able to read and interpret SAS output. You will not need to prove anything, so theory is important only with regard to understanding the concepts.

The final is comprehensive, but a few topics from the old material have been taken out (Chapters 4, 5, 8, 10, 11 are not covered). The points will be weighted a bit more heavily toward the new material (see approximate percentages above).

Materials permitted
You are allowed TWO sheets (8.5 x 11 in - both sides) of handwritten notes. Calculators are allowed. BUT, laptop computers, phones, or any devices capable of wireless communication are not permitted.

Chapter 1: Linear Regression with One Predictor Variable
• Understand the simple linear regression model.
  o \( Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \) for \( i = 1,2,\ldots,n \) where \( \epsilon_i \sim N(0,\sigma^2) \)
  o You should know in the above model which components are parameters, which components are known, and which components are random variables.
  o You should know the three basic assumptions on the error terms of a linear regression model (independent, normal errors with constant variance).
  o You should know how to interpret the regression coefficients \( \beta_0 \) and \( \beta_1 \). In particular you should be aware of when \( \beta_0 \) actually has meaning (i.e. you should understand what is meant by scope of the model).

• Understand the difference between unknowns and their estimates.
  o \( (X_i,Y_i) \) are both known \( (Y_i \) is a random variable)
  o \( E(Y) \) is the mean response – essentially it is the unknown regression line which we estimate by the fitted value \( \hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i \). The mean response is the mean (center) of the probability distribution of \( Y \) corresponding to a specific value of \( X \).
Errors $\varepsilon_i$’s are unknown; Residuals $e_i$ are the difference between the observed $Y_i$ and the fitted value $\hat{Y}_i$. Residuals are used to study whether the assumptions of the model are satisfied.

The parameters $\beta_k$’s are unknown and estimated by the $b_k$’s.

The population variance $\sigma^2$ of the error terms is unknown and is estimated by the sample variance $s^2 = MSE$.

Chapter 2: Inferences in Regression Analysis

- We may estimate parameters by either point estimates or interval estimates (confidence intervals).

- The general form of a confidence interval is: Point Estimate $\pm$ Critical Value * SE(Point Estimate).

- The general form of a test statistic is: $[\text{Point Estimate} - E(\text{Point Est. under the null})]/\text{SE(Point Estimate)}$.

- Know how to compute the standard errors for the parameter estimates $b_0, b_1, \text{etc.}$, fitted values $\hat{Y}_h$, and predictions for new observations $\hat{Y}_{b(\text{new})}$. You should also be able to read these things from PROC REG output from SAS. Again, you will not be asked to calculate $SS_X$, but would need to know how to use it to calculate any of these standard errors, if it was given to you.

- You should be able to discuss the differences between the estimated variance for the mean response and that for the prediction of a new observation.

- Know how to form confidence intervals and hypothesis tests for the parameters, the mean response, and for prediction of new observations. Also know how to get these from PROC REG output. Know how to state your hypotheses, calculate test statistics, degrees of freedom, implement a decision rule, and conclusions for a hypothesis test.

- Be able to discuss the difference between a confidence interval and a confidence band. Understand the Bonferroni correction that is used to produce family confidence limits for multiple CI’s.

- Understand the ANOVA table – degrees of freedom, SS, MS, and overall F-test for model significance. Be able to read the PROC REG output (and fill in the blanks if necessary).

- Understand the General Linear Test approach. Know what is meant by Full and Reduced models. Know how to get degrees of freedom for this test.

- Understand the coefficient of determination $(r^2)$ and how to calculate it. It is a measure of the proportion of variation in $Y$ that has been explained by our model. Also understand its limitations.

- Understand the coefficient of correlation $(r)$ and how to calculate it. It is a measure of the strength of linear association between $Y$ and $X$. It can be positive or negative depending the nature of the relationship between $Y$ and $X$. 
Chapter 3: Diagnostics & Remedial Measures

- You should be able to identify ways to check the assumptions of a simple linear regression model and discuss how any assumption violations could be remedied.

- You should be able to identify problems with assumptions by examining various plots and be able to identify when the assumptions appear to be satisfied. Problems you should be able to identify and discuss remedies for include non-normality of error terms, non-constant variance, dependence of errors, outliers, and non-linear relationships.

- You should understand how transformations (particularly Box-Cox) can be used to decrease problems with assumption violations.

Chapter 6: Multiple Regression I

- Understand the use of several independent variables in the same model and how each variable accounts for a portion of the variation in the response.

