Homework Assignment for Lesson 10

Please submit one document file in the ANGEL Dropbox for Lesson10.

Problem: In a study of seed yield in plants, three fertilizers were compared (C, S, and F) with 10 replications of each treatment level. The experiment was set up as a completely randomized design, that is, by assigning fertilizer levels to plants (in containers) using a simple random process. However, they noticed that the plants differed in heights at the start of the fertilizer treatments, and this reflected the vigor of the plants. They were concerned that this heterogeneity in experimental units would present a problem, perhaps masking treatment effects, so they measured the height of each plant at the start of the experiment.

The data is in the Excel file ‘Plantdata’.

a) Plot the data. This can be done any way you choose (Excel scatter plot, Minitab, or SAS) to show the response (seed\_yield vs. height). It is important here to label the points on the scatter plot to indicate the treatment level they belong to.



b) Fit a simple linear regression for each fertilizer treatment separately and determine if any of the regressions are significant. Include the output for the regression tests of significance.

|  |
| --- |
| **Fert = c** |

**The REG Procedure**

**Model: MODEL1**

**Dependent Variable: yield**

|  |  |
| --- | --- |
| **Number of Observations Read** | 10 |
| **Number of Observations Used** | 10 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 1 | 2.61672 | 2.61672 | 113.60 | <.0001 |
| **Error** | 8 | 0.18428 | 0.02304 |  |  |
| **Corrected Total** | 9 | 2.80100 |  |  |  |

|  |
| --- |
| **Fert = s** |

**The REG Procedure**

**Model: MODEL1**

**Dependent Variable: yield**

|  |  |
| --- | --- |
| **Number of Observations Read** | 10 |
| **Number of Observations Used** | 10 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 1 | 1.95081 | 1.95081 | 310.94 | <.0001 |
| **Error** | 8 | 0.05019 | 0.00627 |  |  |
| **Corrected Total** | 9 | 2.00100 |  |  |  |

|  |
| --- |
| **Fert = f** |

**The REG Procedure**

**Model: MODEL1**

**Dependent Variable: yield**

|  |  |
| --- | --- |
| **Number of Observations Read** | 10 |
| **Number of Observations Used** | 10 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 1 | 2.18703 | 2.18703 | 143.45 | <.0001 |
| **Error** | 8 | 0.12197 | 0.01525 |  |  |
| **Corrected Total** | 9 | 2.30900 |  |  |  |

The highlighted values indicate that all the simple linear regressions are significant, so the slopes are not = 0.

c) Following the Steps of ANCOVA, run the appropriate models to come up with a final model. Include a description of the process (steps) you are using, and include output associated with the steps you follow. Identify your final model and draw a conclusion about the Null Hypothesis of equality of the means for the fertilizer levels.

Step 1 - Are all regression slopes = 0.

|  |
| --- |
| Fert = c |

| Analysis of Variance | | | | | |
| --- | --- | --- | --- | --- | --- |
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 2.61672 | 2.61672 | 113.60 | <.0001 |
| Error | 8 | 0.18428 | 0.02304 |  |  |
| Corrected Total | 9 | 2.80100 |  |  |  |

|  |
| --- |
| Fert = s |

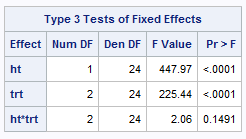
| Analysis of Variance | | | | | |
| --- | --- | --- | --- | --- | --- |
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 1.95081 | 1.95081 | 310.94 | <.0001 |
| Error | 8 | 0.05019 | 0.00627 |  |  |
| Corrected Total | 9 | 2.00100 |  |  |  |

|  |
| --- |
| Fert = f |

| Analysis of Variance | | | | | |
| --- | --- | --- | --- | --- | --- |
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
| Model | 1 | 2.18703 | 2.18703 | 143.45 | <.0001 |
| Error | 8 | 0.12197 | 0.01525 |  |  |
| Corrected Total | 9 | 2.30900 |  |  |  |

The highlighted values indicate that all the simple linear regressions are significant, so the slopes are not = 0.

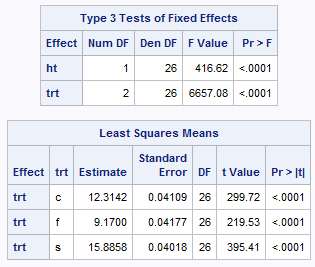
Step 2: Are the slopes equal?



We know that if the slopes differ significantly among treatment levels, the interaction p-value will be < 0.05. However since that is not the case here, we conclude that the slopes are equal and in a plot of the regressions we see that the lines are almost parallel:



Step 3: Fit an Equal Slopes Model



Since in SAS the model statement automatically creates an intercept, and so the ANCOVA model is technically over-parameterized. To get the slopes and intercepts for the covariate directly, we have to re-parameterize the model.

| **Solution for Fixed Effects** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **trt** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **ht** |  | 0.05581 | 0.002734 | 26 | 20.41 | <.0001 |
| **trt** | c | 9.5293 | 0.1336 | 26 | 71.34 | <.0001 |
| **trt** | f | 6.3851 | 0.1535 | 26 | 41.59 | <.0001 |
| **trt** | s | 13.1009 | 0.1396 | 26 | 93.86 | <.0001 |

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **ht** | 1 | 26 | 416.62 | <.0001 |
| **trt** | 3 | 26 | 7859.21 | <.0001 |

| **Least Squares Means** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **trt** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **trt** | c | 12.3142 | 0.04109 | 26 | 299.72 | <.0001 |
| **trt** | f | 9.1700 | 0.04177 | 26 | 219.53 | <.0001 |
| **trt** | s | 15.8858 | 0.04018 | 26 | 395.41 | <.0001 |

| **Differences of Least Squares Means** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **trt** | **\_trt** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** | **Adjustment** | **Adj P** |
| **trt** | c | f | 3.1442 | 0.06037 | 26 | 52.08 | <.0001 | Tukey-Kramer | <.0001 |
| **trt** | c | s | -3.5716 | 0.05703 | 26 | -62.62 | <.0001 | Tukey-Kramer | <.0001 |
| **trt** | f | s | -6.7158 | 0.05851 | 26 | -114.78 | <.0001 | Tukey-Kramer | <.0001 |

Identify your final model and draw a conclusion about the Null Hypothesis of equality of the means for the fertilizer levels.

Thus, the estimated regression equation

Fert c: y-hat = 9.5293 + 0.05581 (ht)

Fert f: y-hat = 6.3851 + 0.05581 (ht)

Fert s: y-hat = 13.1009 + 0.05581 (ht)

From the analysis we can see that fert is significant:

