PLANT-GROWTH EXPERIMENT

5. Working with Your Data

This section contains detailed instructions you should follow to carry out the statistical analysis of your data. The instructions are provided in the form of questions for you to answer while analyzing the data. The answers to all the questions for our data with the corresponding computer instructions are included in the sections that follow.

5.1 Experiment Design

Describe the plant-growth experiment by answering the following questions:

(a) What is the main goal of the experiment?
(b) What were the two seed varieties for your experiment?
(c) What are the levels of the water factor? Briefly describe the thought process that went into selecting these levels.
(d) What is the response variable? How will it be measured and in what units?
(e) How was randomization used in your experiment?
(f) List a factor other than seed variety and water that could be used in the plant-growth experiment. Suggest two levels for this factor.
(g) List another response variable that could have been used.
(h) Give an example of an uncontrollable factor in the current setup. Do you anticipate that this factor will significantly affect your findings?

5.2 Displaying the data

Graphical displays of the data can be very helpful for understanding the information contained in the data.

(a) Use the *Line Chart* feature in the *Graphs* menu to obtain the plot of mean heights versus seed type by watering plan and the plot of mean heights versus watering plan by seed type. What do you conclude about the effects of seed type and watering plan on the plant growth? Can you conclude that one of the two factors affects plant growth more than the other? What combinations of the factor-levels produce the highest and the smallest plants?

(b) Use the Scatterplot feature in the Graphs menu to obtain a scatterplot of plant height versus water with seed as the grouping variable. Is the information contained in this plot consistent with the conclusions reached in (a)?

5.3 Describing the Data

In this part you will provide some numerical summaries of your data. In particular, you will obtain the table of factor-level means.

(a) Use the *Means* procedure in the *Compare Means* submenu in the *Statistics* menu to obtain the table of the marginal means of plant height for all combinations of water and seed. Water and seed should be entered as two layers in the procedure. Estimate the main effects of the two factors by comparing the average responses for each row and column with the overall average response. What do you conclude about the effects of the two factors on plant growth?

(b) Estimate the interaction effects based on the results obtained in part (a). State briefly your conclusions about the strength of the interaction effect between the two factors.

5.4 Two-Way Analysis of Variance

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 procedure can be used for any 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).

In this part you will use the GLM General Factorial Procedure in SPSS to carry out the statistical analysis of the effects of water and seed on the plant height.

(a) Use the *General Factorial Procedure* in the *General Linear Model* menu to examine the main and interaction effects in the model. Report the observed significance levels of the appropriate F-tests and state the conclusions about the significance of the effects.

(b) If the interaction effect in part (a) is found to be significant, find out which combinations of levels of the factors are different.

(c) Obtain the profile (interaction) plots of mean height versus the two factors. Are the estimated marginal means increasing or decreasing across levels? Are the plots consistent with the results obtained in part (a)?

(d) If the hypothesis for main effects for at least one of the factors is rejected, multiple comparisons might be considered for the means of the levels of the factor. Use the Tukey HSD procedure in SPSS to determine which levels of the factor differ. State your conclusions briefly.

5.5 Checking the Assumptions

The results of the above analysis are valid only if the appropriate assumptions are satisfied. The assumptions are that the data come from a normal population and in the population, all cell variances are the same. Analysis of variance is robust to departures from normality, although the data should be symmetric

In order to check the assumptions in SPSS, you can use homogeneity of variance tests and spread-versus-level plots. You can also examine residuals and residual plots. All the features can be accessed from the *GLM General Factorial* dialog box by clicking on *Options*. Then answer the following questions:

(a) Is there any evidence that the assumption of equal error variance might be violated?

(b) Is the normality assumption reasonable?

Answer the above questions by referring to the plots produced by *GLM General Factorial* procedure.

5.6 Final Conclusions

Summarize the model discussed above by answering the following questions:

(a) Write a brief report summarizing your findings. In particular, assess the significance of main and interaction effects based on the graphical displays and inferential methods. Which factor-level combination of seed and water seems to produce the highest plants?

(b) What might be gained by considering a model with more factors?