Lesson 5 (Review) Solutions

**1)** A study was conducted to evaluate the effect of three different drugs on blood cholesterol concentration (mg/100 ml plasma) in women. Each drug could be obtained from two different sources and these were also of specific interest to the researchers.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DRUG 1** | | **DRU** | **G 2** | **DRUG 3** | |
| **Source A Source Q** | | **Source D Source B** | | **Source L Source S** | |
| 102 | 103 | 108 | 109 | 104 | 105 |
| 104 | 104 | 110 | 108 | 106 | 107 |

a) (5pts) State the Null and Alternative hypotheses for this experiment.

b) Source and df columns shown in the output below

Analysis of Variance for cholesterol, using Adjusted SS for Tests

**Source DF**  Seq SS Adj SS Adj MS F P

**Drug 2**  61.167 61.167 30.583 20.39 0.002

**Source(Drug) 3**  1.500 1.500 0.500 0.33 0.802

**Error 6**  9.000 9.000 1.500

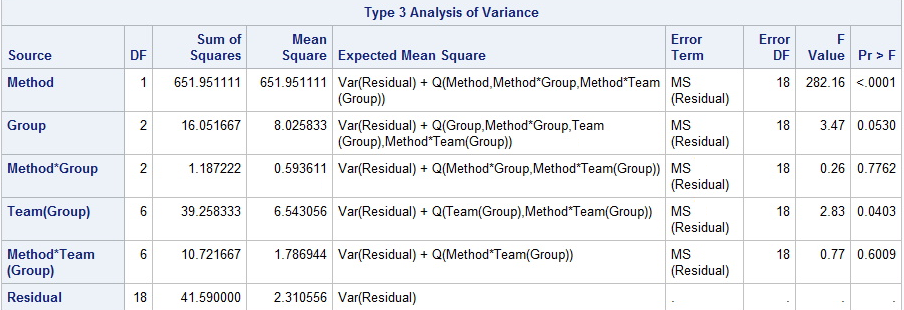
**Total 11**  71.667

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

**2)** A military installation is interested in evaluating the speed of reloading a large gun. Two methods of reloading are considered, and 3 groups of cadets were evaluated (slight, average, and heavy individuals). Three teams were set up within each group and they wanted to identify the fastest team within each group to go on to a demonstration for the military officials. The data is in the Excel File ‘navy\_gun\_data.xlsx’.

a) (5pts) State the Null and Alternative hypotheses for this experiment

b) Source and df shown in the output below;



**3)** An experiment is conducted to evaluate the thrust force encountered when drilling in a machine shop. The factors of interest are Material (two levels), the feed rate (3 levels) and the drill speed (5 levels). Two replications of each of the 2 x 3 x 5 combinations of treatments were used in a completely randomized design (i.e., the order in which the treatment combinations were set up and tested was randomized). The researchers are specifically interested in the significance (or lack of significance) of any interaction terms, so do not ‘reduce’ the model. The data appears in the Excel file ‘Drill\_data.xlsx’.

a) (5pts) State the Null and Alternative hypotheses for this experiment.

b) Source and df columns shown in the output below:

proc mixed data=drill method=type3 plot=all;

class material feed\_rate speed;

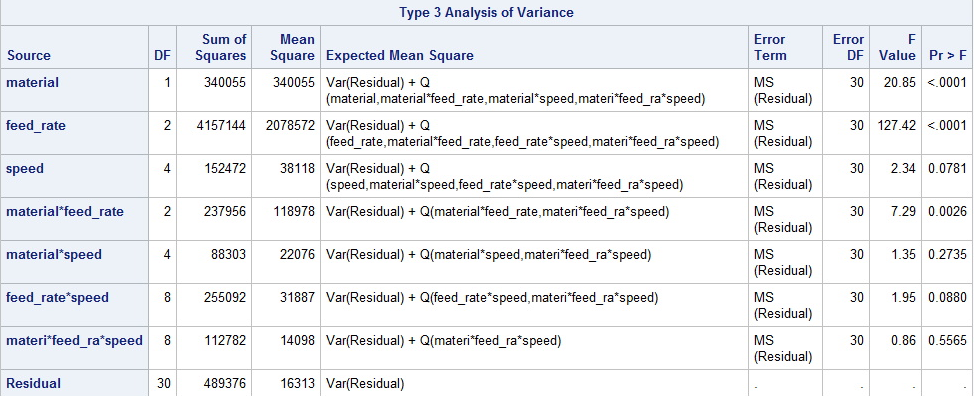
model force = material feed\_rate speed

material\*feed\_rate material\*speed feed\_rate\*speed

material\*feed\_rate\*speed;

store out\_full;

run;



**4)** Consider the following 2-factor factorial, with factor A having 4 levels and factor B having 3 levels, and n=3 observations per treatment combination. Given that the SSA + SSB + SSAxB = 170,

**a)** (5 pts) Complete the ANOVA table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | d.f. | SS | MS | F |
| FactorA | *3* | *120* | 40 | *16* |
| FactorB | *2* | 20 | *10* | *4* |
| AxB | *6* | *30* | *5* | *2* |
| Error | *24* | *60* | *2.5* |  |
| Total | *35* | 230 |  |  |

