PLANT-GROWTH EXPERIMENT

8. Two-Way ANOVA

In this section the effects on the height of plants of all possible combinations of seed type and water levels are investigated. In particular, the main effect for water is shown to be highly significant. After the effect is found to be significant, post hoc range tests and multiple comparisons are used to determine which means differ.

8.1 Plant-Growth Experiment as a Factorial Experiment
8.2 ANOVA Output for the Plant-Growth Experiment
8.3 Exploring the Interaction Effects
8.4 Multiple Comparisons

8.1 Plant-Growth Experiment as a Factorial Experiment

The plant-growth experiment is an example of a factorial experiment. A factorial experiment consists of several factors (seed, water) which are set at different levels, and a response variable (plant height). The purpose of the experiment is to assess the impact of different combinations of the levels of seed and water on plant height. Analysis of variance allows us to test the null hypothesis that seed and water have no impact on plant height. As the plant-growth experiment involved replications, so that responses are available from more than one subject at each combination of levels of seed and water, the presence of interaction can be assessed. In general, the factorial designs allow not only to examine the simultaneous impact of several factors on a response but also the interactive effects of variables.

The General Factorial Procedure available in SPSS 8.0 provides regression analysis and analysis of variance for one dependent variable by one or more factors or variables. The SPSS data file used for this study is available in the SPSS file *plant1.sav* located on the FTP server in the Stat337 directory. In the data file, variables include seed, water, and height. The two-predictor variables in this study, seed and water, are categorical, which means they should be entered as factors in the GLM General Factorial procedure.

To produce the output for your data, select *SPSS Instructions* in the problem menu now. Here, we will display and analyze the output for our data.

8.2 The ANOVA Output for the Plant-Growth Experiment

Analysis of variance allows us to test the null hypothesis that seed and water have no impact on plant height. There are four sources of variation in the experiment: the main effects of *Seed* and *Water*, the interaction effect, and the error variation. Corresponding to these four sources, there are three null hypotheses that may be tested:

1. H_0: No main effect of *Seed*
2. H_0: No main effect of *Water*
3. H_0: No interaction effect between *Seed* and *Water*

The GLM General Factorial procedure in SPSS produces the following output for the experiment:

Tests of Between-Subjects Effects

Dependent Variable: HEIGHT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	393.333[a]	11	35.758	16.503	.000
Intercept	37130.667	1	37130.667	17137.231	.000
SEED	1.333	2	.667	.308	.741
WATER	324.000	3	108.000	49.846	.000
SEED * WATER	68.000	6	11.333	5.231	.007
Error	26.000	12	2.167		
Total	37550.000	24			
Corrected Total	419.333	23			

a. R Squared = .938 (Adjusted R Squared = .881)

The table contains rows for the components of the model that contribute to the variation in the dependent variable. The row labeled *Corrected Model* contains values that can be attributed to the regression model, aside from the intercept. The sources of variation are identified as *Seed*, *Water*, *Seed*Water*, and *Error*. *Error* displays the component attributable to the residuals, or the unexplained variation. *Total* shows the sum of squares of all values of the dependent variable. *Corrected Total* (sum of squared deviations from the mean) is the sum of the component due to the model and the component due to the error.

The total number of degrees of freedom is 23 = 24-1, which is one less than the number of plants. Two degrees of freedom are associated with seed, which is one less than the number of levels of this factor, and similarly 3 degrees of freedom for water. The interaction factor Seed*Water degrees of freedom equals 6 = (3-1)(4-1), the product of the degrees of freedom associated with the two factors.

According to the output the model sum of squares is 393.333 and the error sum of squares is 26.000. The total sum of squares (corrected total) is 419.333. Notice a very small contribution of error term in the total sum of squares. The p-value of the F-test for the model is reported as zero indicating a sufficient evidence of an effect of at least one of the factors on the plant height.

The sum of squares for the seed factor is estimated to be only 1.333, an extremely small value compared to the total sum of squares. The p-value of the F-test is reported as 0.741, indicating a sufficient evidence of no effect of seed type on the plant height. Indeed, in all graphical displays and numerical summaries we found the plant seed not affected by the seed type.

The sum of squares due to water is 324.000, a very large contribution in the total sum of squares of 419.333. The value of the F-statistic is 49.846 with the corresponding reported p-value of zero. Water main effects are highly statistically significant.

The p-value of the interaction term Seed*Water is equal to 0.007, indicating a strong evidence of an interaction between the two factors. Thus, in further analysis, the water effect should be compared at each level of seed.

8.3 Exploring the Interaction Effects

To further explore the interaction effects, we examine the table of estimated marginal means and the profile plot of the same values displayed below.

