UNIVERSITY OF SWAZILAND

SUPPLEMENTARY EXAMINATION PAPER 2015

TITLE OF PAPER	:	DESCRIPTIVE STATISTICS
COURSE CODE	:	ST 132
TIME ALLOWED	:	TWO (2) HOURS
REQUIREMENTS	:	CALCULATOR
INSTRUCTIONS	:	THIS PAPER HAS FIVE (5) QUESTIONS. ANSWER ANY FOUR (4) QUESTIONS.

Question 1 [25 marks, 14+6+5]

(a) The table below shows the numbers of units sold of a company's products quarter by quarter over a three-and-a-half years.

	Sales in SZL'000			
Year	Q1	Q2	Q3	Q4
1987	100	125	127	102
1988	104	128	130	107
1989	110	131	133	107
1990	109	132		

Use the method of moving averages to construct a deseasonalised series.

(b) The following are the average weekly wages of part-time legal office employees in a large city for the years 1985 through 1990: 187.55, 196.92, 203.82, 217.88, 239.67, and 252.85 emalangeni.

(i) Construct an index showing the changes in these wages from the base year 1987.

(ii) If a consumer price index for this city showed an increase of 10 percent from 1988 to 1990 how much did these employees earn in 1990 in real wages (constant 1988 emalangeni).

Question 2 [25 marks, 6+6+8+5]

(a) A company wishes to measure the change in its performance using an index calculated from the data given below on numbers of times sold and their prices in 1990 and 1991.

	1990		1991	
Item	Price	Number	Price	Number
A	2.50	90	2.70	200
B	3.80	150	4.00	160
C	4.10	180	4.50	120

Use 1990 as base and calculate for 1991:

(i) the Laspeyres quantity index;

(ii) the simple aggregate quantity index.

(b) The number of senior civil servants (a random sample) who joined work before 8:45 am, almost every day, was recorded as follows:

17	17	18	18	18	19	20	21	22	24
24	25	25	26	26	27	27	27	27	28

(i) Calculate the coefficient of skewness.

(ii) Estimate the interquartile range.

Question 3 [25 marks, 8+6+6+5]

(a) It is believed that the price of a house in a certain city may be related to its distance from the centre of the city. These distances (in kilometres) can easily be obtained from a map and are given below for the 12 houses in the sample.

House	A	B	C	D	E	F	G	H	I	J	K	L
Price (SZL 000)	63	75	59	75	100	108	100	90	70	96	84	100
Distance	5.5	5.7	5.2	4.9	3.3	2.1	2.2	3.1	4.2	3.1	3.5	2.8

(i) Using least squares, find the regression coefficients of house price on distance from the city centre. Explain to a manager, with no statistical knowledge the meaning of the terms: slope, intercept and coefficient of determination.

(ii) The overall average distance from the city centre is 4.5 kilometres. Use this information to estimate the population mean house price..

(iii) Comment on the advantages of using linear regression for forecasting and the limitations of the technique.

(b) The summary statistics for two data sets are as follows:

	Sample size	**Sample mean**
X data	19	7.0
Y data	25	5.1

Compute the mean of the combined data sets.

Question 4 [25 marks, 8+6+5+6]

(a) *The Wall Street Journal* Stock Market Data Bank reports the numbers of shares traded on the New York Stock Exchange in half-hourly intervals. Following are the combined numbers of shares traded (in millions of shares) at half-hourly intervals for three recent days.

Shares traded (in millions)	**Number of half-hourly periods**
1	1
2-4	3
10-14	17
15-19	8
20-24	3
25-29	2
30-34	1
5-9	8

(i) Calculate the coefficient of skewness.

(ii) Estimate the interquartile range.

(b) In the UK Index of Retail Prices for December 1986 (January 1974=100) the approximate index for beer was around 500 and that for cheese was 400. Consider the following statements about December 1986:

(i) The price of beer was lower than the price of cheese.

(ii) The price of beer was higher than the price of cheese.

(iii) The change in the price of beer was 20 percent greater than the change in the price of cheese since January 1974.

Which of the statement(s) is/are true?

(c) For a certain product, data is available on last quarter's sales, by value, and on current quarter's prices and sales volume. Which of the following index number types can be calculated, using the last quarter as base?

(i) Laspeyres price index;

(ii) Laspeyres quantity (volume) index;

(iii) Paasche price index;

(iv) Paasche quantity (volume) index; and

(v) Sales value index.

Question 5 [25 marks, 5+5+5+5+5]

A police officer classifies a total of 150 reported crimes in 2009 by age (in years) of the criminal and whether the crime is violent or non-violent.

