**IB MATH STUDIES EXAM REVIEW: Topic 4**

**Chi-Squared Test, Correlation Coefficient, Line of Best Fit, Normal Distribution**

**1.** Some of the customers in each café were given survey forms to complete to find out if they were satisfied with the standard of service they received.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Pete’s Eats** | **Alan’s Diner** | **Sarah’s Snackbar** | **Total** |
| **Dissatisfied** | 16 | 8 | 16 | 40 |
| **Satisfied** | 26 | 20 | 34 | 80 |
| **Total** | 42 | 28 | 50 | 120 |

A *χ*2 test at the 5 % significance level was carried out to determine whether there was any difference in the level of customer satisfaction in each of the cafés.

(a) Write down the null hypothesis, H0, for the *χ*2 test.

(1)

(b) Write down the number of degrees of freedom for the test.

(1)

(c) Using your graphic display calculator, find .

(2)

(d) State, giving a reason, the conclusion to the test.

(2)

(Total 6 marks)

**2.** A survey was carried out at a university amongst 150 Science students. These students all studied one of either French, Spanish or Russian. The results of the survey are shown below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **French** | **Spanish** | **Russian** |
| **Female** | 9 | 29 | 12 |
| **Male** | 31 | 40 | 29 |

Ludmila decides to use the *χ*2 test at the 5 % level of significance to determine whether the choice of language is independent of gender.

(a) State Ludmila’s null hypothesis.

(1)

(b) Write down the number of degrees of freedom.

(1)

(c) Find the expected frequency for the females studying Spanish.

(2)

(d) Use your graphic display calculator to find the χ2 test statistic for this data.

(2)

(e) State whether Ludmila accepts the null hypothesis. Give a reason for your answer.

(2)

(Total 8 marks)

**3.** At the end of the year, only seven female students sat examinations in Science and French.

The marks for these seven students are shown in the following table.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Science (*S*)** | 23 | 51 | 56 | 62 | 12 | 73 | 72 |
| **French (*F*)** | 65 | 45 | 45 | 40 | 70 | 36 | 30 |

(a) Using a scale of 2 cm to represent 10 marks for each axis, draw a labelled scatter diagram for this data.

(4)

(b) Use your graphic display calculator to find

(i) *,* the mean of *S*;

(ii) , the mean of *F.*

(2)

(c) Plot the point M()on your scatter diagram.

(1)

(d) Use your graphic display calculator to find the equation of the regression line of *F* on *S.*

(2)

(e) Draw the regression line on your scatter diagram.

(2)

Carletta’s mark on the Science examination was 44. She did not sit the French examination.

(f) Estimate Carletta’s mark for the French examination.

(2)

Monique’s mark on the Science examination was 85. She did not sit the French examination. Her French teacher wants to use the regression line to estimate Monique’s mark.

(g) State whether the mark obtained from the regression line for Monique’s French examination is reliable. Justify your answer.

(2)

(Total 15 marks)

**4.** The heat output in thermal units from burning 1 kg of wood changes according to the wood’s percentage moisture content. The moisture content and heat output of 10 blocks of the same type of wood each weighing 1 kg were measured. These are shown in the table.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Moisture content % (*x*)** | 8 | 15 | 22 | 30 | 34 | 45 | 50 | 60 | 74 | 82 |
| **Heat output ( *y*)** | 80 | 77 | 74 | 69 | 68 | 61 | 61 | 55 | 50 | 45 |

(a) Draw a scatter diagram to show the above data. Use a scale of 2 cm to represent 10 % on the *x-*axis and a scale of 2 cm to represent 10 thermal units on the *y-*axis.

(4)

(b) Write down

(i) the mean percentage moisture content, ;

(ii) the mean heat output, *.*

(2)

(c) Plot the point ()on your scatter diagram and label this point M.

(2)

(d) Write down the product-moment correlation coefficient, *r.*

(2)

The equation of the regression line *y* on *x* is *y* = –0.470*x* + 83.7.

(e) Draw the regression line *y* on *x* on your scatter diagram.

(2)

(f) Find the heat output in thermal units of a 1 kg block of wood with 25% moisture content.

(2)

(g) State, with a reason, whether it is appropriate to use the regression line *y* on *x* to estimate the heat output in part (f).

(2)

(Total 16 marks)

**5.** The times, in minutes, that runners take to complete a 10 km race are distributed such that .

(a) Find the probability that a randomly selected runner takes more than 80 minutes to complete the 10 km race.

(2)

A total of 5000 runners participate in the 10 km race.

(b) Find the number of runners that will complete the 10 km race in less than one hour.

(4)

(c) Simon is the 1000th person to complete the 10 km race. How long does Simon take?

(3)

(Total 9 marks)

**6.** The weights of babies born at Prince Louis Maternity Hospital last year followed a normal distribution with a mean of 3.0 kg and a standard deviation of 200 **grams**.

(a) Find the percentage of babies that weigh

(i) between 2.8 kg and 3.4 kg;

(ii) less than 2.91 kg.

(4)

(b) 600 babies were born at the hospital last year. Find the number of babies that weighed more than 3.1 kg.

(4)

(c) All babies born at the hospital are weighed, and 18% of them are classified as overweight.

(i) Draw a diagram of this situation with the appropriate region shaded and labeled;

(ii) Find the maximum weight of the babies that are not classified as overweight.

(5)

(Total 13 marks)