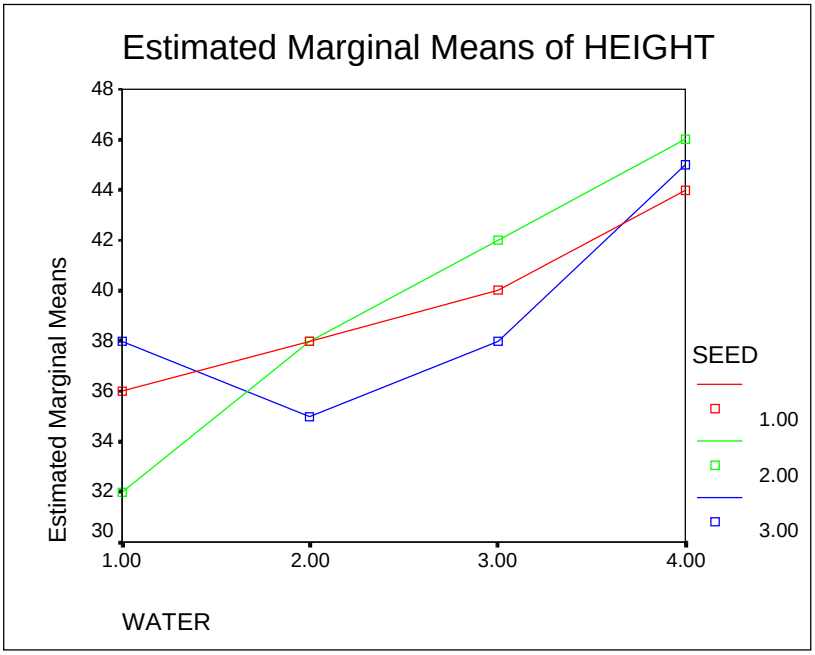
SEED * WATER

Dependent Variable: HEIGHT

SEED	WATER	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1.00	1.00	36.000	1.041	33.732	38.268
	2.00	38.000	1.041	35.732	40.268
	3.00	40.000	1.041	37.732	42.268
	4.00	44.000	1.041	41.732	46.268
2.00	1.00	32.000	1.041	29.732	34.268
	2.00	38.000	1.041	35.732	40.268
	3.00	42.000	1.041	39.732	44.268
	4.00	46.000	1.041	43.732	48.268
3.00	1.00	38.000	1.041	35.732	40.268
	2.00	35.000	1.041	32.732	37.268
	3.00	38.000	1.041	35.732	40.268
	4.00	45.000	1.041	42.732	47.268

The combination of seed 2 and water level 4 produces the highest plants. The combination of seed 2 and water level 1 produces the lowest plants. The pooled estimate of the standard deviation is 1.041.

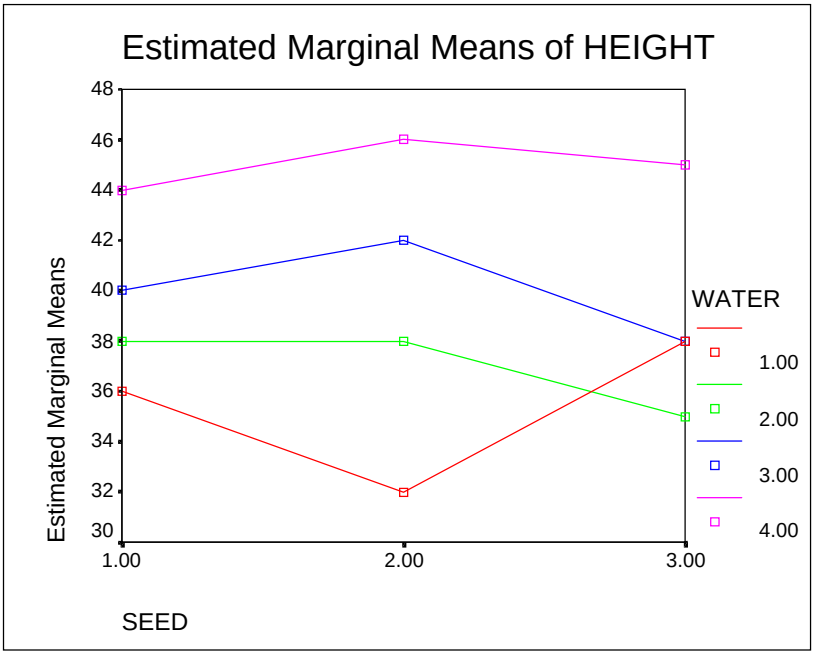
Now we examine the interaction effects with a profile plot. In general, profile plots (interaction plots) are useful for comparing marginal means in your model. A profile plot is a line plot in which each point indicates the estimated marginal mean of a dependent variable at one level of a factor. The plot for our data is displayed below.



The graph indicates a lack of additivity (interaction) between the means for the different seeds when taken across the water levels.

The strongest interaction effect is shown for the water level 1 with seeds 2 and 3. This corresponds to the point where the above graph displays the greatest degree of non-additivity.

The profile plot of marginal means versus seed type is shown below.



The three lines corresponding to the three water levels 1-3 show that the three seeds respond in similar manner across the three seed levels. These lines indicate that the means for the water level 4 are higher than the means for water level 3, and those are higher than the means for water level 2. The combination of seed 2 and water level 4 produces the highest plants. The combination of seed 2 and water level 1 produces the lowest plants.

8.4 Multiple Comparisons

Since the hypothesis for main effects for water was strongly rejected, multiple comparisons might be considered for the means of the levels of the factor. The following SPSS output obtained with the GLM model shows the results of the multiple comparisons with the Tukey HSD procedure.

Multiple Comparisons

Dependent Variable: HEIGHT

Tukey HSD

(I) WATER	(J) WATER	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.6667	.850	.255	-4.1898	.8564
	3.00	-4.6667*	.850	.001	-7.1898	-2.1436
	4.00	-9.6667*	.850	.000	-12.1898	-7.1436
2.00	1.00	1.6667	.850	.255	-.8564	4.1898
	3.00	-3.0000*	.850	.019	-5.5231	-.4769
	4.00	-8.0000*	.850	.000	-10.5231	-5.4769
3.00	1.00	4.6667*	.850	.001	2.1436	7.1898
	2.00	3.0000*	.850	.019	.4769	5.5231
	4.00	-5.0000*	.850	.000	-7.5231	-2.4769
4.00	1.00	9.6667*	.850	.000	7.1436	12.1898
	2.00	8.0000*	.850	.000	5.4769	10.5231
	3.00	5.0000*	.850	.000	2.4769	7.5231

Based on observed means.

*. The mean difference is significant at the .05 level.

The above table shows significant differences in water main effects on the plant height across the water levels. As you can see, there are significant differences in the water main effects for any pair of water levels except for the levels 1 and 2.