- Be able to write down and/or identify the parts of a multiple regression model and interpret the regression coefficients.

- Understand the general linear model in terms of matrices.

- Be able to form / read the ANOVA table.

- Realize that the F-test in the ANOVA table tests the hypothesis that all of the predictors are insignificant (against the alternative that at least one is not).

- Be able to do inference (i.e. hypothesis testing / CI’s) on the model parameters. Know how to interpret the p-values from the SAS output and understand the variable-added-last T-tests.

- Understand the R² value – know that it represents the proportion of variation explained by the model. Also know the difference between R² and adjusted R².

- Understand simultaneous confidence intervals (Bonferroni).

- Be able to do diagnostics to assess the assumptions of the model.

Chapter 7: Multiple Regression II

- Understand the Extra SS, both in terms of formulae and Type I/II SS in SAS output (or Type III in PROC GLM).

- Know how to perform general linear tests.

- Understand the coefficients of partial determination (pcorr1 and pcorr2 in SAS).

- Understand what multicollinearity is and how to recognize it. Understand the use of a VIF in detecting multicollinearity.
Chapter 9: Building the Regression Model: Selection of Predictors and Model Validation

- Know and understand how to use the various criteria for model selection ($R^2$, Adjusted $R^2$, Mallow’s $C_p$, AIC, BIC, PRESS). Be able to compare and contrast the various methods. Know how to read the SAS output in an effort to determine the “best” (or a couple of good candidate) model(s).

- Understand the “Best” subsets algorithm approach to model selection.

- Understand the concepts and advantages/disadvantages of forward selection, backward elimination, and stepwise regression as automated selection procedures. Be able to interpret SAS output for forward stepwise regression.

- Understand when to use different model selection procedures.

- Understand the basic concept of model validation.

Chapter 15: Introduction to Design of Experiments

- Know and understand the following terminology:
  - Categorical/qualitative variable
  - Factor/ factor levels
  - Experimental vs. observational factor
  - Qualitative vs. quantitative factor
  - Treatment/Treatment combination
  - Experimental unit
  - Balanced design
  - Crossed vs. nested factors
  - Control group
  - Fixed vs. random effect
  - Randomization

- Be able to read the description of an experiment and identify any of the above terms.

Chapters 16-18: One-Way Analysis of Variance (ANOVA)

- Understand the difference between a regression analysis and an ANOVA in terms of the assumed relationship between the dependent and independent variable(s).

- Be able to understand the connection between regression and ANOVA by interpreting SAS output from a regression analysis using indicator variables for r-1 factor levels of the ANOVA.

- Be able to interpret plots (scatterplots/boxplots) to identify visual differences between groups.

- Understand the cell means model and the factor effects model (with constraints).

- Be able to compute parameter estimates for either model given a table of means.

- Understand the partitioning of SS and the ANOVA table in general. Recall that the pooled variance estimate $\hat{\sigma}^2 = \text{MSE}$.

- Be able to conduct the overall F-test for testing the equality of all factor level means and give details associated with the test.
• Be able to compute CI’s and conduct hypothesis tests for means and differences between means (which may require calculation of the standard errors)
• Be able to determine the minimum significant difference under LSD, Tukey, Bonferroni, or Scheffe methods and test for differences among the means. Understand the differences between these methods! Know how to interpret SAS output for these methods.
• Be able to set up and use linear combinations/contrasts to construct CI’s or test hypotheses about different groups of means.
• Know the three basic assumptions (on the error terms) of ANOVA. Be able to suggest diagnostics for determining whether these are satisfied as well as remedial measures if they are not satisfied (Lecture 25 material).
• Understand the differences between one vs. two-sided tests. Understand the different types of errors that can be made in testing and discuss the importance of sample size.
• Be able to determine the appropriate sample size needed given other required information either from table B.12 (which would be provided, if needed) or output from Proc Power.

Chapters 19-21, 23: Two-Way ANOVA

• Understand the notation involved for both the cell means model and the factor effects model – and why the factor effects model is advantageous when there are two factors.
• Understand the assumptions/constraints for both models and how to check.
• Know how to get parameter estimates for either model from a table of cell and marginal means.
• Understand the partitioning of SS and the ANOVA table in general.
• Understand and be able to discuss main effects and interaction. Be able to construct and interpret an interaction plot. Know the differences involved in analysis depending on whether or not interaction is present.
• Know how to conduct the hypothesis tests for the interaction and main effects (and in what order they should be considered).
• Be able to do multiple comparisons/contrasts when appropriate.
• Understand that the additive model used when either pooling an insignificant interaction in with the error or when there is only one observation per cell. Understand the concepts related to both situations and how to apply the additive model.
• Have a basic understanding of the randomized complete block design.
• Understand how things are different when there are unequal numbers of observations in the different cells. Know differences between the different types of Sums of Squares, as well as differences between MEANS and LSMEANS.