	Age (in years)		
Type of crime	**Under 20**	**20 to 40**	**Over 40**
Violent	27	41	14
Non-violent	12	34	22

You must define the respective event(s) in each case and must use one of the probability rules to compute the following probabilities:

(a) What is the probability of selecting a case to analyse and finding it involved a violent crime?

(b) What is the probability of selecting a case to analyse and finding the crime was committed by some one 40 or less than 40 years old?

(c) What is the probability of selecting a case that involved a violent crime or an offender less than 20 years old?

(d) Given that a violent crime is selected for analysis, what is the probability the crime was committed by a person under 20 years old?

(e) Two crimes are selected for review by a Judge. What is the probability that both are violent crime?

APPENDIX 2: LIST OF KEY FORMULAE

MEASURES OF CENTRAL LOCATION

Arithmetic mean

Ungrouped data

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3.1}$$

Grouped data

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{n}$$

Mode

Grouped data

$$M_o = O_{mo} + \frac{c(f_m - f_{m-1})}{2f_m - f_{m-1} - f_{m+1}} \tag{3.3}$$

Median

Grouped data

$$M_e = O_{me} + \frac{c\left[\frac{n}{2} - f(<)\right]}{f_{me}} \tag{3.2}$$

Lower quartile

Grouped data

$$Q_1 = O_{q1} + \frac{c\left(\frac{n}{4} - f(<)\right)}{f_{q1}} \tag{3.7}$$

Upper quartile

Grouped data

$$Q_3 = O_{q3} + \frac{c\left(\frac{3n}{4} - f(<)\right)}{f_{q3}} \tag{3.8}$$

Geometric mean

Ungrouped data

$$GM = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n} \tag{3.4}$$

Weighted arithmetic mean

Grouped data

$$\text{weighted } \bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} \tag{3.5}$$

MEASURES OF DISPERSION AND SKEWNESS

Range

$$\begin{aligned} \text{Range} &= \text{Maximum value} - \text{Minimum value} + 1 \\ &= x_{max} - x_{min} + 1 \end{aligned} \tag{3.9}$$

Variance

Mathematical – ungrouped data

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{(n-1)} \tag{3.10}$$

Computational – ungrouped data

$$s^2 = \frac{\Sigma x_i^2 - n\bar{x}^2}{(n-1)} \tag{3.11}$$

Standard deviation

$$s = \sqrt{s^2} \tag{3.12}$$

Coefficient of variation

$$CV = \frac{s}{\bar{x}} \times 100\% \tag{3.13}$$

Pearson's coefficient of skewness

$$sk_p = \frac{n\Sigma(x_i - \bar{x})^3}{(n-1)(n-2)s^3} \tag{3.14}$$

$$sk_p = \frac{3\,(\text{Mean} - \text{Median})}{\text{Standard deviation}} \quad \text{(approximation)} \tag{3.15}$$

PROBABILITY CONCEPTS

Conditional probability

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \tag{4.2}$$

Addition rule

Non-mutually exclusive events

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{4.3}$$

Mutually exclusive events

$$P(A \cup B) = P(A) + P(B) \tag{4.4}$$

Multiplication rule	*Statistically dependent events* $P(A \cap B) = P(A/B) \times P(B)$	4.5
	Statistically independent events $P(A \cap B) = P(A) \times P(B)$	4.6
$n! = n$ factorial	$n \times (n-1) \times (n-2) \times (n-3) \times \ldots \times 3 \times 2 \times 1$	4.8
Permutations	${}_nP_r = \frac{n!}{(n-r)!}$	4.10
Combinations	${}_nC_r = \frac{n!}{r!\,(n-r)!}$	4.11

PROBABILITY DISTRIBUTIONS

Binomial distribution	$P(x) = {}_nC_x\, p^x(1-p)^{(n-x)}$ for $x = 0, 1, 2, 3, \ldots, n$	5.1
	$P(x \text{ successes}) = \frac{n!}{x!\,(n-x)!} p^x(1-p)^{(n-x)}$ for $x = 0, 1, 2, 3, \ldots, n$	
Binomial descriptive measures	Mean $\mu = np$ Standard deviation $\sigma = \sqrt{np(1-p)}$	5.2
Poisson distribution	$P(x) = \frac{e^{-a} a^x}{x!}$ for $x = 0, 1, 2, 3 \ldots$	5.3
Poisson descriptive measures	Mean $\mu = a$ Standard deviation $\sigma = \sqrt{a}$	5.4
Standard normal probability	$z = \frac{x-\mu}{\sigma}$	5.6

