Uva Wellassa University of Sri Lanka
Faculty of Science and Technology
Department of Computer Science and Technology
300 Level 2nd Semester Examination–Dec. 2018/ Jan. 2019
CST 314-2/ IIT 313-2 Statistical Methods II

Uva Wellassa University

UVA WELLASSA UNIVERSITY OF SRI LANKA LIBRARY

Instructions to candidates.

Duration: Two (02) hours and fifteen minutes (15)

Number of questions: Four (04) esssay questions

Mark allocation: 100

Answer All Questions

1.

a. Explain the meaning of the following terms used in hypothesis tests.

I. Type I error. (2 mark)

II. Level of significance. (2 mark)

III. Power of the test. (2 mark)

b. *"Experiments are conducted to obtain correct information about real world phenomena"*

I. Briefly discuss the importance of Experimental Designs in research with an examples. (4 mark)

II. What are the main principles of Experimental Designs? (4 mark)

c.

I. Explain why Linear Regression cann't be used for Time Series data? (3 mark)

II. Describe the main component of Time Series analysis. (3 mark)

2.

a. A taste-testing experiment has been designed. Four brands of coffee are to be rated by nine experts. To avoid any carryover effects, the tasting sequence of the four brews is randomly determined for each of the nine expert tasters until a rating on a 7-point scale (1 = extremely unpleasing, 7 = extremely pleasing) is given for each of four characteristics: taste, aroma, richness, and acidity. The following table displays the ratings accumulated over all four characteristics.

Blocks of Experts	Coffee Brand			
	A	B	C	D
E.B.	24	26	25	22
N.B.	27	27	26	24
M.D.	19	22	20	16
M.H.	24	27	25	23
B.J.	22	25	22	21
R.J.	26	27	24	24
B.K.	27	26	22	23
B.M.	25	27	24	21
J.S.	22	23	20	19

75

"At a level of significance of 0.01 is the sufficient evidence to indicate that there is a difference in the summated ratings of the four brands of coffee? " Write down a full report to achieve the above target. (Your report shouldinclude: appropriate statistical test, Suitable hypothesis, mean comparisons and other necessary analysis with the appropriate conclusions.) (10 mark)

b. In a study of super market checkout equipment it is found that, although checkout prices are often correct, customers can sometimes be charged more or less than the prices posted on the shelves. It is suggested that discrepancies in prices may be associated with whether or not items are on special offer compared with their normal prices. A random sample of 819 items is investigated and for each item it is noted whether the checkout equipment is undercharging, overcharging or charging the correct price. The results are shown in the contingency table below

	Normal-Priced items	Special offer items	Total
Undercharged	20	7	27
Overcharged	15	29	44
Correct price	384	364	748
Total	419	400	819

I. Test whether or not price accuracy is associated with items being on special offer or not. State your null and alternative hypotheses and report your conclusions clearly. (6 mark)

II. Adapt the table to classify the normal-priced and special offer items according to whether or not they are charged correctly. Calculate a 95% confidence interval for the difference in the proportions of items which are charged incorrectly in the normal priced and special offer categories. (5 mark)

III. Comment briefly on your answers to parts (a) and (b). (4 mark)

3. A bank has recorded the number y of businesses that it has driven to bankruptcy by the recalling of loans during a single financial year, classified into five geographical regions (1–5) and four types of business (A–D).

	Region				
Type	**1**	**2**	**3**	**4**	**5**
A	58	56	51	55	54
B	45	39	53	54	50
C	45	35	49	37	38
D	48	36	41	51	50

a. Write down an appropriate linear model for these data, relating y to region and type of business, stating your assumptions about the term that represents random variation. Comment on how appropriate your assumptions are in this case. (5 mark)

b. Construct the suitable analysis of variance table for these data under your model. State appropriate hypotheses and hence use the analysis of variance table to test at the 5% level of significance whether there is any evidence of a difference in mean number of bankruptcies between the types of business and whether there is any evidence of a difference in mean number of bankruptcies between the regions. Carefully state your conclusions. (7 mark)

c. Managers are most interested in seeing how easy it is to force different types of business in different regions into bankruptcy. Comment on what this data set and this analysis show, and do not show, with respect to this question. What other information would be useful? (7 mark)

d. A manager comments that businesses of the same type often experience very different business conditions in the different regions Explain the concept of interaction and hence explain if and how it could be modelled in this case, through a different model and/or more data. (6 mark)

4. The data below (Table 03) consist of 16 indexed observations on rubber consumptions.
(The data are stored in a Minitab work sheet named "Rubber.mtw" given on your computer.)

Observation No	Total Rubber Consumption Y_1	Tire Rubber Consumption Y_2	Car Production X_1	Gross National Product X_2	Disposable Personal Income X_3	Motor Fuel Consumption X_4
1	0.909	0.871	1.287	0.984	0.987	1.046
2	1.252	1.220	1.281	1.078	1.064	1.081
3	0.947	0.975	1.061	1.061	1.007	1.051
4	1.022	1.021	0.787	1.013	1.012	1.046
5	1.044	1.002	0.796	1.028	1.029	1.036
6	0.905	0.890	1.392	0.969	0.993	1.020
7	1.219	1.213	0.892	1.057	1.047	1.057
8	0.923	0.918	1.400	1.001	1.024	1.034
9	1.001	1.014	0.721	0.996	1.003	1.014
10	0.916	0.914	1.032	0.976	0.993	1.013
11	1.173	1.170	0.685	1.046	1.027	1.037
12	0.936	0.952	1.291	1.004	1.001	1.007
13	0.965	0.946	1.170	1.002	1.014	1.008
14	1.106	1.096	0.817	1.049	1.032	1.024
15	1.011	0.999	1.231	1.023	1.020	1.030
16	1.080	1.093	1.086	1.035	1.053	1.029

a. Examine data carefully and state preliminary conclusions. (7 mark)

b. Use these data to develop suitable fitted equations for predicting Y_1 and Y_2 separately in terms of the predictor variables X1, X_2, X_3 and X_4. (8 mark)

c. Write a brief report of your findings (it should include all relevant hypothesis with ANOVA, model checking with appropriate test and validation and necessary and relevant analysis). (15 mark)

UVA WELLASSA UNIVERSITY OF SRI LANKA
LIBRARY