Chapter 22: Analysis of Covariance (ANCOVA)

• Understand the use of a covariate in an ANOVA model – why it is needed and what it accomplishes. Know the model and special assumptions for the covariate and the other factors, as well as how to check the assumptions.
• Be able to interpret and discuss the output/graphs for an ANCOVA model including the common slope parameter estimate and the least squares means.
• Know how to apply ANCOVA when there are one or two factors.
• Understand the basic idea behind the analyzing a difference (e.g., post-pre) and how it compares to ANCOVA.
Chapter 24: Multi-Factor ANOVA

- Be able to extend concepts from two-way ANOVA to an analysis with multiple factors. Know how to formulate the model and how the general steps for proceeding with the analysis.

Chapter 25: Random and Mixed Effects Models

- Understand the difference between random effects and fixed effects. Be able to identify factors as either random or fixed.
- Understand the single and two-factor random effects models, as well as the two-factor mixed model. Make sure you know the differences in the assumptions and analysis (looking at variances, etc) when the effects are random.
- Understand the SAS output from GLM, VARCOMP, and MIXED.
- Be able to use the expected mean squares to obtain variance parameter estimates and determine the appropriate hypothesis tests.
- Understand the intraclass correlation coefficient and its usefulness.

Review Problems for Final Exam

Review problems cover only the new material. Please see previous reviews for practice problems of the old material. Note that these questions do not provide an extensive review, but are meant to get you thinking about possible topics and questions you may see. Note that this portion of the review does not contain questions related to SAS output, which may be covered on any of these topics.

1. Suppose that you have a two-way ANOVA examining the effects of gender (A, 2 levels) and paper color (B, 3 levels = red, blue, green) on the answer to a survey question written on that paper. Further suppose that your design is balanced with three observations per cell. The estimates for the cell means are shown in the following table:

<table>
<thead>
<tr>
<th></th>
<th>BLUE</th>
<th>GREEN</th>
<th>RED</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEMALE</td>
<td>15</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>MALE</td>
<td>25</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

a. Use the estimated cell means above to produce the following estimates for the factor effects model:

\[
\hat{\alpha}_{Male} = \\
\hat{\beta}_{Green} = \\
(\alpha \beta)_{FemaleGreen} =
\]

2. Suppose that you have a two-way ANOVA examining effects A and B, each with two levels and one observation per treatment combination (A*B). Write down the model that would need to use. What is this model called? What would happen if you tried to include the interaction term?
3. In a two-factor ANOVA, factor A has 2 levels and factor B has 4 levels. There are 5 observations per treatment, except when A=2, B=1 and A=2, B=2; there are only 4 observations for these two treatments because subjects dropped out.
   a. Give the degrees of freedom for each of the following (as they would appear in the ANOVA tables):
      i. factor A
      ii. factor B
      iii. interaction A*B
      iv. error
      v. total
   b. The Type III SSA = 2276 and SSE = 5598. Calculate the F-statistic for testing for a factor A main effect, and give the degrees of freedom for that test. Also, state the null and alternative hypotheses.
   c. Will the type I and type III sums of squares be the same or different in this analysis? Give a clear answer and a brief explanation.
   d. In this particular case, would you consider the type I or type III sums of squares (and tests) to be most informative? Give a clear answer and explain your reasoning.
   e. Explain how the Means and LSMeans will be different for this case.

