MUNHUMUTAPA SCHOOL OF COMMERCE

DEPARTMENT OF ECONOMICS AND FINANCE

MAIN EXAMINATION

BACHELOR OF COMMERCE	PART 2 SEMESTER 2
COURSE	RESEARCH METHODS IN FINANCE AND INSURANCE
CODE	HBF 2211
DATE	2024
DURATION	3 HOURS

INSTRUCTIONS TO CANDIDATES

1. THE PAPER COMPRISES 5 QUESTIONS.
2. YOU ARE REQUIRED TO ANSWER ANY FOUR QUESTIONS.
3. BEGIN THE ANSWER TO EACH QUESTION ON A FRESH PAGE OF THE ANSWER BOOKLET.
4. NON-PROGRAMMABLE FINANCIAL OR SCIENTIFIC CALCULATORS ARE ALLOWED INTO THE EXAMINATION.
5. CANDIDATES WILL OBTAIN CREDIT FOR SHOWING ALL WORKINGS.

Question 1

Discuss the roles played by literature review, methodology and data presentation and analysis in documentation of a degree research project. [25 marks]

[Total 25 Marks]

Question 2

2.1 Distinguish between descriptive and inferential statistics. [2 marks]

2.2 The following data refer to the pension benefits to be realized by the contributors from Danda Limited on attainment of the 60 year retirement age:

Pension Benefit ($000)	600-900	900-1200	1200-1500	1500-1800	1800-2100	2100-2400	2400-2700	2700-3000
Number of Clients	18	20	24	36	50	40	30	32

2.2.1 Compute the mean, median and standard deviation of the data. [8 marks]

2.2.2 Draw a histogram and use it to estimate the mode of the data. [4 marks]

2.2.3 Evaluate the lower and upper quartiles of the data. Draw a Box and Whisker plot to determine the skewness of the data. [6 marks]

2.2.4 Construct a less than ogive of the data, and use it to read the 8^{th} decile benefit. [5 marks]

[Total 25 Marks]

Question 3

3.1 Explain the concepts of regression and correlation analyses. [2 marks]

3.2 The following data refer to the ages and prices of various car models on the market:

Age (In Years)	5	6	3	2	4	7	8	9
Car Price ($000)	16	14	22	25	18	12	10	8

3.2.1 Determine the independent and dependent variables of the data. [2 marks]

3.2.2 Show the data on a scatter plot and comment on the distribution. [3 marks]

3.2.3 Estimate the Ordinary Least Squares (OLS) model connecting the two variables. Interpret your estimated linear regression coefficients. [5 marks]

3.2.4 Estimate the pension benefit for a contributor who is 48 years today. [2 marks]

3.2.5 Calculate Pearson's product moment correlation coefficient and the coefficient of determination. Appraise your findings. [6 marks]

3.2.6 Test the claim that the age has no effect on a client's pension benefit at 5% level. [5 marks]

[Total 25 Marks]

Question 4

4.1 Discuss the conditions that suit the application of the Chi-Square and ANOVA approaches to hypothesis testing? [7 marks]

4.2 Three pills A, B and C are delivered to COVID patients by a doctor at Masvingo General Hospital in January 2024. The data below represent the number of days taken by five patients given each of the three drugs to recover:

Drug A: 27; 23; 24; 25; 26.
Drug B: 19; 21; 22; 18; 20.
Drug C: 21; 20; 22; 19: 23.

Test the claim that the three drugs have the same effectiveness levels at 5% level of significance. [18 marks]

[Total 25 Marks]

Question 5

5.1 Briefly explain any 4 statistical tests that can be used to determine the suitability of independent variables for inclusion in a research model. [6 marks]

5.2 The following data refer to the total quarterly sales generated by Duranta Limited for the period 2020 to 2022 ($000):

Period (In Years)	**Quarter 1**	**Quarter 2**	**Quarter 3**	**Quarter 4**
2020	360	480	440	460
2021	480	420	440	320
2022	380	340	420	380

5.2.1 Estimate the trend line connecting time and quarterly sales. Estimate the sales for quarter 3 of 2025. [5 marks]

5.2.2 Determine the 4 point centred moving averages of the data and use them to generate the seasonal indices of the data. [9 marks]

5.2.3 De-seasonalize the above data and comment on the effects of the use of moving averages in time series data analysis. [5 marks]

[Total 25 Marks]

