Solutions to Methods in Fall 2002

1. Omitted (see problem 1 of Winter 2004).

2. (a) (Omitted detail) How many free variables means how many degrees of freedom.

   (b) (Omitted).

3. (a) The six numbers are 10.96, 16.68, 17.35, 6.41, 14.01 and 20.58.

   (b) We can choose the goodness of fit, such as G^2 or X^2. We have

$$G^2 = 2\sum_{i,j,k} n_{ijk} \log \frac{n_{ijk}}{\hat{n}_{ijk}} = 4.7865$$

   and

$$X^2 = \sum_{i,j,k} \frac{(n_{ijk} - \hat{n}_{ijk})^2}{\hat{n}_{ijk}} = 4.8483.$$

   There are 18 observations and the model has 5 parameters. Thus, based on 13 degrees of freedom, the values are not too large. Thus, we say the fit of the model is good.

4. (a) The relationship between the average yield and variance looks like a linear function. This means we probably need a square root transformation on the response.

   (b) If we use the weighted least square, then we needs to use the the group average as the weighted, i.e., we minimize $\sum w_i(y_i - \mu_i)^2$, where $w_i = 1/\hat{\mu}_i$, where μ_i is the i-th mean.

   (c) In the third method, the heterogeneous variance is ignored $\cdots$.

5. (a)

$$\hat{\eta} = \log(\frac{\hat{p}}{1-\hat{p}}) = 2.54617 - 0.11577(3) \Rightarrow \hat{\eta} = 2.19886.$$

   Thus,

$$\hat{p} = 0.9001.$$

   Note that

$$V(\hat{\eta}) = 0.40341^2 + 3^2(0.04403)^2 - 2(3)(0.40341)(0.04403) = 0.0736.$$

   Thus, the 95% confidence interval for η is

$$[2.19886 - 1.96\sqrt{0.0736}, 2.19886 + 1.96\sqrt{0.0736}] = [1.667, 2.731].$$

   and the 95% confidence interva for p is

$$[\frac{e^{1.667}}{1+e^{1.667}}, \frac{e^{2.731}}{1+e^{2.731}}] = [0.8412, 0.9388].$$

(b) The p-value is 0.14288 indicating that it is not significantly different from the placebo.

(c) Yes, since it is relatively small enough.

6. (a) When independent variable changes one unit, the dependent variable changes $\hat{\beta} = 0.10934$ units. The 95% confidence interval for estimate of the expected value of change is

$$0.10934 \pm t_{0.025,32}\sqrt{0.0000909} = [0.08992, 0.1288].$$

Another question, if considered the confidence interval of the changes, we need to consider the difference of two predicted values. It follows

$$Y_{01} - Y_{02} \sim N(\beta(X_{01} - X_{02}), 2\sigma^2).$$

$$s^2(\hat{Y}_{01} - \hat{Y}_{02}) = s^2(\hat{\beta}x) + 2MSE = 0.0000909x^2 + 1.6360.$$

Thus, we have the 95% confidence interval if independent variable changes x units is

$$0.10934 \pm t_{0.025,32}(0.0000909x^2 + 1.6360)^2.$$

Another question:

Note that the difference between the two cases is the MSE.

(b) When estimated the mean, the variance is

$$V(\hat{E}(Y)) = (\,1 \quad 4\,)\begin{pmatrix} 0.1240363 & -0.002627 \\ -0.002627 & 0.0000900 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix} = 0.10446.$$

If estimate the dependent variable, the variance is

$$V(\hat{Y}) = V(\hat{Y}) + MSE = 1.7405.$$

7. (a) The numbers of trt is 6, the number of tree is 3 and the number of trees is $24(6) = 144$. The degree of freedom for trt is 5, for field is 2 and for error term is $144 - 1 - 5 - 2 = 136$.

Source	MS	df	Expectation
Trt	$\frac{3(24)}{5}\sum_i(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2$	5	$\sigma^2 + \frac{3(24)}{5}\sum_i(T_i - \bar{T})^2$
Field	$48\sum_j(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot})^2$	2	$\sigma^2 + 24\sum_j(F_j - \bar{F})^2$
Error	$\frac{1}{136}\sum_{i,j,k}(\bar{y}_{ijk} - \bar{y}_{i\cdot} - y_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot})^2$	136	σ^2
Total	$\sum_{ijk}(y_{ijk} - \bar{y}_{\cdot\cdot\cdot})^2$	143	

(b)

$$\hat{\mu} = \bar{y}_{...} = \frac{1}{144} \sum_{i,j,k} y_{ijk}$$

and

$$V(\hat{\mu}) = \frac{\sigma^2}{144}.$$

(c) It is

$$D = \frac{T_1 + T_2 + T_3 + T_4}{4} - \frac{T_5 + T_6}{2}.$$

Thus,

$$\hat{D} = \frac{1}{4} \sum_{i=1}^{4} \bar{y}_{i..} - \frac{1}{2} \sum_{i=5}^{6} \bar{y}_{i..}$$

and

$$V(\hat{D}) = \frac{\sigma^2}{96} + \frac{\sigma^2}{48} = \frac{\sigma^2}{32}.$$

(d) It is

$$D = \frac{T_1 + T_4}{2} - \frac{T_2 + T_4}{2}.$$

Thus,

$$\hat{D} = \frac{1}{2}(\bar{y}_{1..} + \bar{y}_{4..}) - \frac{1}{2}(\bar{y}_{2..} + \bar{y}_{3..})$$

and

$$V(\hat{D}) = \frac{\sigma^2}{24}$$

(e) Since not undercut has not been done in the second year, it can not be assessed.