Answers to Exam, Semester 1, 2000

1. (a)  

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>treat</td>
<td>2</td>
<td>1759.5</td>
<td>879.75</td>
<td>57.13</td>
</tr>
<tr>
<td>Error</td>
<td>21</td>
<td>323.4</td>
<td>15.4</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>2082.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$H_0: \mu_1 = \mu_2 = \mu_3, \mu_i =$ mean relief time for treatment $i$; $H_1$: the means are not all equal.

$F = 57.13 > F_{2,21}(.999)$, so reject $H_0$.

(b) Response variable is normally distributed; equal variance in response variable for each treatment; all observations are independent.

(c) LSD = 4.967, so all pairs of treatment means are significantly different.

2. (a) Higher oxygen concentration results in less ethanol; galactose results in more ethanol than glucose; not much interaction between type of sugar and oxygen concentration.

(b) $F = 0.441 < F_{3,8}(.95) = 4.07$, so there is no significant interaction.

(c) Under additive model, error MS = 0.0115; $F = 3.26 < F_{3,11}(.95) = 3.59$, so effect of oxygen concentration not significant.

(d) (0.094, 0.330); interval doesn’t include 0, so galactose produces significantly more ethanol than glucose.

(e) (i) 0.299. (ii) 0.130.

3. (a) Each extra pump running reduces the fish intake by 19.2 fish, provided the other variables are held constant.

(b) 0.400 or 40.0%.

(c) (62.9, 216.9).

(d) $T = 2.27 > t_{21}(.975) = 2.08$, so effect is significant.

4. (a) Higher lead content associated with greater traffic flow; a linear relationship looks appropriate.

(b) 361.8 $\mu$g/g dry weight; (124.7, 598.9).

(c) Within groups SS from ANOVA = 60040; for adequacy test, $F = 0.472 < F_{2,6}(.95) = 5.14$, so simple linear regression model is adequate.

5. (a) Enter $x_1$ ($F = 40.2$), enter $x_2$ ($F = 34.3$), don’t enter $x_1x_2$ ($F = 1.37$), so selected model is $x_1, x_2$.

(b) (i) 0.959; 0.949; 3.36; 0.387. (ii) the only two models in contention are $x_1, x_2$ and $x_1, x_2, x_1x_2$; $x_1, x_2$ is better because it is simpler, and the four statistics are only slightly worse.
6. (a) (i) \( \mu(i, x) = \mu + \alpha_i \). (ii) \( \mu(i, x) = \mu + \beta x \).
(b) (i) Interaction not significant (output 1, \( P = 0.912 \)), so parallel lines appropriate; both treatment and pretest are significant (output 2, \( P = 0.000 \) for both), so accept parallel line model: 
\( \mu(i, x) = \mu + \alpha_i + \beta x \). (ii) Both therapy and pretest score significantly affect posttest score, and the effect of therapy does not depend on pretest score. (iii) \( \mu(2, x) = 1.374 + 1.115x \).

7. (a) (i) Completely randomised. (ii) Yes, since its effect on posttest score is significant.
(iii) Group subjects with similar pretest scores into blocks, then randomise treatments within blocks.
(b) 6.

8. (a) \( Y_{i,j} = \mu + A_i + E_{i,j} \), \( A_i \overset{d}{=} \text{N}(0, \sigma^2_A) \), \( E_{i,j} \overset{d}{=} \text{N}(0, \sigma^2) \); \( Y_{i,j} \) = \( j \)th measurement from vat \( i \), \( \mu \) = overall mean, \( A_i \) = random effect of vat \( i \), \( E_{i,j} \) = random error.
(b) (i) \( H_0: \mu = 3.3 \). (ii) \( H_0: \sigma^2_A = 0 \).
(c) \( \hat{\sigma}^2 = 0.0918 \); \( \hat{\sigma}^2_A = 0.7238 \); variation between vats is much greater than variation within vats, so the manufacturing process is very inconsistent.
(d) the method for this question is not required by semester 2 students.