4. A two-way ANOVA has cell means given in the following table. Construct an interaction plot and interpret that plot.

<table>
<thead>
<tr>
<th></th>
<th>B=1</th>
<th>B=2</th>
<th>B=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A=1</td>
<td>7</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>A=2</td>
<td>5</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

5. A leading consumer magazine wishes to compare the performance of three specific brands of ear thermometer: Braun, Vicks, and Safety First. They purchase one of each brand of thermometer, with purchases made at a variety of stores. Six employees of the company are chosen at random and agree to participate. Each person’s temperature is measured in both ears, four times with each thermometer.
   a. Identify by name the response variable and all factors in this experiment.
   b. For each factor, state whether it should be considered a fixed or random effect. Justify your answers briefly.
   c. State the following:
      i. the number of levels for each factor,
      ii. the total number of treatments,
      iii. the total sample size \( n_T \), and
      iv. \( n \), the number of observations per treatment.
6. Complete the following ANOVA table assuming that Factor A (fixed) has two levels, Factor B (random) has 3 levels, and there are 3 observations for each treatment. Also test for main effects and interaction (use the F-table in the back of the book for appropriate critical F-values). Make sure to state your null and alternative hypotheses, decision rules, and conclusions. Also, give the restricted factor effects model with any assumptions.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>___</td>
<td>____</td>
<td>75</td>
<td>_____</td>
</tr>
<tr>
<td>B</td>
<td>___</td>
<td>96</td>
<td>____</td>
<td>_____</td>
</tr>
<tr>
<td>A*B</td>
<td>___</td>
<td>____</td>
<td>____</td>
<td>_____</td>
</tr>
<tr>
<td>Error</td>
<td>___</td>
<td>____</td>
<td>3</td>
<td>_____</td>
</tr>
<tr>
<td>Total</td>
<td>___</td>
<td>237</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. An ANCOVA model was fit using a factor A with two levels and a quantitative predictor X1. Explain how the LSMEANS for each factor level are computed. Write the model and state the null and alternative hypothesis for testing whether the factor and covariate are significant. Explain how one could check the assumption of equal slopes for all levels of A.

8. Explain the difference between fixed and random effects.

9. Assume in a one way random effects model you have MSA = 7, MSE = 3, and n=15. Compute and interpret the intraclass correlation coefficient.

10. Be able to extend any of the concepts from two-way ANOVA to multi-factor ANOVA. Although you did not complete a homework for this, I am not going to give a specific SAS problem, but see Lecture 33 for an example.
SAS Review Problems for Final Exam

Again, these are not designed to be comprehensive, but can help give you practice reading and interpreting the SAS output on Random and Mixed Effect Models, since you have not completed a homework on these.

1. The book describes a study on the production of imitation pearls (25.17, page 1081, KNNL). The researchers were investigating the effect of the number of coats of a special lacquer (factor A) applied to an opalescent plastic bead used as the base of the pearl on the market value of the pearl. Four batches of 12 beads (Factor B) were used. The three levels of coat (6,8, and 10) were fixed in advance, while the four batches can be regarded as a random sample of batches from the bead production process.

Below is the SAS Output from PROC GLM for the two-way mixed effects model.

```
Source                      DF         Squares     Mean Square    F Value    Pr > F
Model                       11     305.0916667      27.7356061       5.75    <.0001
Error                       36     173.6250000       4.8229167
Corrected Total             47     478.7166667

R-Square     Coeff Var      Root MSE    value Mean
0.637312      2.904593      2.196114      75.60833

Source                      DF       Type I SS     Mean Square    F Value    Pr > F
batch                        3     152.8516667      50.9505556      10.56    <.0001
coat                         2     150.3879167      75.1939583      15.59    <.0001
batch*coat                   6       1.8520833       0.3086806       0.06    0.9988

Source                      DF     Type III SS     Mean Square    F Value    Pr > F
batch                        3      152.851667       50.950556     165.06    <.0001
coat                         2      150.387917       75.193958     243.60    <.0001
Error                        6        1.852083        0.308681

Source                  Type III Expected Mean Square
batch                  Var(Error) + 4 Var(batch*coat) + 12 Var(batch)
coat                  Var(Error) + 4 Var(batch*coat) + Q(coat)
batch*coat          Var(Error) + 4 Var(batch*coat)
```

Error: MS(batch*coat)

```
Source                  DF     Type III SS     Mean Square    F Value    Pr > F
batch*coat                   6        1.852083        0.308681       0.06    0.9988
Error: MS(Error)            36      173.625000        4.822916
```
a. Write the **restricted** two-way mixed effects model with any assumptions/constraints. What is the difference between this model and the unrestricted model? Which does SAS use and how does that affect us when conducting the analysis?

b. Using the restricted model, what are the Expected Mean Squares?

c. Using the restricted model, conduct the tests for the main effects and the interaction. Give the F-statistic you are using (say which mean squares are in the numerator and denominator, as well as the value from the output), the p-values, and your conclusions.

2. Now, assume that the coat is random (just for practice – I’m not sure if it would make sense here for coat to be random unless a large number of coats could possibly be applied and these three levels were selected at random).

