SOLUTION TO NOVEMBER 2002 EXAMINATION

Question 1.

(a) \( H : \beta_1 = \beta_2 = \beta_3 = 0 \)
\[
F = 50.53, \; P \approx 0 \text{ Highly significant}
\]
\( \therefore \) Evidence at least one of the variables has a relationship with daily soil evaporation.

(b) \( H_1 : \beta_{\text{Maxt}} = 0 \) given Avat and Avh are in the model.
\( t = -0.428 \; P = .67 \text{ Not significant} \)

\( H_2 : \beta_{\text{Avat}} = 0 \) given Maxst and Avh in model
\( t = 3.197 \; P = .0026 \text{ Highly significant} \)

\( H_3 : \beta_{\text{Avh}} = 0 \) given Maxst and Avat not in model
\( t = -5.79 \; P = 7.96e-07(\approx 0) \text{ Highly significant} \)

(c) From the pairs graph it does appear evaporation increased with average temperature. This with the positive sign in the multiple regression supports this contention. Similarly, evaporation appears to decrease with Avh. (negative slope in graph, negative sign in equation). However this contention that evaporation decreases with Maxst is very doubtful from the pairs graph which indicates an increase even though there is a negative sign in the multiple regression. This can be explained by correlation (which is evident from the pairs graph) between the three predictors.

The hypothesis tested is that there is insufficient ADDITIONAL information supplied by maxst (once Avat and Avh) are fitted to justify including it in the equation. It does not imply there is no relationship between maxst and evaporation. Indeed the pairs plot indicates otherwise.

(d) Several outliers seem to be possible in both graphs. If these are ignored the residual plot does appear reasonably random so the assumption of equal variance does not appear to be violated and the model fits reasonably well.

Similarly if the outliers are ignored the remainder of the plot resembles a straight line – the assumption of a normal distribution for the residuals appears reasonable.

(e) \( \text{evap} = 123.58 - 0.17 \text{ maxst} + 0.29 \text{ Avst} - 0.33 \text{ Avh} \)
\( \text{evap} = 123.58 - 0.17 \times 80 + 29 \times 200 - .33 \times 400 = 35.98 \)
Question 2.

(a) \( H : \beta_{\text{grazed}} = 0 \)
\[ F = 115.6 \quad P \approx 0 \text{ Highly significant} \]
Intercepts appear different.

(b) 95% CI
\[ \hat{\beta}_{\text{root}} \pm t_{37.975} \times se(\hat{\beta}) \]
\[ 23.56 \pm 2.03 \times 1.149 \quad \text{(or 2.02)} \]
\[ 23.56 \pm 2.33 \]
95% CI is (21.23, 25.89)
This CI does not contain zero so the slope is unlikely to be zero. The slope is significantly different from zero.

(c) Line 1 (grazed) Fruit = \(-127.829 + 23.56 \times \text{root} \)
Line 2 (ungrazed): Fruit = \((-127.829 + 36.103) + 23.56 \times \text{root} \)
\[ = -91.726 + 23.56 \times \text{root} \]

(d) If root = 9. For grazed
\[ \text{Fruit} = -127.829 + 23.56 \times 9 \]
\[ = 84.21 \text{ mg.} \]

(e) No.

Question 3.

(a) Contrasts

<table>
<thead>
<tr>
<th></th>
<th>( C \lor R )</th>
<th>( A \lor DD )</th>
<th>( \text{Dim} \lor \text{Dol} )</th>
<th>( \text{DimB} \lor \text{K} )</th>
<th>( \text{Dol} 1 \lor 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cont</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>-1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dim B</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Dol 1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>

(b) \[ 5 \times 0 + (-1 \times 4) + (-1) \times (-1) + (-1) \times (-1) + (-1) \times (-1) + (-1) \times (-1) = 0 \]
\[
\therefore \text{Orthogonal}
\]
\( H_{CVR} : \mu_i = \frac{\mu_A + \mu_{DimB} + \mu_{DimK} + \mu_{Dol1} + \mu_{Dol2}}{5} \)

OR

\[ 5\mu_C - \mu_A - \ldots = 0 \]

\[ F = 26, P \approx 0 \text{ Highly significant} \]

\[ H_{AVDD} : \quad 4\mu_A - \mu_{DimB} - \mu_{DimK} - \mu_{Dol1} - \mu_{Dol2} = 0 \]
\[ F = 0.17 \quad P = 0.68 \quad \text{Not significant} \]

\[ H_{DimB\lor Dol} : \quad \mu_{DimB} + \mu_{DimK} - \mu_{Dol1} - \mu_{Dol2} = 0 \]
\[ F = 15.749 \quad P = 0.0009 \quad \text{Highly significant.} \]

\[ H_{DimB\lor K} : \quad \mu_{DimB} = \mu_{DimK} \]
\[ F = .42 \quad P = .53 \quad \text{Not significant} \]

\[ H_{Dol1\lor Dol2} : \quad \mu_{Dol1} = \mu_{Dol2} \]
\[ F = 4.09 \quad P = .058 \quad \text{Almost significant} \]

