|  |  |  |
| --- | --- | --- |
| **CHUKA** |  | **UNIVERSITY** |

**UNIVERSITY EXAMINATIONS**

**EXAMINATION FOR THE AWARD OF MASTERS OF SCIENCE DEGREE IN AGRICULTURAL EDUCATION**

**AGED 810: SOCIAL STATISTICS**

**STREAMS: MSc. AGED**

**TIME: 3 HOURS**

**DAY/DATE: TUESDAY 13 /07/ 2021 8.30 AM – 11.30 PM**

**INSTRUCTIONS:**

* Answer Question One and any other Three Questions
* Do not write on the question paper

1. a) What is a Singular Matrix

b) Differentiate between:

i. Parametric and non-parametric test. [4 Marks]

ii. One way ANOVA and Two-way ANOVA. [4 Marks]

c) Solve the following linear system of equations using the method of determinants:

x + 4z = 4

4x + y – 2z = 0

3x + y – z = 2

[7 Marks]

2. a) Calculate the covariance for the following data:

|  |  |  |  |
| --- | --- | --- | --- |
| **Item Number** | **1** | **2** | **3** |
| x | 2 | 1 | 3 |
| y | 1 | 2 | -3 |
| z | -1 | -1 | 4 |

[7 Marks]

b) Find eigen values and eigen vectors of matix A =

[8 Marks]

3. a) distinguish between chi-square test of goodness of fit and chi-square test of independence. [2 Marks]

b) Students in a social studies class hypothesize that the literacy rates across the world for every region are 82%. The table below shows the actual literacy rates across the world broken down by region. Using 5% level of significance test if their hypothesis fits the regional adult literacy rates? [7 Marks]

|  |  |
| --- | --- |
| **MDG Region** | **Adult Literacy Rate (%)** |
| Developed Regions | 99.0 |
| Commonwealth of Independent States | 99.5 |
| Northern Africa | 67.3 |
| Sub-Saharan Africa | 62.5 |
| Latin America and the Caribbean | 91.0 |
| Eastern Asis | 93.8 |
| Southern Asia | 61.9 |
| South-Eastern Asia | 91.9 |
| Western Asia | 84.5 |
| Oceania | 66.4 |

c) The following table shows the results of 8 students in “Business Mathematics” and “Statistics”.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Business Mathematics [%] | 68 | 54 | 19 | 72 | 50 | 44 | 92 | 37 |
| Statistics [%] | 51 | 76 | 32 | 85 | 62 | 25 | 74 | 59 |

Find and interpret the value of Spearman’s rank correlation coefficient? [6 Marks]

4. a) The World Bank collects information on the life expectancy of a person in each country and the fertility rate per woman in the country. The data for 24 randomly selected countries for the year 2011 are in table#1.

Table #1: Data of Fertility Rates versus Life Expectancy

|  |  |
| --- | --- |
| Fertility Rate | Life Expectancy |
| 1.7. | 77.2 |
| 5.8 | 55.4 |
| 2.2 | 69.9 |
| 2.1 | 76.4 |
| 1.8 | 75.0 |
| 2.0 | 78.2 |
| 2.6 | 73.0 |
| 2.8 | 70.8 |
| 1.4 | 82.6 |
| 2.6 | 68.9 |
| 1.5 | 81.0 |
| 6.9 | 54.2 |
| 2.4 | 67.1 |
| 1.5 | 73.3 |
| 2.5 | 74.2 |
| 1.4 | 80.7 |
| 2.9 | 72.1 |
| 2.1 | 78.3 |
| 4.7 | 62.9 |
| 6.8 | 54.4 |
| 5.2 | 55.9 |
| 4.2 | 66.0 |
| 1.5 | 76.0 |
| 3.9 | 72.3 |

i. Create a scatter plot of the data and find a linear regression equation between fertility rate and life expectancy. [5 Marks]

ii. Then use the regression equation to find the life expectancy for a country that has a fertility rate of 2.7 and for a country with fertility rate of 8.1. [3 Marks]

iii. Which life expectancy that you calculated do you think is closer to the true-life expectancy? Why? [2 Marks]

b) A pharmaceutical company is considering introducing new, easier to open, packaging for a drug used in the treatment of arthritis. The company seeks the views of two groups of patients. One group consists of those who have been using the existing packaging for along time and the other group consists of new users. The preferences are shown below.

|  |  |  |
| --- | --- | --- |
|  | Prefer new packaging | Prefer existing packaging |
| Long-term users | 35 | 32 |
| Recent users | 25 | 8 |

Investigate, at the 5% significance level, whether there is evidence of a difference between the two groups of patients in their packaging preferences and briefly state your conclusions.

[5 Marks]

5. a) What are the assumptions for Two-way ANOVA. [2 Marks]

b) Waste cooling water from a large engineering works is filtered before being released into the environment. Three separate discharge pipes are used, each with its own filter. Five samples of water are taken on each of four days from each of the three discharge pipes and the concentrations of a pollutant, in parts per million, are measured. The data are given below. Analyse the data to test for differences between the discharge pipes. Allow for effects due to pipes and days and for an interaction effect. Treat the pipe effects as fixed and the day effects as random. Use the 5% level of significance. [13 Marks]

|  |  |
| --- | --- |
| Day | Pipe A |
| 1  2  3  4 | 160 181 163 173 178  175 170 219 166 171  169 186 179 178 183  230 206 216 195 250 |
| Day | Pipe B |
| 1  2  3  4 | 172 164 186 185 172  177 170 156 140 155  193 194 189 156 181  212 235 195 206 209 |
| Day | Pipe C |
| 1  2  3  4 | 214 196 207 219 200  186 184 181 189 179  209 220 199 185 228  254 293 283 262 259 |

6. a) Suppose you computed r =0.627 with 12 data points, test the significance of the correlation. [5 Marks]

b) An investigator theorizes that people who participate in a regular program of exercise will have levels of systolic blood pressure that are significantly different from that of people who do not participate in a regular program of exercise. To test this idea the investigator randomly assigns 21 subjects to an exercise program for 10 weeks and 21 subjects to a non-exercise comparison group. After ten weeks the mean systolic blood pressure of subjects in the exercise group is 137 and the standard deviation of blood pressure values in the exercise group is 10. After ten weeks, the mean systolic blood pressure of subjects in the non-exercise group is 127 and the standard deviation on subjects in the non-exercise group is 9.0. Go through the steps of hypothesis testing to confirm the investigator’s theory using an alpha level of .0.5. [10 Marks]

………………………………………………………………………………………………………