CAKE-BAKING EXPERIMENT

8. Two-Way Analysis of Variance

In this section two-way analysis of variance will be used to investigate the effects of all possible combinations of baking time and temperature on the taste of the cakes. In particular, the main effects for both time and temperature are shown to be very small. However, the interaction between the two factors is highly significant. After the interaction is found to be significant, custom contrast coefficients can be discussed to find out which combinations of levels of the factors are different.

8.1 Cake-Baking Experiment as a Factorial Experiment
8.2 ANOVA Output for the Cake-Baking Experiment
8.3 Exploring the Interaction Effects
8.4 Contrasts

8.1 Cake-Baking Experiment as a Factorial Experiment

The cake-baking experiment is an example of a factorial experiment. A factorial experiment consists of several factors (baking time, baking temperature) which are set at different levels, and a response variable (taste score). The purpose of the experiment is to assess the impact of different combinations of the levels of baking time and temperature on the taste of the cakes. Analysis of variance allows us to test the null hypothesis that baking time and temperature have no impact on the taste score. As the experiment involved replications, so that responses are available from more than one subject at each combination of levels of time and temperature, the presence of interaction can be assessed.

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 *cake1.sav* located on the FTP server in the Stat337 directory. In the data file, variables include time, temperature and taste score. The two-predictor variables in this study, time level and temperature level, 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 baking time and baking temperature have no impact on taste score. There are four sources of variation in the experiment: the main effects of *Time* and *Temperature*, 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 *Time*
2. H_0: No main effect of *Temperature*
3. H_0: No interaction effect between *Time* and *Temperature*

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

Tests of Between-Subjects Effects

Dependent Variable: RATING

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	66.000[a]	8	8.250	3.908	.008
Intercept	243.000	1	243.000	115.105	.000
TIME	2.000	2	1.000	.474	.630
TEMP	2.000	2	1.000	.474	.630
TIME * TEMP	62.000	4	15.500	7.342	.001
Error	38.000	18	2.111		
Total	347.000	27			
Corrected Total	104.000	26			

a. R Squared = .635 (Adjusted R Squared = .472)

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 *Time*, *Temp*, *Time*Temp* (interaction), 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 26 =27-1, which is one less than the number of cakes tested. Two degrees of freedom are associated with time, which is one less than the number of levels of this factor, and similarly 2 degrees of freedom for temperature. The interaction factor Seed*Water degrees of freedom equals 4 = (3-1)(3-1), the product of the degrees of freedom associated with the two factors. Interaction can be thought of as the joint effect of the two factors.

According to the output the model sum of squares is 66.000 and the error sum of squares is 38.000. The total sum of squares (corrected total) is 104.000. Notice a very small contribution of error in the total sum of squares. The p-value of the F-test for the model is reported as 0.008 indicating a sufficient evidence of an effect of at least one of the factors on the taste score.

The sum of squares for the time factor is estimated to be only 2.000, an extremely small value compared to the total sum of squares. The p-value of the F-test is reported as 0.630, indicating no evidence of effect of time on the taste score. Indeed, in all graphical displays and numerical summaries we found no evidence of time effect on the taste score.

The sum of squares due to temperature is also 2.000, a very small contribution in the total sum of squares of 104.000. The value of the F-statistic is 0.474 with the corresponding reported p-value of 0.630. Temperature main effects are not statistically significant.

The p-value of the interaction term *Time*Temperature* is equal to 0.001, indicating a strong evidence of an interaction between the two factors. Thus, in further analysis, the baking time effect should be compared at each level of baking temperature.

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.