Below is the SAS Output from for the two-way random effects model.

**Output from GLM**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>305.0916667</td>
<td>27.7356061</td>
<td>5.75</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>173.6250000</td>
<td>4.8229167</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>47</td>
<td>478.7166667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE value Mean  
0.637312 2.904593 2.196114 75.60833

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.8516667</td>
<td>50.9505556</td>
<td>10.56</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>coat</td>
<td>2</td>
<td>150.3879167</td>
<td>75.1939583</td>
<td>15.59</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>batch*coat</td>
<td>6</td>
<td>1.8520833</td>
<td>0.3086806</td>
<td>0.06</td>
<td>0.9988</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.851667</td>
<td>50.950556</td>
<td>165.06</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>coat</td>
<td>2</td>
<td>150.387917</td>
<td>75.193958</td>
<td>243.60</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>batch*coat</td>
<td>6</td>
<td>1.8520833</td>
<td>0.3086806</td>
<td>0.06</td>
<td>0.9988</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>Var(Error) + 4 Var(batch*coat) + 12 Var(batch)</td>
</tr>
<tr>
<td>coat</td>
<td>Var(Error) + 4 Var(batch*coat) + 16 Var(coat)</td>
</tr>
<tr>
<td>batch*coat</td>
<td>Var(Error) + 4 Var(batch*coat)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.851667</td>
<td>50.950556</td>
<td>165.06</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>coat</td>
<td>2</td>
<td>150.387917</td>
<td>75.193958</td>
<td>243.60</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>173.625000</td>
<td>4.822917</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Error: MS(batch*coat)**
Output from VARCOMP

MIVQUE(0) Estimates

<table>
<thead>
<tr>
<th>Variance Component</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var(batch)</td>
<td>4.22016</td>
</tr>
<tr>
<td>Var(coat)</td>
<td>4.68033</td>
</tr>
<tr>
<td>Var(batch*coat)</td>
<td>-1.12856</td>
</tr>
<tr>
<td>Var(Error)</td>
<td>4.82292</td>
</tr>
</tbody>
</table>

da. Write the two-way random effects model with any assumptions.
b. What are the Expected Mean Squares?
c. Give the estimates of the variance components.
d. Conduct the tests for the main effects and the interaction. Give the F-statistic you are using (say which mean squares are in the numerator and denominator, as well as the value from the output), the p-values, and your conclusions. (You might also note the similarities/differences between these results and the results from the two-way mixed effects model).

3. Now, we will just consider the random batch effect (again, just for practice – it doesn’t really make sense to leave out coat since it is both significant and the main effect of interest).

Output from GLM

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>152.8516667</td>
<td>50.9505556</td>
<td>6.88</td>
<td>0.0007</td>
</tr>
<tr>
<td>Error</td>
<td>44</td>
<td>325.8650000</td>
<td>7.4060227</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>47</td>
<td>478.7166667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE value Mean
0.319295 3.599340 2.721401 75.60833

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.851667</td>
<td>50.9505556</td>
<td>6.88</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

Source | Type III SS | Mean Square | F Value | Pr > F |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.851667</td>
<td>50.950556</td>
<td>6.88</td>
</tr>
</tbody>
</table>

Source | Type III Expected Mean Square
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>Var(Error) + 12 Var(batch)</td>
</tr>
</tbody>
</table>

Source | DF | Type III SS | Mean Square | F Value | Pr > F |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3</td>
<td>152.851667</td>
<td>50.950556</td>
<td>6.88</td>
<td>0.0007</td>
</tr>
<tr>
<td>Error: MS(Error)</td>
<td>44</td>
<td>325.865000</td>
<td>7.406023</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Output from MIXED

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Estimate</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>3.6287</td>
<td>0.05</td>
<td>1.0222</td>
<td>109.33</td>
</tr>
<tr>
<td>Residual</td>
<td>7.4060</td>
<td>0.05</td>
<td>5.0757</td>
<td>11.8176</td>
</tr>
</tbody>
</table>
a. Write the one-way random effects model with any assumptions.
b. What are the Expected Mean Squares?
c. Give the estimates of the variance components, as well as the intraclass correlation coefficient.
d. Conduct the tests for batch. Give the F-statistic you are using (say which mean squares are in the numerator and denominator, as well as the value from the output), the p-values, and your conclusions. Would this test change if you considered batch to be fixed? Why or why not?