

EASTERN UNIVERSITY, SRI LANKA,
THIRD EXAMINATION IN SCIENCE – 2004/2005
(Oct./Nov.' 2006)
SECOND SEMESTER
ST 304 – DATA ANALYSIS

Answer all questions
 Time: Two hours

- Q1. The data below gives 15 measurements of two variables X and Y
- | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| X: | 23 | 24 | 26 | 25 | 30 | 24 | 23 | 22 | 29 |
| Y: | 7.6 | 7.7 | 5.4 | 5.9 | 5.0 | 6.5 | 8.3 | 8.2 | 5.2 |
| | | | | | | | | | |
| X: | 24 | 25 | 28 | 22 | 22 | 24 | | | |
| Y: | 8.2 | 6.0 | 4.9 | 8.7 | 8.1 | 6.0 | | | |

Using MINITAB,

- (i) Draw a scatter diagram
- (ii) Find the linear regression of Y on X
- (iii) Draw the fitted line on the graph in (i)
- (iv) Test the hypothesis that the slope of the regression line is zero
- (v) Form the columns of fitted values and residuals
- (vi) Test whether the model is simple linear regression or not
- (vii) Draw the confidence band and fitted line in same graph without scatter diagram
- (viii) Check the assumption of normality.

Q2. Using MINITAB,

- (a) i) Enter the four columns of data in the format given below

Name	Height	Weight	Colour
Dennis	165	67	Blue
Stuart	172	68	Red
Richard	166	69	Green
Michael	164	68	Black
William	173	72	Green
Gareth	162	69	Red
Graham	177	64	Black
Leslie	162	67	Red
Tom	170	69	Red
John	169	68	Blue

- ii) Compute the proportion of individuals having weight higher than 67.
 - iii) Obtain the names and corresponding height and weight values for the individuals whose weight is higher than 68.
 - iv) Rank the selected students according to weight values.
 - v) Arrange the selected data according to the order of ranks.
- (b) Let the scores assign to colours are as follows; Red 4, Blue 5, Green 9 and Black 11.
- i) Compute the total for each colour.
 - ii) Sort the data (Names and colours only) according to the score of colour.

- Q3. Tennis balls are tested in a machine to see how many bounces they can withstand before they fail to bounce 30% of their dropping height. Two brands of balls (W and P) are compared. In addition, the effect of shelf life on these brands is tested. Half of the balls of each brand are 6 months old, the other half, fresh. The data are shown below.

		BRAND	
		W	P
AGE	NEW	67 72 74 82 81	75 76 80 72 73
	OLD	46 44 45 51 43	63 62 66 62 60

Write a SAS program to solve this problem using a two way analysis of variance. What conclusion can you reach?

- Q4. Using SAS, Number treated (y) out of sample size of n in 5 districts are given below.

District	y	n
1	22	27
2	41	64
3	12	34
4	11	43
5	24	28

Enter the data as given above. Write a programme to obtain a cross tabulation, District \times Status as below

District	Status		Total
	treat	untreated	
1	22	5	27
2	41	23	64
3	12	22	34
4	11	32	43
5	24	4	28
Total	110	86	196