Report

RATING

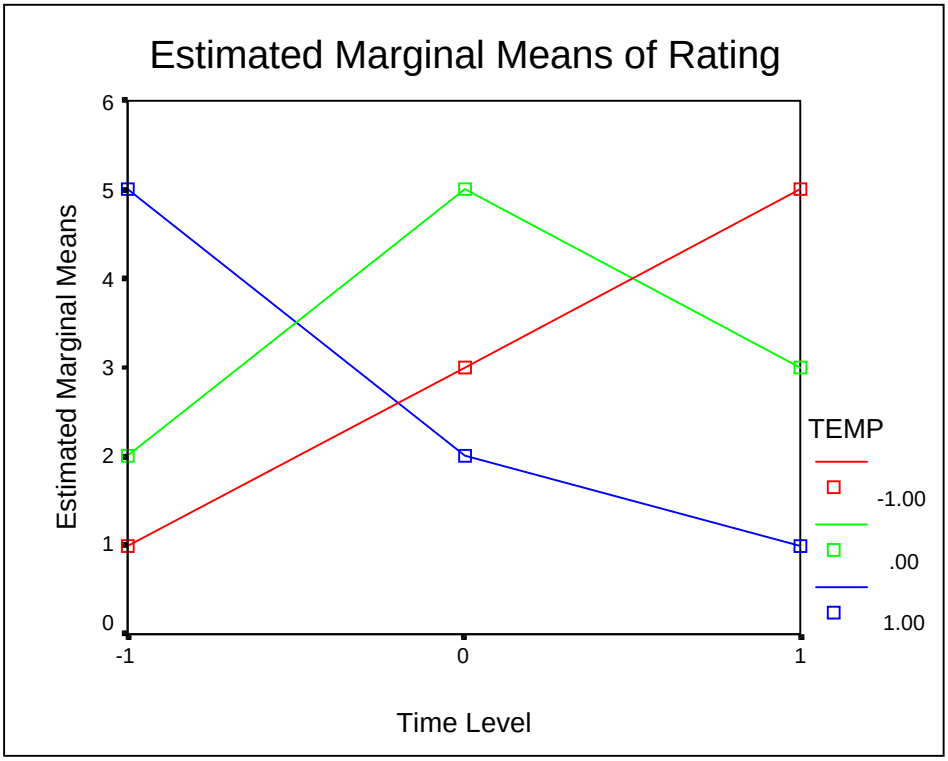
TIME	TEMP	Mean	N	Std. Deviation
-1	-1.00	1.0000	3	1.7321
	.00	2.0000	3	2.0000
	1.00	5.0000	3	1.0000
	Total	2.6667	9	2.2913
0	-1.00	3.0000	3	1.0000
	.00	5.0000	3	1.7321
	1.00	2.0000	3	1.0000
	Total	3.3333	9	1.7321
1	-1.00	5.0000	3	1.0000
	.00	3.0000	3	2.0000
	1.00	1.0000	3	1.0000
	Total	3.0000	9	2.1213
Total	-1.00	3.0000	9	2.0616
	.00	3.3333	9	2.1213
	1.00	2.6667	9	2.0000
	Total	3.0000	27	2.0000

The table shows the means of the three ratings for each combination of levels of the factors. The overall mean is 3, the average of all 27 ratings. The marginal means differ very little from the overall mean of 3, which is why the Time and Temperature main effects are not statistically significant. On the other hand, the means vary from a low of 1 to a high of 5.

The table shows that the highest mean rating is achieved with the following combinations of time and temperature: (+1, -1), (0, 0), and (-1, +1). Thus it is possible to compensate

for a lower time by using a higher temperature and vice versa. The lowest average rating is achieved with the following combinations of time and temperature: (-1, -1) and (1, +1).

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 lines in the above graph cross each other, there is strong interaction between time and temperature. The strongest interaction effects are shown for the time level -1 (low) and +1 (high) with temperature levels at -1 (low) and +1 (high). This corresponds to the point where the above graph displays the greatest degree of non-additivity.

The plot shows that the highest mean rating is achieved with the following combinations of time and temperature: (+1, -1), (0, 0), and (-1, +1). The lowest average rating is achieved with the following combinations of time and temperature: (-1, -1) and (1, +1).

8.4 Contrasts

After the interaction between baking time and temperature is found to be significant, it is possible to discuss now custom contrasts to find out which combinations of levels of the factors are different. In order to find out how to customize your analysis by using syntax commands (LMMATRIX) that are not available in the dialog boxes, use help in SPSS.