END OF EXAMINATION

STATISTICAL FORMULAE FOR UNDER AND POST GRADUATE PROGRAMMES

A. SAMPLE AND POPULATION MEASURES

Sample Mean and Variance

1. Sample mean for a) Ungrouped data = $\frac{\sum x_i}{n}$ *and* b) Grouped data=$\frac{\sum f_i x_i}{n}$.

2. Sample variance for a) Ungrouped data = $s_x^2 = \frac{1}{n-1}[\sum x^2 - \frac{(\sum x_i)^2}{n}]$ and b) Groued data $= s_x^2 = \frac{1}{n-1}[\sum f_i x^2 - \frac{(\sum f_i x_i)^2}{n}]$

Measures of Position for Grouped Data

3. Estimated Mode for grouped data = Lm+$\frac{Cm(fm-fm-1)}{2fm-(fm+1+fm-1)}$

Position of median = 0.5(n+1).

Estimated Median for grouped data = Lm+$\frac{Cm(0.5n-Fm- \ )}{fm}$

4. Estimated Quartiles, Q1 and Q3= = Lq+$\frac{Cq(0.25or\ 0.75n-Fq-1)}{fq}$
5. Estimated Percentile, = Lp+$\frac{Cp(X\% \times n-Fp- \ )}{fp}$.

Population Mean and Variance

Population mean for a) Ungrouped data = $\frac{\sum X_i}{N}$ *and* b) Grouped data=$\frac{\sum F_i X_i}{N}$.

Sample variance for a) Ungrouped data = $\sigma_X^2 = \frac{1}{N}[\sum X^2 - \frac{(\sum X_i)^2}{N}]$ and b) Groued data $= \sigma_X^2 = \frac{1}{N}[\sum F_i X^2 - \frac{(\sum F_i X_i)^2}{N}]$

Measures of Position for Ungrouped Data

6. Estimated Mode for grouped data = Lm+$\frac{Cm(fm-fm-1)}{2fm-(fm+1+fm-1)}$
7. Position of median = 0.5(N+1).

Estimated Median for grouped data = Lm+$\frac{Cm(0.5n-Fm-1)}{fm}$

8. Estimated Quartiles, Q1 and Q3= = Lq+$\frac{Cq(0.25or\ 0.75n-F\quad)}{fq}$
9. Estimated Percentile, = Lp+$\frac{Cp(X\%\ \times n-Fp-\ )}{fp}$

B. THE THEORY OF PROBABILITY AND ITS DISTRIBUTIONS

1. Addition Law (Either----Or) = p(A or B) = p(A) + p(B).

2. Multiplication or Product Law (Both-----And) = p(A) $\times\ p(B)$.

Discrete Probability Distributions

The mean and variance are given by the formula E(X)=$\sum p_i \times X_i$ and σ_X^2 = E(X^2) -μ^2 respectively.

Binomial Probability Distributions

The mean or expected value of X and variance are given by E(X) =np and variance, σ_X^2= np(1-p).

The probability of X is given by the formula, P(X=x) =$nC_x p^x q^{n-x}$

Poisson Probability Distribution

The mean and variance are given by E(X) =m and σ_X^2=m respectively.

The probability of observing X successes, p(X) =$\frac{e^{-m} \times m^x}{x\ Factorial} = \frac{e^{-m} \times m^x}{x\,!}$.

Normal Probability Distributions

The Normal Distribution is characterised by two parameters which are the mean, μ and variance, σ_X^2, to constitute, ND (μ; σ^2)

The Standard Normal Distribution, SND (0;1) is used for calculation of probability and is given by Z = $\frac{X-\mu}{\sigma}$ where X =A raw score or mark recorded.

C. CONFIDENCE INTERVALS (CI) AND HYPOTHESIS TESTING

1. 95% CI for small samples = $\bar{x} \pm\ t\frac{\alpha}{2}$(t-1)$\frac{s}{\sqrt{n}}$.
2. 95% CI for large samples =$\bar{x} \pm\ Z\frac{\alpha}{2} \times \frac{s}{\sqrt{n}}$.

