



WIT

**BACHELOR OF SCIENCE IN
INFORMATION TECHNOLOGY**

EXAMINATION:

**STATISTICAL ANALYSIS
SEMESTER TWO YEAR 1**

MAY 2008

DURATION: 2 HOURS

**INTERNAL EXAMINERS: MS A VEREKER (FULL TIME)
MR JONATHAN BRAZIL (ACCS)**

EXTERNAL EXAMINER: PROF EAMON MURPHY

INSTRUCTIONS TO CANDIDATES

- 1. ATTEMPT ALL QUESTIONS**
- 2. TOTAL MARK = 100**
- 3. LINEAR REGRESSION FORMULAE AND STATISTICS FORMULAE
ARE APPENDED FOR USE WITH QUESTIONS AS APPROPRIATE**

EQUIPMENT/MATERIALS SUPPLIED

- GRAPH PAPER**

WATERFORD INSTITUTE OF TECHNOLOGY

Question 1

- (a) Explain the term Ordinal data and provide an example of this data type. **(6 marks)**
- (b) Briefly describe how the stratified sampling method is used to generate a sample from a given population and provide an example of the method being used. **(6 marks)**
- (c) The following table shows the amounts of money donated to a certain charity in a given year.

Amount of Donation (€)	No of Donors
0 - 20	8
20 - 40	30
40 - 60	55
60 - 80	40
80 - 100	7

Note: The interval (0 - 20) ranges from €0.00 to €19.99.

- (i) Present the data shown above in a cumulative frequency distribution table. **(2 marks)**
- (ii) Draw an ogive of the data. **(4 marks)**
- (iii) Draw a histogram of the data set. **(4 marks)**
- (iv) Calculate the mean of the data set. **(4 marks)**
- (v) Calculate the mode of the data set. **(4 marks)**
- (vi) Calculate the inter-quartile range of the data set. **(10 marks)**

(Total 40 marks)

Question 2

The following table shows the mortgage interest rate charged by a bank along with the number of mortgages set up with the bank over an eight month period.

Mortgage Interest Rate (%)	4.1	4.2	3.9	4.0	4.2	4.3	4.4	4.5
No of Mortgages Set Up per Month	18	17	25	23	19	16	15	14

- (i) Calculate the regression line equation for the above data. **(10 marks)**
- (ii) Give a brief description of the value of **b** in the regression line equation calculated in (i). **(4 marks)**
- (iii) Plot a scatter diagram of the above data. **(2 marks)**
- (iv) Draw the regression line on the diagram created for (iii). **(4 marks)**
- (v) Calculate the coefficient of determination (r^2) and comment on its value. **(6 marks)**
- (vi) How many mortgages would you expect the bank to set up during a month if the mortgage interest rate charged was 3.8%. Explain the reliability of your answer. **(4 marks)**

(Total 30 marks)

Question 3

- (a) The following are the sales figures for a company over a twelve-month period in 2006.

JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
85	81	76	68	57	42	35	31	49	58	71	79

Calculate the moving averages forecasts with $n = 4$ from May 2007 to January 2008.

(10 marks)

- (b) The table below represents the sales of a particular item over a 3-year period.

Year	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
2005	460	300	360	440
2006	490	325	390	450
2007	470	350	410	480

- (i) Given that the regression line, plotting sales versus time, is given by $Y = 375 + 5.5x$, determine the seasonally adjusted forecasts for each quarter in the above table.

(18 marks)

- (ii) Determine the seasonally adjusted forecasts for each quarter in 2008.

(2 marks)

(Total 30 marks)

Statistical Formulae

1. Quartiles for grouped data

$$Q = L + \frac{r}{f}(U - L)$$

2. Mode for grouped data

$$Mode = L + \frac{d_1}{d_1 + d_2}(U - L)$$

3. Arithmetic mean for grouped data

$$\bar{x} = \frac{\sum fx}{\sum f}$$

Linear Regression

1. Linear parameters a and b

$$S_{xy} = \frac{\sum xy}{n} - (\bar{x})(\bar{y})$$

$$S_{xx} = \frac{\sum x^2}{n} - (\bar{x})^2$$

$$S_{yy} = \frac{\sum y^2}{n} - (\bar{y})^2$$

$$b = \frac{S_{xy}}{S_{xx}} \quad a = \bar{y} - b\bar{x}$$

2. Coefficient of determination

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}}$$