**1** A sports club carries out a survey of prospective members to find out which of three sports they play regularly.

 A random sample of 120 prospective members gave the following results.

|  |  |  |
| --- | --- | --- |
|  |  | Sport |
|  |  | Hockey | Cricket | Rugby |
| Gender | **Male** | 25 | 17 | 16 |
| **Female** | 18 | 21 | 23 |

 A test is carried out at the 5% level of significance to determine whether or not there is an association between gender and the sport the members play regularly.

**a** State the null hypothesis for this test. **(1 mark)**

**b** Calculate the number of degrees of freedom. **(1 mark)**

**c** Find the critical value appropriate for the test. **(1 mark)**

**d** Show that the expected frequency for male rugby players is 18.85. **(1 mark)**

**e** Calculate the test statistic for this test. **(3 marks)**

**f** State whether or not the null hypothesis is rejected. Justify your answer. **(1 mark)**

**g** Explain whether or not the null hypothesis would be rejected if the test was carried out at the 10% level of significance. **(1 mark)**

**2** A second hand car dealer records the number of cars sold each day over a 30 day period.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Number of cars sold | 0 | 1 | 2 | 3 | 4 | 5 |
| Frequency | 9 | 12 | 5 | 3 | 0 | 1 |

 The dealer uses a Poisson distribution with mean 1.2 to calculate expected frequencies.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Number of cars sold | 0 | 1 | 2 | 3 | 4 or more |
| Expected frequency | 9.04 | *x* | 6.51 | 2.60 | *y* |

**a** Find, correct to two decimal places, the values of *x* and *y*. **(2 marks)**

The dealer tests, at the 10% level of significance, whether or not the data can be modelled by a Poisson distribution.

**b** Calculate the test statistic and state the conclusion for this test. State clearly your hypotheses and the number of degrees of freedom used in the test. **(6 marks)**

**3** A five-sided spinner with sections numbered 1–5 is spun 80 times and the results shown in a table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Number on spinner** | 1 | 2 | 3 | 4 | 5 |
| **Observed frequency** | 18 | 15 | 18 | 19 | 10 |

**a** State the distribution that should be used to model the spinner if it is ‘fair’. **(1 mark)**

**b** Test, at the 10% level of significance, whether or not the observed frequencies could be modelled by your stated distribution. State clearly your hypotheses and the number of degrees of freedom used. **(6 marks)**

**4** A cat breeder was investigating the number of female kittens born when the litter contained five kittens. She collected data on 100 litters and recorded the data in a table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Number of female kittens | 0 | 1 | 2 | 3 | 4 | 5 |
| Frequency | 12 | 19 | 37 | 19 | 8 | 5 |

A colleague suggests that the distribution may be modelled by a binomial distribution with *p* = 0.5.

**a** By testing at the 5% level of significance, show that this model is not a good model. **(10 marks)**

The breeder still believes that a binomial model is suitable for modelling the distribution.

**b** Test, at the 5% level of significance, the breeder’s claim. **(8 marks)**

**5** Each day, Susane, who works as a statistician, picks a random pen from a large pot and records the number of pens that she has to pick before finding one that works. She records this number on 120 separate occasions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Number of pens** | 1 | 2 | 3 | 4 | 5 or more |
| **Frequency** | 80 | 25 | 10 | 4 | 1 |

 She believes that the number of pens she has to pick can be modelled as a Geo(0.6) distribution.

 Test Susane’s belief at the 2.5% level of significance. State your hypotheses clearly. **(8 marks)**