Homework for Lesson 11

Submit a Word document or pdf file to the ANGEL dropbox for this assignment.

A study is conducted to compare two teaching methods ( A and B) and five training times (1, 3, 5, 7, or 9 hours) in a crossed design. High school students are assigned completely at random to receive the one of the method x training treatment combinations. There were 3 students randomly assigned to each method x training combination, and at the end of the training session mastery scores were recorded.

The data are in the Week 11 Lessons folder as HW11\_Schools.xlsx.

1. Plot the data to show the response variable vs. hours training for each method (on one graph).



1. Run an ANOVA to compare the two methods at each level of training. Just show the output ANOVA table and appropriate mean comparisons.

**General Linear Model: mastery versus Method, hours**

Method

Factor coding (-1, 0, +1)

Factor Information

Factor Type Levels Values

Method Fixed 2 A, B

hours Fixed 5 1, 3, 5, 7, 9

Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Method 1 32764 32763.5 32119.63 0.000

hours 4 173523 43380.8 42528.25 0.000

Method\*hours 4 25111 6277.7 6154.32 0.000

Error 20 20 1.0

Total 29 231418

The interaction term is significant and the comparison is:

**Comparisons for mastery**

**Tukey Pairwise Comparisons: Response = mastery, Term = Method\*hours**

Grouping Information Using the Tukey Method and 95% Confidence

Method\*hours N Mean Grouping

B 9 3 299.958 A

B 7 3 184.268 B

A 9 3 138.231 C

B 5 3 99.120 D

A 7 3 86.537 E

A 5 3 49.367 F

B 3 3 42.117 G

A 3 3 22.145 H

B 1 3 8.972 I

A 1 3 7.683 I

Means that do not share a letter are significantly different.

1. Use ANCOVA to characterize the effect of increasing number of training hours on mastery scores for the two teaching methods. Fit the response as a polynomial (order 2) function of training hours. Show the output and indicate what factors are significant.

|  |  |
| --- | --- |
| **Minitab** | **SAS** |
| **General Linear Model: mastery versus x, x2, Method**  Method  Factor coding (-1, 0, +1)  Factor Information  Factor Type Levels Values  Method Fixed 2 A, B  Analysis of Variance  Source DF Adj SS Adj MS F-Value P-Value  Method 1 7891 7891 5086.49 0.000  x 1 165253 165253 106522.24 0.000  x2 1 8259 8259 5323.99 0.000  x\*Method 1 23836 23836 15365.07 0.000  x2\*Method 1 1269 1269 817.87 0.000  Error 24 37 2  Lack-of-Fit 4 17 4 4.13 0.013  Pure Error 20 20 1  Total 29 231418  Model Summary  S R-sq R-sq(adj) R-sq(pred)  1.24553 99.98% 99.98% 99.98%  Coefficients  Term Coef SE Coef T-Value P-Value VIF  Constant 74.008 0.354 208.84 0.000  Method  A -25.274 0.354 -71.32 0.000 2.43  x 26.2403 0.0804 326.38 0.000 1.00  x2 2.4790 0.0340 72.97 0.000 1.00  x\*Method  A -9.9659 0.0804 -123.96 0.000 1.00  x2\*Method  A -0.9716 0.0340 -28.60 0.000 2.43  Regression Equation  Method  A mastery = 48.734 + 16.274 x + 1.5074 x2  B mastery = 99.282 + 36.206 x + 3.4506 x2 |  |

All the factors (including the interactions) above are significant.

1. Write a 250 word summary of your results.

Designed experiments often contain treatment levels that have increasing numerical values. Here for example we have two factors: the method ( A and B) and five training times (1, 3, 5, 7, or 9 hours). This is a factorial treatment design, and we also have 3 replications of each method × hours combination administered in a completely randomized design. Since the treatment levels are quantitative and there are at least 3 levels of measurement, we can use ANCOVA to investigate the quantitative factor level with regression.

We can proceed as usual with a 2 × 5 factorial ANOVA to evaluate the Null Hypotheses

H0:μA=μB

H0:μ1=μ3=μ5=μ7=μ9

and H0:no interaction

We do that in the part 2 of the homework and get this output:

Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Method 1 32764 32763.5 32119.63 0.000

hours 4 173523 43380.8 42528.25 0.000

Method\*hours 4 25111 6277.7 6154.32 0.000

Error 20 20 1.0

Total 29 231418

The interaction term is significant

Further the design matrix enables us to add new columns for fitting a quadratic polynomial function to model the effect of the hours. We will want to add into the design matrix hours and hours2, to allow us to look at linear and quadratic trends, respectively. In addition, we want to test to see if any of the quantitative factor trend terms interact with the method type. To do this, we have to center the covariate by subtracting the mean of the covariate (5) from each hour level. We get the output:

Analysis of Variance

Source DF Adj SS Adj MS F-Value P-Value

Method 1 7891 7891 5086.49 0.000

x 1 165253 165253 106522.24 0.000

x2 1 8259 8259 5323.99 0.000

x\*Method 1 23836 23836 15365.07 0.000

x2\*Method 1 1269 1269 817.87 0.000

Error 24 37 2

Lack-of-Fit 4 17 4 4.13 0.013

Pure Error 20 20 1

Total 29 231418

We see that 1) the method effect is significant (p < 0.001), the linear and quadratic terms are significant in describing the trend in the response, and linear and quadratic effects are not the same for each of the methods (the interaction terms are significant).