**b)** (5pts) What is the FCritical (or Fα) for the test of the AxB interaction? (*Appendix B, Table B4)* ***2.51***

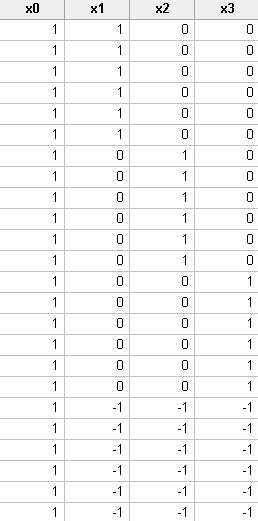
**c)** (5pts) What decision would you make for H0: there is no AxB interaction? *Do not Reject H0.*

**5)** An experiment was conducted to compare volume delivery (vol) by 4 machines. For each machine, 6 trials (replications) were recorded and we want to fit the ANOVA model

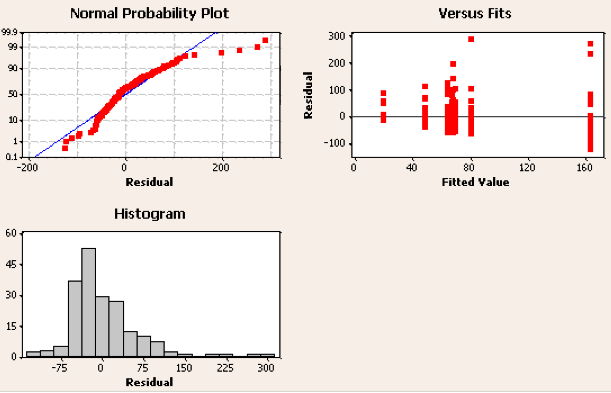
The data, arranged in unstacked format (*i columns* by *j rows)*, was:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| trial | machine1 | machine2 | machine3 | machine4 |
| 1 | 18 | 18 | 15 | 19 |
| 2 | 18 | 19 | 14 | 20 |
| 3 | 19 | 18 | 15 | 18 |
| 4 | 17 | 18 | 16 | 19 |
| 5 | 19 | 19 | 14 | 18 |
| 6 | 18 | 19 | 16 | 19 |
|  |  |  |  |  |

a) Construct the design matrix **X,** the matrix containing the coded variables as columns, to generate the estimates of for this model using regression.



**6)** (5pts) The following graphs resulted from a one-way ANOVA. From the graphs, what can you conclude about the model?



*There are a few concerns here. Most important is non-constancy of variance and suggests a log transformation (residuals are ‘fanning out’ in upper right panel). In addition, the residual distribution is not centered well on 0, and is right skewed.*

7) ANS: If there is no variability among treatment means, they are all equal to each other and equal to the grand mean. The SS for the treatment would be 0. As treatment means start to differ, the SS due to the treatments increases. If the MS computed from the SStrt is large enough (compared to residual variance) then we declare significance.

**8)** The F-test is

a) a non-parametric method

**b) a variance ratio test**

c) a waste of time

d) a and c

e) a and b

f) none of the above

**9)** In the context of a statistical hypothesis test, answer the following with regard to the Null Hypothesis. Your answer must be in words, NOT using symbols or the terms “type I” or “type II” error.

**a)** What is α? Probability of rejecting H0 when H0 is true.

**b)** What is 1-α? (Confidence) Probability of not rejecting H0 when H0 is true. (not getting false positive results).

**c)** What is β? Probability of accepting H0 when it is false.

**d)** What is 1-β? (Power) Probability of not accepting Ho when false: not missing significant effects.

(Other answers…

**a)** What is α?

This is the test's probability of incorrectly rejecting the null hypothesis when the null hypothesis is true, also known as the false positive rate. A test's significance level, a probability threshold below which the null hypothesis will be rejected is denoted by α.

**b)** What is 1-α?

Also known as the test specificity, 1-α denotes the probability of correctly accepting the null hypothesis when the null hypothesis is true.

**c)** What is β?

A test's probability of accepting the null hypothesis when the null hypothesis is false

**d)** What is 1-β?

The power of a statistical test, the probability of correctly rejecting the null hypothesis when the null hypothesis is false.

**10) 10)** In the textbook (Section 26.3) the authors state that, in a balanced nested ANOVA, where Factor B is nested within the levels of Factor A:

SSB(A) = SSB + SSAB.

*Briefly,* how would you interpret the relationship SSB(A) = SSB + SSAB?

*By definition of interaction, the response to levels of B differ within levels of A. So effect of B when nested can be thought of as the main effect of B and in addition, the ‘unique’ aspect of levels of B being different in each A, the interaction.*

**11)** Factors affecting fabric durability were evaluated in a 3-factor factorial experiment. There were two surface preparations (S\_trt 1 and S\_trt 2), two Filler types (F\_trt 1 and F\_trt 2) and three filler proportions (25%, 50% and 75%). Note that as a factorial all of the 2 x 2 x 3 combinations of treatments appear in the experimental design. Two replications of the 12 treatment combinations were prepared for testing. The data is provided in Excel as “Fabric\_data” in the Lesson 5 (Review) folder.

a) State the Null and Alternative hypotheses for this experiment.

b) Write out the first two columns of the ANOVA table (i.e. Source and df columns).

Source DF

S\_trt 1

F\_trt 1

Filler 2

S\_trt\*F\_trt 1

S\_trt\*Filler 2

F\_trt\*Filler 2

S\_trt\*F\_trt\*Filler 2

Residual 12

Total 23