Tutorial Letter 014/0/2023

Applied Statistics III

STA3701

Year module

Department of Statistics

ASSIGNMENT 4 QUESTIONS

ASSIGNMENT 04

Unique Nr.: 859766

Due date: 18 September 2023

Instructions

1. **Do not PLAGIARISE. Students suspected of plagiarism will be subjected to disciplinary processes.**
2. **Use R to answer all the questions. Present or attach R outputs. Label all the figures and tables.**

Question 1 **[10]**

1.1 Define analysis of variance (ANOVA). (1)

1.2 How is ANOVA different from regression analysis? (2)

1.3 Explain the difference between one-factor (one-way) ANOVA and two-factor (two-way) ANOVA. (2)

1.4 The objective of a study was to determine whether there are a different responses to different calcium channel blockers[1]. Two hundred and fifty patients with mild-to-moderate hypertension were randomly assigned to 4 weeks of treatment with once-daily doses of: (i) lercanidipine; (ii) felodipine; or (iii) nifedipine. Prior to treatment and at the end of 4 weeks, each of the subjects had his or her systolic blood pressure measured. Researchers then calculated the change in systolic blood pressure.

1.4.1 What is the treatment variable in this study? (1)

1.4.2 What is the response variable in this study? (1)

1.4.3 Mention or state one extraneous variable that you think its effects would be included in the error term. (1)

1.4.4 What are the degrees of freedom for the F critical value in this study? (2)

[1]Daniel, W. W., & Cross, C. L. (2013). Biostatistics: a foundation for analysis in the health sciences. Tenth edition. Hoboken, NJ, Wiley.

Question 2 **[40]**

Fifty-five pigs were randomly assigned to five different treatments and the weight gain was recorded[1]. The weight gains of the pigs are given in the table below.

Treatment 1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
165	168	164	185	201
156	180	156	195	189
159	180	156	195	189
159	180	189	184	173
167	166	138	201	193
170	170	153	165	164
146	161	190	175	160
130	171	160	187	200
151	169	172	177	142
164	179	142	166	184
158	191	155	165	149

Table 1:Table 1: Weight gain in pigs

2.1 Create a boxplot of the data. (4)

2.2 Does the boxplot suggest any differences among the means? Comment on the homogeneity of variances. (2)

2.3 The model for this data is $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$. Define all the terms/notations in the model. (2)

2.4 What are the assumptions underlying the model in part 2.3? (2)

2.5 Use the Anova to test the hypothesis of equal treatment means

($H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$) at five percent level of significance. (5)

2.6 Use Tukey's *honestly significant differences* (HSD) test to test all possible pairwise mean differences at 5% level of significance. (15)

2.7 Perform residual analyses to investigate whether any model assumptions were violated. (10)

[1]Daniel, W. W., & Cross, C. L. (2013). Biostatistics: a foundation for analysis in the health sciences. Tenth edition. Hoboken, NJ, Wiley.

Question 3 **[50]**

Researchers at a trauma center wished to develop a program to help brain-damaged trauma victims regain an acceptable level of independence[1]. An experiment involving 72 subjects with the same degree of brain damage was conducted. The objective was to compare different combinations of psychiatric treatment and physical therapy. Each subject was randomly assigned to one of 24 different combinations of four types of psychiatric treatment and six physical therapy programs. There were three subjects in each combination. The response variable is the number of months elapsing between initiation of therapy and time at which the patient was able to function independently. Use in R to analyze the data below and answer the questions that follow.

Physical Therapy Program	**Psychiatric treatment**			
	A	**B**	**C**	**D**
1	11.0 9.6 10.8	9.4 9.6 9.6	12.5 11.5 10.5	13.2 13.2 13.5
2	10.5 11.5 12.0	10.8 10.5 10.5	10.5 11.8 11.5	15.0 14.6 14.0
3	12.0 11.5 11.8	11.5 11.5 12.3	11.8 11.8 12.3	12.8 13.7 13.1
4	11.5 11.8 10.5	9.4 9.1 10.8	13.7 13.5 12.5	14.0 15.0 14.0
5	11.0 11.2 10.0	11.2 11.8 10.2	14.4 14.2 13.5	13.0 14.2 13.7
6	11.2 10.8 11.8	10.8 11.5 10.2	11.5 10.2 11.5	11.8 12.8 12.0

Table 2:the number of months elapsing between initiation of therapy and time at which the patient was able to function independently.

3.1 Use interaction plots to investigate whether there are significant psychiatric treatment type and physical therapy program. (6)

[1]Daniel, W. W., & Cross, C. L. (2013). Biostatistics: a foundation for analysis in the health sciences. Tenth edition. Hoboken, NJ, Wiley.

3.2 Give the appropriate model and interaction effects assumptions. (Define the terms/notations used in the model using the given information). (8)

3.3 Conduct a two-way analysis of variance (ANOVA) and present the results in an ANOVA table. (8)

3.4 Using a 5% level of significance, test the hypotheses, which follow.

3.4.1 Can you conclude based on these data that the different psychiatric treatment programs have different effects? In other words, does the mean number of months elapsing between initiation of therapy and time at which the patient was able to function independently differ among different types of psychiatric treatments? (H_0: $\alpha_A = \alpha_B = \alpha_C = \alpha_D$)? (6)

3.4.2 Can you conclude that the physical therapy programs differ in effectiveness? In other words, does the type of physical therapy program affect the mean number of months elapsing between initiation of therapy and time at which the patient was able to function independently? (H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6$)? (6)

3.4.3 Can you conclude that there are significant interaction effects between psychiatric treatment type and physical therapy program (H_0: $(\alpha\beta)_{ij} = 0$)? Does your conclusion corroborate your finding in part 3.1? (6)

NB: Provide the null and alternative hypotheses, critical regions (or rejection regions), test statistics and conclusions.

3.5 Perform a residual analysis of your model and report your observations. (10)

Grand total = [100]

[1]Daniel, W. W., & Cross, C. L. (2013). Biostatistics: a foundation for analysis in the health sciences. Tenth edition. Hoboken, NJ, Wiley.