3. 95% CI for small samples with population variance known $=\bar{x} \pm Z\frac{\alpha}{2} \times \frac{\sigma}{\sqrt{n}}$..
4. 95% CI for small proportions$=\hat{p} \pm t\frac{\alpha}{2}(\mathrm{t}-1)\sqrt{\frac{\hat{p}\times\hat{q}}{n}}$.
5. 95%CI for large samples $=\bar{x} \pm Z\frac{\alpha}{2} \times \sqrt{\frac{\hat{p}\times\hat{q}}{n}}$.
6. The procedure for hypothesis testing follows 4 steps namely:

   .Formulation of hypothesis that is the null and alternate (or alternative) hypotheses (H_0) and H_1).
   .Reading the suitable statistic from tables (Critical value)
   Calculating the suitable test statistic using collected or given sample data (Calculated value)
   .By comparing the critical and calculated values we draw a conclusion.

.Illustrate the procedure for hypothesis testing in Statistics.

7. **Chi-Square, X^2 Test for Independence**

X^2 crit = X^2 ∝(r-1) (c-1) and Expected frequency $= \frac{Row\ Total \times Column\ Ttotal}{Ggrand\ Ttotal}$.

The formula for Chi-Square calculated, X^2 cal $=\frac{\sum(O-E)^2}{E}$.

8. **Chi-Square Goodness of Fit Test**

$X^2\ crit = X^2\ \alpha\ (k - r - 1)$ where r= the number of estimated parameters.

The formula for Chi-Square calculated, X^2 cal $=\frac{\sum(O-E)^2}{E}$.

D. REGRESSION, CORRELATION AND TIME SERIES ANALYSIS

9. The Simple Linear Regression Model

The general simple linear regression model is given by:

$Y_i = \beta_0 + \beta_1 X_i + e_i$ or $Y_i =$ a + $bX_i + e_i$ where

$Y_i =$ the response or dependent variable, β_0 and β_1 are regression coefficients, X_i = the predictor or independent variable and $+e_i$ = the random or error term.

Therefore b$=\frac{n\times\sum XY-\sum X\times\sum Y}{n\sum X^2-(\sum X)^2}$ and a $=\frac{\sum Y-b\sum X}{n}$.

.The correlation coefficient by Pearson, $r_{A,B} = \frac{n \times \sum XY - \sum X \times \sum Y}{\sqrt{(n\sum X^2 - (\sum X)^2)(n\sum Y^2 - (\sum Y)^2)}}$

The correlation coefficient by Spearman, $r_S = 1 - \frac{6\sum d^2}{n\,(n^2-1)}$.

10. The Multiple Time Series Model

It is given by $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + ------ + \beta_n X_n + e_i$.

11. The Simple Time Series Model

The model is given by:

$Y_t = \beta_0 + \beta_1 X_{t1} + e_i$ or $Y_t = a + bX_t + e_t$ where

Y_t= the response or dependent time variable, β_0 and β_1 are time series coefficients, X_t = the predictor or independent time variable and $+e_i$ = the random or error term.

12. The Multiple Time Series Model

It is given by $Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t3} + ------ + \beta_n X_{tn} + e_i$.

13. ANOVA Tests

ANOVA Under One Comparable Variable

Ho: $\mu_A = \mu_B = \mu_C$

H1:. $\mu_A \neq \mu_B \neq \mu_C$.

$Fcrit = F_{v2}^{v1}\alpha = F_{12}^{2}\,0.05 = 3.89$; where v1 and v2 are numerator and denominator degrees of freedom from error.

Total sum of squares, SST = $\sum n_i(x_i - x)^2$

Within groups sum of squares, SSW = $SS_A + SS_B + SS_C$ that is sum of squares for groups A, B and C.

ANOVA Under Regression and Correlation Analysis

Ho: b=0

H1: b $\neq$ 0.

$Fcrit = F_{v2}^{v1} \alpha;$

Total sum of squares, SST = $\sum Y^2 - n\bar{Y}^2$.

Regression sum of squares, SSR = b($\sum XY - nXY$) where, b is calculated as under the simple linear regression model above.

14. Construction of Indices

Unweighted Price Index =$\frac{\sum P_n}{\sum P_0}$ and weighted Price index =$\frac{\sum P_n \times Q_n}{\sum P_0 \times Q_0}$.

Laspeyres PI (LPI) =$\frac{\sum P_n \times Q_0}{\sum P_0 \times Q_0}$ and Paasche PI (PPI)=$\frac{\sum P_n \times Q_n}{\sum P_0 \times Q_n}$.

Fisher PI = $\sqrt{\frac{\sum P_n \times Q_0}{\sum P_0 \times Q_0} \times \frac{\sum P_n \times Q_n}{\sum P_0 \times Q_n}}$ and do the same for quantity indices.