SIE 330R Homework, Spring 2023

**Agustin Espinoza**

**HW 6 (Chapter 4)**

**3/14/2023**

Homework must be readable! Do not just send in numbers or charts. You must explain the homework answers Preferred to receive homework in Word doc format with any excel or Minitab results pasted into word document. You may choose to use pdf which is also OK.

# Put answers to all questions in one document NOT in separate documents

* 4.3. Blocking is a technique that can be used to control the variability transmitted by uncontrolled

nuisance factors in an experiment. True False

When the source of nuisance variability is known and uncontrollable, the blocking technique can be implemented to reduce the variability of the dependent variable results and lead to a more accurate model.

**4.14.** An article in *Communications of the ACM* (Vol. 30, No. 5, 1987) studied different algorithms for estimating software development costs. Six algorithms were applied to several different software development projects and the percent error in estimating the development cost was observed. Some of the data from this experiment is show in the table below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Project | | | | | | |
| Algorithm | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 (SLIM) | 1244 | 21 | 82 | 2221 | 905 | 839 |
| 2 (COCOMO-A) | 281 | 129 | 396 | 1306 | 336 | 910 |
| 3 (COCOMO-R) | 220 | 84 | 458 | 543 | 300 | 794 |
| 4 (COCOMO-C) | 225 | 83 | 425 | 552 | 291 | 826 |
| 5 (FUNCTION POINTS) | 19 | 11 | -34 | 121 | 15 | 103 |
| 6 (ESTIMALS) | -20 | 35 | -53 | 170 | 104 | 199 |

1. Do the algorithms differ in their mean cost estimation accuracy?

The ANOVA below provides the F-value, 5.377, which is greater than the F critical value. Therefore, the mean cost estimation accuracies differ significantly.



1. Analyze the residuals from this experiment.

The residual analysis provided below is a result of the Blocking ANOVA technique.



1. Which algorithm would you recommend for use in practice?

Based on the analysis provided above, I would recommend the FUNCTIONAL POINTS algorithm for practical use as it has the lowest cost estimation error.

**4.54.** Physics graduate student Laura Van Ertia has conducted a complete randomized design with a single factor, hoping to solve the mystery of the unified theory and complete her dissertation. The results of this experiment are summarized in the following ANOVA display:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | *DF* | *SS* | *MS* | *F* |
| Factor | ? | ? | 14.18 | ? |
| Error | ? | 37.75 | ? |  |
| Total | 23 | 108.63 |  |  |

Answer the following questions about this experiment.

1. The sum of squares for the factor is?

SSTotal = SSFactor + SSError

108.63 = SSFactor + 37.75

SSFactor­­­ = 108.63 – 37.75

SSFactor­­­ = 70.88

1. The number of degrees of freedom for the single factor in the experiment is?

dfFactor = (sum of squares for the factor) / (MS for the factor)

dfFactor = 70.88 / 14.18

dfFactor = 5

1. The number of degrees of freedom for the error is?

dfTotal = dfFactor + dfError

23 = 5 + dfError

dfError = 18

1. The mean square for error is?

dfError = (sum of squares for the error) / (MS for the error)

MS for the error = (sum of squares for the error) / (dfError)

MS for the error = 37.75 / 18

MS for the error = 2.10

1. The value of the test statistic is?

F = MSFactor / MS­Error

F = 14.18 / 2.10

F = 6.75

1. If the significance level is 0.05, your conclusions are not to reject the null hypothesis.

Yes or No

The statement above is False. If the significance level is 0.05, and your statistical test produces a test statistic value greater than 0.05, then your conclusions would be to reject the null hypothesis.

1. The P-value for the test statistic is?

P-value = 0.001

* 1. Laura used how many levels of the factor in this experiment?

6 levels. This problem is solved by adding 1 to the factor DF value. In this case, the factor DF is 5. Therefore, there are 6 levels of the factor.

* 1. Laura replicated this experiment how many times?

Laura replicated this experiment 4 times.

* 1. Suppose that Laura had actually conducted this experiment as a random complete block design and the sum of squares for the blocks was 12. Reconstruct the ANOVA display above to reflect this new situation. How much has the blocking reduced the estimate of the experimental error?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | *DF* | *SS* | *MS* | *F* |
| Factor | ? | ? | 14.18 | ? |
| Error | ? | 37.75 | ? |  |
| Total | 23 | 108.63 |  |  |

The blocking reduced the *SSError*by 12 and the *MSError by* 0.67.

# Categorical Question 1:

A one sample test of proportions can be evaluated with a Z score True False

A one-sample test of proportions can be evaluated with a Z score if the sample size is sufficiently large and if the assumptions for using a normal approximation are met.

# Categorical Question 2:

A Chi-Square contingency table has 4 rows and 6 columns. How many Degrees of Freedom will the test of independence have?

* + 1. 24
    2. 23
    3. 15
    4. 9

The formula for calculating Degrees of Freedom (DF) for a Chi-Square test of independence is:

DF = (# of rows - 1)(# of cols - 1)

In this case, we have 4 rows and 6 columns:

DF = (4 - 1)(6 - 1)

DF = 3(5)

DF = 15

Therefore, a Chi-Square contingency table with 4 rows and 6 columns will have 15 DF.