March 1st Exam

**Coverage:**
- Chapters 16, 17, 18, 19, 20, 22.

**Not included:**
- Scheffe Multiple Comparisons
- Holm Simultaneous Testing
- Analysis of Factor Effects for Quantitative Factors
- Tests for Constancy of Error Variance
- Weighted Least Squares
- Empty cells in two factor studies

You will be allowed to bring 2 pages of notes (front and back OK) into the exam. You may use a calculator, but a calculator is not necessary. You may leave all your answers as un-simplified expressions like 3.1231+12.1233.

The in class part of the exam will focus on the meanings and interpretations of, and the relations between, every number in the “summary” and “anova” regression outputs from R (except Adjusted R-squared). For example:

```r
summary(fit)
Call:
 lm(formula = sqrt(days) ~ type * sex, data = data)
```

Residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-8.0976</td>
<td>-2.6848</td>
<td>-0.2110</td>
<td>2.3659</td>
<td>8.4327</td>
</tr>
</tbody>
</table>

Coefficient estimates:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 30.3545 | 1.8268 | 16.616 | < 2e-16 *** |
| typeL | -27.9314 | 2.7905 | -10.010 | < 2e-16 *** |
| typeM | -10.0291 | 2.0560 | -4.878 | 4.2e-06 *** |
| sexM | 3.6274 | 2.0714 | 1.751 | 0.0831 . |
| typeL:sexM | 0.8687 | 2.9926 | 0.290 | 0.7722 |
| typeM:sexM | -5.9637 | 2.8018 | -2.129 | 0.0358 * |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.654 on 97 degrees of freedom
Multiple R-squared: 0.8934, Adjusted R-squared: 0.8879
F-statistic: 162.5 on 5 and 97 DF, p-value: < 2.2e-16
anova(fit)

Analysis of Variance Table

Response: sqrt(days)

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>2</td>
<td>10729.0</td>
<td>5364.5</td>
<td>401.8749</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>sex</td>
<td>1</td>
<td>24.4</td>
<td>24.4</td>
<td>1.8249</td>
<td>0.17987</td>
</tr>
<tr>
<td>type:sex</td>
<td>2</td>
<td>94.9</td>
<td>47.4</td>
<td>3.5543</td>
<td>0.03238 *</td>
</tr>
<tr>
<td>Residuals</td>
<td>97</td>
<td>1294.8</td>
<td>13.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The following are examples of the types of questions I will ask.

1. What is the model for the mean of sqrt(days) that is being estimated in the table above?
2. What is the estimated “beta0” and what does it mean?
3. What is the estimated “beta1” and what does it mean?
4. Draw a graph (by hand) with sex on the X-axis and mean sqrt(days) on the y-axis. Draw a separate line for each tissue type. On the plot, specify (with numbers) the means for each that correspond to each of the six unique levels of the combination of sex and type.
5. Is the effect of sex the same for all tissue types? Why or why not? How statistically sure are you of your conclusions?
6. What is the estimated difference in the mean between males with type=low and females with type=high?
7. Describe how to make a 95% confidence interval for what you estimated in question 6. Do you have the information you need to make this interval? If so, please make it. If not, what else do you need?
8. What is an estimate of the total variance of sqrt(days)?
9. What is an estimate of the standard deviation of the errors in the model?
10. What is the sample size?
11. Write down a general model for the mean of sqrt(days) as an additive function of the main effects of type and sex.
12. Explain specifically how the additive model you specified in question 11 is different from the model that is fit in the tables above.
13. What is mean square regression for Sex, and what could you use that number for?
14. Suppose the model that is fit in the tables above uses data where TYPE has 3 levels (L,M,H) and SEX has 2 levels (M,F). Using those data, can you estimate a model for mean sqrt(days) that assumes a more flexible relationship between the mean of sqrt(days) and the effects of sex and type? If there is a more flexible model, please suggest it. If there isn’t, why not?