"Investing in Africa's Future"

COLLEGE OF BUSINESS, PEACE, LEADERSHIP AND GOVERNANCE

NMEC209: COMPUTER AND STATISTICAL ANALYSIS

END OF SECOND SEMESTER EXAMINATION

LECTURER: DR. A MAUNE

MAY 2022

DURATION: 3 HOURS

INSTRUCTIONS TO CANDIDATES

1. Answer all questions in Section A and any **three** in Section B.
2. The marks allocated to **each** question are shown at the end of the question.
3. Candidates are allowed to use a non-programmable calculator.
4. Start **each** question on a new page in your answer booklet.

SECTION A: ANSWER ALL QUESTIONS

QUESTION ONE

Answer True/False (T/F) (Each question carries 2 marks)

1. The Chi-square test only tells us whether two variables are independent, it does not say anything about the magnitude of the dependency if one is found to exist.

2. As the number of degrees of freedom increases, the chi-square distribution becomes more nearly symmetric.

3. The shape of the chi-square distribution depends on the degrees of freedom, just like Student's t-distribution and as the number of degrees of freedom increases, the chi-square distribution becomes more nearly symmetric.

4. Dependent samples are often referred to as matched pairs. It is possible for an individual to be matched against him or herself.

5. A sampling method is independent when the individuals selected for one sample do not dictate which individuals are to be in a second sample.

6. If the p-value is smaller than 0.05 $p < 0.05$ then H0 can be rejected.

7. Given a survey of 1,000 investment companies, it was found that 700 invest in derivatives. Based on this sample, the point estimate is 700/1000.

8. The p-value is the probability, assuming that H0 is true, of observing a value for the test statistic that is as extreme as or more extreme than the value actually observed.

9. The sampled values are independent of each other for a hypothesis test for a population mean with sigma known.

10. The sample is obtained using simple random sampling or from a randomized experiment for a hypothesis test for a population mean with sigma known.

11. The correlation coefficient can take values between 0 and 1.

12. A regression line can be written in the following form: $y = a + b$.

13. Type II Error: Accepting Null Hypothesis when Null Hypothesis is true.

14. A sample is known as the entire group of individual to be studied.

15. Mode is the process when the smallest and largest values are dropped prior to finding the mean.

16. Given a survey of 2,500 traders, it was found that 1,500 of them expected an increase of a stock market index. Based on this sample, the point estimate is 1500.

17. Values of the chi-square statistic are always negative.

18. Chi-Square Distribution is symmetric.

19. The conventional dictum that correlation implies causation means that correlation can be validly used to infer a causal relationship between the variables.

20. If $p = 0.25$ and $n = 1000$, considering the binomial probability distribution, the standard deviation is 1000*0.25.

[40 marks]

SECTION B: ANSWER ANY THREE (3) QUESTIONS

QUESTION TWO

a) Explain the value of business statistics in management. (6 marks)

b) What is the difference between descriptive statistics and inferential statistics? (4 marks)

c) Explain three factors that influence data quality. (6 marks)

d) Describe two sources of data. (4 marks)

[20 marks]

QUESTION THREE

a) Explain the following terms:

    i) Data relevancy (2 marks)

    ii) Data cleaning (2 marks)

    iii) Data enrichment (2 marks)

    iv) Data quality (2 marks)

b) Discuss the main methods of data collection. (6 marks)

c) Explain the differences between a bar chart and a histogram. (6 marks)

[20 marks]

QUESTION FOUR

    a) Define the following terms;

        i)   Nominal data   (2 marks)

        ii)   Ordinal data   (2 marks)

        iii)   Interval data   (2 marks)

    b) Explain the importance of using softwares in data analysis.   (10 marks)

    c) Distinguish reliability testing from validity testing of a questionnaire.   (4 marks)

        [20 marks]

QUESTION FIVE

    a) Describe the measures of central tendency   (7 marks)

    b) Explain the measures of dispersion   (7 marks)

    c) Explain the assumptions underpinning analysis of variance (ANOVA)   (6 marks)

        [20 marks]

END OF EXAMINATION