Stat/For/Hort 571
Final Exam, Fall 2003
Brief Solutions

1. This is a case of paired samples.

   (a) For testing $H_0 : \mu_D = 0$ versus $H_A : \mu_D \neq 0$, use $T = \frac{\bar{d}-0}{s_D/\sqrt{n}}$. Since $n = 5, \bar{d} = 8.5/5 = 1.7$, and $s_D = \sqrt{(22.56 - 5 \times 1.7^2)/4} = \sqrt{2.03} = 1.4247$, the observed $t = \frac{1.7}{1.4247/\sqrt{5}} = 2.67$. Compare with T_4, the p-value $= 2 \times P(T_4 \geq 2.67)$ is between 0.05 and 0.10. There is weak evidence of change in bacteria count.

   (b) Use the formula $n = \frac{(Z_{\alpha/2}+Z_{\beta})^2}{(\mu_A-\mu_0)^2}$. Since $\alpha/2 = 0.025, \beta = 1 - 0.95 = 0.05, Z_{0.025} = 1.96, Z_{0.05} = 1.645, \mu_A - \mu_0 = 1.5$, we have $n = \frac{(1.96+1.645)^2}{(1.5)^2} = 5.77 \uparrow 6$.

2. (a) Let p_1 denote the preference of BS1 for young leaves. Then $1-p_1$ is the the preference of BS1 for old leaves. The hypotheses are $H_0 : p_1 = 0.5$ versus $H_0 : p_1 \neq 0.5$ and the test statistic is $Z = \frac{\hat{p}_1 - 0.5}{\sqrt{0.5 \times 0.5/n_1}}$. Since $n_1 = 60, \hat{p}_1 = 38/60 = 0.633$, the observed $z = \frac{0.633-0.5}{\sqrt{0.5\times0.5/60}} = 2.06$. Compare with standard normal distribution, the p-value $= 2 \times P(Z \geq 2.06) < 0.05$. Reject H_0 at 5% and there is some evidence of a preference. Normal approximation is appropriate because $n_1\hat{p}_1 \geq 5$ and $n_1(1 - \hat{p}_1) \geq 5$.

   (b) Let p_2 denote the preference of BS2 for young leaves. The hypotheses are $H_0 : p_1 = p_2$ versus $H_0 : p_1 \neq p_2$ and the test statistic is $Z = \frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1+1/n_2)}}$. Here $n_1 = 60, \hat{p}_1 = 38/60 = 0.633, n_2 = 48, \hat{p}_2 = 18/48 = 0.375$, and $\hat{p} = \frac{38+18}{60+48} = 0.5185$, the observed $z = \frac{0.633-0.375}{\sqrt{0.5185\times0.4815\times(1/60+1/48)}} = 2.665$. Compare with standard normal distribution, the p-value $= 2 \times P(Z \geq 2.665) < 0.05$. Reject H_0 at 5%. There is evidence of a difference between the two species. Normal approximation is appropriate because $n_i\hat{p}_i \geq 5$ and $n_i(1 - \hat{p}_i) \geq 5$ for $i = 1, 2$.

3. (a) SSError $= 4\times10^2+4\times10^2+4\times11^2+4\times11^2 = 1768$ and SSAgent $= 5 \times [(20 - 34.5)^2 + (32 - 34.5)^2 + (36 - 34.5)^2 + (50 - 34.5)^2] = 2295$ (as overall mean $\bar{y}_{\cdot\cdot} = 34.5$). Hence SSTotal $=$ SSAgent $+$ SSError $= 4063$ and the ANOVA table is:

Source	df	SS	MS
Agent	3	2295	765
Error	16	1768	110.5
Total	19	4063	–

   For testing $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$, use $F = \frac{\text{MSAgent}}{\text{MSError}}$ and the observed $f = 6.923$. Compare with $F_{3,16}$, the p-value is less than 0.01. Reject H_0 at 5% and there is strong evidence of treatment differences.

   (b) The hypotheses of interest is $H_0 : \frac{1}{2}(\mu_1 + \mu_3) - \frac{1}{2}(\mu_2 + \mu_4) = 0$ versus not H_0. Use $T = \frac{\frac{1}{2}(\bar{y}_1+\bar{y}_3)-\frac{1}{2}(\bar{y}_2+\bar{y}_4)}{\sqrt{\text{MSE}}\sqrt{4\times\frac{1}{4}/n}}$. Hence the observed $t = \frac{\frac{1}{2}(20+36)-\frac{1}{2}(32+50)}{\sqrt{110.5}\sqrt{4\times\frac{1}{4}/5}} = -2.765$. Compare with T_{16}, the p-value $= 2 \times P(T_{16} \leq -2.765)$ is between 0.01 and 0.05. Reject H_0 at 5% and there is evidence of a difference between treatment 1,3 and treatment 2,4. That is, the additive A has an overall effect on milk production.

4. (a) The MSE is 120.0 and is on 36 df. Thus LSD $= t_{36,0.025}\sqrt{\text{MSE}}\sqrt{\frac{2}{n}} = 2.028 \times \sqrt{120.0} \times \sqrt{\frac{2}{10}} = 9.935$. The display is:

   20.5 32.0 38.3 48.2

   (b) An equivalent of LSD $t_{36,0.025/6}\sqrt{\text{MSE}}\sqrt{\frac{2}{n}} = 2.719 \times \sqrt{120.0} \times \sqrt{\frac{2}{10}} = 13.32$.
   The display is:

   20.5 32.0 38.3 48.2

   Bonferroni is preferred if one wants to be conservative and LSD is preferred if one wants to be liberal.

5. (a) $\hat{b}_1 = \frac{17.7457-2.09\times45.95/6}{0.9079-2.09^2/6} = 9.67$ and $\hat{b}_0 = 45.95/6 - 9.67 \times 2.09/6 = 4.29$.

   (b)

Source	df	SS	MS
Regression	1	16.82	16.82
Error	4	0.33	0.083
Total	5	17.21	–

   The observed $f = \frac{\text{MSReg}}{\text{MSErr}} = 201.5$. Compare with $F_{1,4}$, the p-value $= P(F_{1,4} \geq 203.45) < 0.01$. Reject H_0 at 5% and there is strong evidence of a linear relationship.

   (c) Since $\hat{y}_{pred} = 4.34 + 9.53 \times 0.4 = 8.158$ and $s.e.(\hat{y}_{pred})$ is $\sqrt{0.083}\sqrt{1 + 1/6 + (0.4 - 0.348)^2/(0.91 - 2.09^2/6)} = 0.313$, the 95% PI is
   $\hat{y}_{pred} \pm t_{4,0.025} \times s.e.(\hat{y}_{pred}) = 8.158 \pm 0.868$.