ANALYSIS OF COVARIANCE

Residual standard error: 6.747 on 37 degrees of freedom
Multiple R-Squared: 0.9291, Adjusted R-squared: 0.9252
F-statistic: 242.3 on 2 and 37 DF, p-value: < 2.2e-16

You know when you have got the minimal adequate model, because every row of the coefficients table has one or more significance stars (three in this case, because the effects are all so strong). In contrast to our initial interpretation based on mean fruit production, grazing is associated with a 36.103 mg reduction in fruit production.

anova(ancova2)

Analysis of Variance Table

Response: Fruit

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grazing</td>
<td>1</td>
<td>2910.4</td>
<td>2910.4</td>
<td>63.929</td>
</tr>
<tr>
<td>Root</td>
<td>1</td>
<td>19148.9</td>
<td>19148.9</td>
<td>420.616</td>
</tr>
<tr>
<td>Residuals</td>
<td>37</td>
<td>1684.5</td>
<td>45.5</td>
<td></td>
</tr>
</tbody>
</table>

These are the values we obtained the long way on p. 495.

Now we repeat the model simplification using the automatic model-simplification function called step. It couldn’t be easier to use. The full model is called ancova:

step(ancova)

This function causes all the terms to be tested to see whether they are needed in the minimal adequate model. The criterion used is AIC, Akaike’s information criterion (p. 353). In the jargon, this is a ‘penalized log-likelihood’. What this means in simple terms is that it weighs up the inevitable trade-off between degrees of freedom and fit of the model. You can have a perfect fit if you have a parameter for every data point, but this model has zero explanatory power. Thus deviance goes down as degrees of freedom in the model go up. The AIC adds 2 times the number of parameters in the model to the deviance (to penalize it). Deviance, you will recall, is twice the log-likelihood of the current model. Anyway, AIC is a measure of lack of fit; big AIC is bad, small AIC is good. The full model (four parameters: two slopes and two intercepts) is fitted first, and AIC calculated as 157.5:

Start: AIC = 157.5
Fruit ~ Grazing * Root

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum of Sq</th>
<th>RSS</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grazing: Root</td>
<td>1</td>
<td>4.81</td>
<td>1684.46</td>
</tr>
<tr>
<td>&lt;none&gt;</td>
<td></td>
<td></td>
<td>1679.65</td>
</tr>
</tbody>
</table>

Step: AIC = 155.61
Fruit ~ Grazing + Root

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum of Sq</th>
<th>RSS</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;none&gt;</td>
<td></td>
<td>1684.5</td>
<td>155.6</td>
</tr>
<tr>
<td>Grazing</td>
<td>1</td>
<td>5264.4</td>
<td>6948.8</td>
</tr>
<tr>
<td>Root</td>
<td>1</td>
<td>19148.9</td>
<td>20833.4</td>
</tr>
</tbody>
</table>
Call: 
\texttt{lm(formula = Fruit \sim Grazing + Root)}

Coefficients:

\[
\begin{array}{ccc}
\text{(Intercept)} & \text{GrazingUngrazed} & \text{Root} \\
-127.83 & 36.10 & 23.56 \\
\end{array}
\]

Then \texttt{step} tries removing the most complicated term (the Grazing by Root interaction). This reduces AIC to 155.61 (an improvement, so the simplification is justified). No further simplification is possible (as we saw when we used \texttt{update} to remove the Grazing term from the model) because AIC goes up to 210.3 when Grazing is removed and up to 254.2 if Root size is removed. Thus, \texttt{step} has found the minimal adequate model (it doesn’t always, as we shall see later; it is good, but not perfect).

**ANCOVA and Experimental Design**

There is an extremely important general message in this example for experimental design. No matter how carefully we randomize at the outset, our experimental groups are likely to be heterogeneous. Sometimes, as in this case, we may have made initial measurements that we can use as covariates later on, but this will not always be the case. There are bound to be important factors that we did not measure. If we had not measured initial root size in this example, we would have come to entirely the wrong conclusion about the impact of grazing on plant performance.

A far better design for this experiment would have been to measure the rootstock diameters of all the plants at the beginning of the experiment (as was done here), but then to place the plants in matched pairs with rootstocks of similar size. Then, one of the plants would be picked at random and allocated to one of the two grazing treatments (e.g. by tossing a coin); the other plant of the pair then receives the unallocated gazing treatment. Under this scheme, the size ranges of the two treatments would overlap, and the analysis of covariance would be unnecessary.

**A More Complex ANCOVA: Two Factors and One Continuous Covariate**

The following experiment, with Weight as the response variable, involved Genotype and Sex as two categorical explanatory variables and Age as a continuous covariate. There are six levels of Genotype and two levels of Sex.

\begin{verbatim}
Gain <- read.table("c:\temp\Gain.txt",header=T)
attach(Gain)
names(Gain)
[1] "Weight" "Sex" "Age" "Genotype" "Score"
\end{verbatim}

We begin by fitting the maximal model with its 24 parameters: different slopes and intercepts for every combination of Sex and Genotype.

\begin{verbatim}
m1<-lm(Weight~Sex*Age*Genotype)
summary(m1)
\end{verbatim}