(d) The average rice yield is higher for the treated plots although the average yield of the Azodrin treated plots is not different from those treated with Dimecron or Dolmix. Dimecron appears less effective than Dolmix with no significant difference between the two Dimension or Dolmix Treatments. However the two Dolmix treatments are verging on reaching significance with Dolmix 2 perhaps being the most effective.

Recommendation: Use Dolmix 2 (average yield 7.88) with Dolmix 1 or Azodrin being the second choice.

Question 4.

(a) \( H_{space} : \) The mean yield for different row spacings are the same when averaged over levels of variety.

\( H_{space:var} : \) The difference in yields for any two levels of spacing is the same for each variety

\[ F = 39.54 \quad P = 3.679e.05 \quad \text{which is highly significant} \]

(There is no need to test within main effect since not is significant).

(b) \( H_{lin} : \) no linear trend in yield due to spacing when averaged over variety

\( H_{quad} \) no quadratic trend (after linear removed) when averaged over variety

\( H_{lin:var} \) Linear trend is same for each variety

\( H_{quad:var} \) quadratic trend is same for each variety (after linear trend removed).
(c) \( H_{tinc:var} \ F = 21.5 \ P = 4.5e - 06 \) High significant
\( H_{quad:var} \ F = 0.18 \ P = .83 \) Not significant
\( H_{quad} \ F = 0.0008 \ P = .98 \) Not significant

(d) The linear trends are not the same for all varieties. There is however no evidence of a quadratic : var interaction or of a quadratic trend.

From the graph we see that increasing spacing increases the yield for both varieties 1 and 3 while increasing spacing decreases the yield for variety 2 (which is a clear reason for the interaction).

Variety 3 gives consistently higher yields than variety 1. However for spacing 1 we should use variety 2.

At spacing’s 2 and 3 use variety 3.

**Question 5**

(a) \( a = 3, \ b = 4, \ n = 2 \)

\[
E(MS_{operator}) = \sigma^2 + 2\sigma_{O:M}^2 + 8\sigma_O^2
\]
\[
E(MS_M) = \sigma^2 + 2\sigma_{O:M}^2 + 6\sigma_M^2
\]
\[
E(MS_{O:M}) = \sigma^2 + 2\sigma_{O:M}^2
\]
\[
E(Res) = \sigma^2
\]

(b)
\[
\hat{\sigma}^2 = 3.792
\]
\[
\hat{\sigma}_{O:M}^2 = \frac{1}{2}(7.44 - 3.792) = 1.824
\]
\[
\hat{\sigma}_M^2 = \frac{1}{6}(4.153 - 7.44) < 0 \quad \therefore \hat{\sigma}_M^2 = 0
\]
\[
\hat{\sigma}_O^2 = \frac{1}{8}(80.167 - 7.44) = 9.09
\]

(c) \[
\text{Var}(y) = 9.09 + 0 + 1.824 + 3.792 = 14.71
\]

(d) The largest source of variance is from the operators. Further training for the operators may result in reduced variation in the breaking strength of the fibres.
Question 6.

(a) Test of Fitness

\[ G = 1.365 \sim \chi^2_3 \Rightarrow P > 0.2 : \text{ Not significant. Model appears to fit.} \]

(b) \( H_{\text{sex mort}} \): Mortality of both sexes is the same (averaged over site).

\[ X^2 = 13.973 \sim \chi^2_1 \Rightarrow P = 1.855e - 104 \text{ Highly sig.} \]

\( H_{\text{site mort}} \): mortality is not affected by site of population

\[ X^2 = 8.959 \sim \chi^2_3 \Rightarrow P = .030 \text{ Significant.} \]

(c) It appears females are more resistant than males (\( \frac{93}{107} = 87\% \) to \( \frac{57}{87} = 65.5\% \))

Site does matter. We can conclude that pupa located at the margin of the medium have a better chance of survival than those on the wall of the vial. To make further comparisons we would need to have devised suitable contrasts before collecting the data.

\[
\begin{align*}
P(H|AM) & = 86\% \\
P(H|IM) & = 79.5\% \\
P(H|OM) & = 68.4\% \\
P(H|OW) & = 52.6\%
\end{align*}
\]