6

INDEX NUMBERS

Price relative	Price relative $= \frac{p_1}{p_0} \times 100\%$	13.2
Laspeyres price index	*Weighted aggregates method* Laspeyres price index $= \frac{\Sigma(p_1 \times q_0)}{\Sigma(p_0 \times q_0)} \times 100\%$	13.5
Laspeyres price index	*Weighted average of relatives method* Laspeyres price index $= \frac{\Sigma\left[\left(\frac{p_1}{p_0}\right) \times 100 \times (p_0 \times q_0)\right]}{\Sigma(p_0 \times q_0)}$	13.9
Paasche price index	*Weighted aggregates method* $= \frac{\Sigma(p_1 \times q_1)}{\Sigma(p_0 \times q_1)} \times 100\%$	13.8
Paasche price index	*Weighted average of relatives method* $= \frac{\Sigma\left[\left(\frac{p_1}{p_0}\right) \times 100 \times (p_0 \times q_1)\right]}{\Sigma(p_0 \times q_1)}$	13.10
Quantity relative	Quantity relative $= \frac{q_1}{q_0} \times 100\%$	13.11
Laspeyres quantity index	*Weighted aggregates method* Laspeyres quantity index $= \frac{\Sigma(p_0 \times q_1)}{\Sigma(p_0 \times q_0)} \times 100\%$	13.12
Laspeyres quantity index	*Weighted average of relatives method* Laspeyres quantity index $= \frac{\Sigma\left[\left(\frac{q_1}{q_0}\right) \times 100 \times (p_0 \times q_0)\right]}{\Sigma(p_0 \times q_0)}$	13.14

Paasche quantity index *Weighted aggregates method*

$$= \frac{\Sigma(p_1 \times q_1)}{\Sigma(p_1 \times q_0)} \times 100\% \qquad 13.13$$

Paasche quantity index *Weighted average of relatives method*

$$= \frac{\sum\left[\left(\frac{q_1}{q_0}\right) \times 100 \times (p_1 \times q_0)\right]}{\Sigma(p_1 \times q_0)} \qquad 13.15$$

Link relatives *Price*

$$= \frac{p_t}{p_{t-1}} \times 100\% \qquad 13.17$$

Quantity

$$= \frac{q_t}{q_{t-1}} \times 100\% \qquad 13.18$$

Composite

$$= \frac{\text{Basket value}_t}{\text{Basket value}_{t-1}} \times 100\% \qquad 13.19$$

or $$= \frac{\text{Composite index}_t}{\text{Composite index}_{t-1}} \times 100\%$$

REGRESSION AND CORRELATION

Formula $$\hat{y} = b_0 + b_1 x \qquad 12.1$$

Coefficients $$b_1 = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \qquad 12.2$$

$$b_0 = \frac{\Sigma y - b_1 \Sigma x}{n} \qquad 12.3$$

Pearson's correlation coefficient $$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2] \times [n\Sigma y^2 - (\Sigma y)^2]}} \qquad 12.4$$

t-stat $$= r\sqrt{\frac{(n-2)}{1-r^2}} \qquad 12.8$$

TIME SERIES ANALYSIS

Regression trend coefficients $$b_1 = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \qquad 12.2$$

$$b_0 = \frac{\Sigma y - b_1 \Sigma x}{n} \quad \text{where } x = 1, 2, 3, 4 \ldots n \qquad 12.3$$

De-seasonalised *y* $$= \frac{\text{Actual } y}{\text{Seasonal index}} \times 100 \qquad 14.5$$

FINANCIAL CALCULATIONS

Simple interest $$F_v = P_v(1 + in) \qquad 15.1$$

Compound interest $$F_v = P_v(1 + i)^n \qquad 15.2$$

$$F_v = P_v\left(1 + \frac{i}{k}\right)^{nk} \qquad 15.3$$

where k = number of compounding periods in a year.

Effective rate of interest $$r = \left(1 + \frac{i}{m}\right)^m - 1 \qquad 15.7$$

Ordinary annuity certain $$F_v = R\frac{(1+i)^n - 1}{i} \qquad 15.8$$

$$P_v = R\frac{1-(1+i)^{-n}}{i} \qquad 15.9$$

Ordinary annuity due $$F_v = R\frac{[(1+i)^n - 1](1+i)}{i} \qquad 15.10$$

$$P_v = R\frac{[1-(1+i)^{-n}](1+i)}{i} \qquad 15.11$$

Deferred annuity $$P_v = R\left(\frac{1-(1+i)^{-(m+n)}}{i} - \frac{1-(1+i)^{-m}}{i}\right) \qquad 15.12$$