Stat/For/Hort 571 — Midterm II, Fall 01 — Brief Solutions

1. (a) SSError $= 7 \times 1.84 + 5 \times 1.94 + 4 \times 0.51 + 7 \times 0.38 = 27.28$. The ANOVA table is

Source	df	SS	MS
Insecticide	3	16.19	5.40
Error	23	27.28	1.19
Total	26	43.47	

We get $F = 5.40/1.19 = 4.55$. Comparing to an F-distribution with 3 and 23 degrees of freedom, we have $0.01 <$ p-value < 0.05. Hence there is moderate to strong evidence against the null hypothesis that all the insecticides have the same population mean.

(b) Here, $\hat{p}_A = 7/38 = 0.1842$ and $\hat{p}_B = 32/102 = 0.3137$. The 99% CI is given by

$$\hat{p}_A - \hat{p}_B \pm Z_{.01/2}\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{38} + \frac{\hat{p}_B(1-\hat{p}_B)}{102}}$$

Plugging in gives a CI of $-0.33 \leq p_A - p_B \leq 0.07$. Because 0 falls within the 99% CI, we might conclude that we do not reject H_0 at the 1% level. A potential limitation of this conclusion is that the correspondence between the hypothesis test and the CI is not exact for proportions.

2. (a) For non-normal data with small sample sizes, we use the Mann-Whitney test. The ranks for the two samples are

30 seconds	2	2	4	8	9	11	14
	15	16	17				
2 minutes	2	5	6	7	10	12	13

$T^* = 2 + 5 + 6 + 7 + 10 + 12 + 13 = 55$, $T^{**} = 7 \times 18 - 55 = 71$, and $T = 55$. Comparing to the table for the Mann-Whitney test with sample sizes of 7 and 10, we have p-value > 0.05. Hence there is relatively little evidence against the null hypothesis that there is no difference between the treatments in terms of the numbers of bacteria surviving.

(b) The normal scores plot for the 30 second group is curved suggesting that the data do not come from a normal distribution. The normal scores plot for the 2 minute group is roughly linear suggesting that the data may come from a normal distribution.

3. To test $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_A : \sigma_1^2 \neq \sigma_2^2$, we use Levene's test. The absolute deviations from the sample medians (after eliminating exactly one zero from the first group) used for Levene's test are

Exercise	1	3				
No exercise	5	3	1	1	2	9

The summary statistics for the absolute deviations are $\bar{x}_1 = 2$, $\bar{x}_2 = 3.5$, $s_1^2 = 2$, and $s_2^2 = 9$. The pooled standard deviation is $s_p^2 = (1 \times 2 + 5 \times 9.5)/(1 + 5) = 8.25$. We get $t = -0.64$. Comparing to a T distribution with 6 df, we have p-value > 0.2. Hence there is little evidence against the null hypothesis of equal variances.

4. Under H_0, $V^2 = (n-1) \times S^2/\sigma_0^2 \sim \chi^2_{n-1}$. We want $0.10 = P(S^2 \leq 4 | \sigma^2 = \sigma_0^2) = P(V^2 \leq 18 \times 4/\sigma_0^2)$. From tables, $0.10 = P(V^2 \leq 10.86)$. Therefore, $18 \times 4/\sigma_0^2 = 10.86$, and so It follows that $\sigma_0^2 = 18 \times 4/10.86 = 6.63$.

5. This is a paired experiment with pairing within the same cow. We assume independence of the different cows, and this is assured by the design of the experiment. If we were to use the T distribution as the basis for the confidence interval, we would also require the assumption that the data follow a normal distribution. We can check that assumption using the stem and leaf display. The stem and leaf display is multimodal, and we would conclude the normality assumption is (strongly) violated. Given the large sample size, we can use the Central Limit Theorem to provide the desired CI. The form of the 90% CI is given by

$$\bar{D} \pm Z_{.10/2}\frac{s_D}{\sqrt{N}}$$

Plugging in gives a 90% CI of $9.14 \leq \mu_D \leq 10.30$.

Grade Distribution

```
90-99|48
80-89|46
70-79|32      mean = 81.3
60-69|8       median = 85
50-59|9
  <50|5
```