4.7 Hypothesis Testing

Question Paper

Course	DP IB Maths
Section	4. Statistics & Probability
Topic	4.7 Hypothesis Testing
Difficulty	Medium

Time allowed: 90

Score: /71

Percentage: /100

Question 1a

At a school in Copenhagen, it is believed that favourite music genre is related to gender. 400 students were asked to indicate their favourite music genre from a selection and the results are indicated in the table below.

	Pop	Rock	Classical	Rap
Female	58	63	17	44
Male	23	96	12	87

It is decided to test this hypothesis by using a χ^2 test at the 5% significance level. The critical value is 7.815.

(a) State the null and alternative hypotheses for this test.

[2 marks]

Question 1b

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

[1 mark]

Question 1c

(c) Calculate the χ^2 test statistic for this data.

[2 marks]

Question 1d

(d) What conclusion can be drawn from this test? Give a reason for your answer.

[2 marks]

Question 2a

An environmental organisation is trying to establish if altitude affects the growth of pine needles. A number of needles have been taken from trees at both high and low altitudes and their lengths, in inches, recorded. The results are shown in the table below.

Low altitude	6.1	8.2	7.7	8.0	11.9	6.9	7.5	7.1	8.1
High altitude	7.4	7.9	8.3	6.6	9.5	7.9	8.2	8.1	8.5

Perform a t-test to compare the mean lengths of the pine needles.

(a) Write down the null and alternative hypotheses.

[2 marks]

Question 2b

(b) State whether this is a one-tailed test or a two-tailed test.

[1 mark]

Question 2c

(c) Perform a *t*-test at the 10% significance level. Write down the *p*-value.

[2 marks]

Question 2d

(d) Write down the conclusion of the test. Give a reason for your answer.

[2 marks]

Question 3a

A carpet salesman in interested how his sales are distributed and records his sales results over a period of six months. The data is shown in the table.

Month	January	February	March	April	May	June
Number of sales	16	12	14	20	15	19

A chi-squared goodness of fit test is to be performed on the data at the 5% significance level to find out whether the data fits a uniform distribution.

(a)   Find an estimate of how many carpets the salesman expects to sell each month.

[1 mark]

Question 3b

(b)   Write down the null and alternative hypotheses.

[2 marks]

Question 3c

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

[1 mark]

Question 3d

(d)   Calculate the p-value.

[2 marks]

Question 3e

(e)   State the conclusion of the test. Give a reason for your answer.

[2 marks]

Question 4a

A supermarket is interested in how the applications for its loyalty scheme are distributed throughout the working week. It is expected that the distribution of the data will be uniform. Over the course of one week the number of applications has been collected and recorded in the table.

Month	Monday	Tuesday	Wednesday	Thursday	Friday
Number of sales	473	405	512	467	503

A goodness of fit test at the 10% significance level is to be performed.
The critical value is 7.779.

(a) Calculate the expected value of the number of applications on any given workday.

[1 mark]

Question 4b

(b) State the null and alternative hypotheses.

[2 marks]

Question 4c

(c) Write down the

    (i) χ^2 statistic

    (ii) p-value.

[3 marks]

Question 4d

(d) State whether the data fits a uniform distribution, giving a reason for your answer.

[2 marks]

Question 5a

A chi-squared test is performed to see if there is any dependence between eye colour (blue, green or brown) and hair colour (blond, red, brown or black). The test is completed at a significance level of 10%.

(a) State the null and alternative hypotheses.

[2 marks]

Question 5b

(b) Find the number of degrees of freedom for this test.

[2 marks]

Question 5c

(c) The *p*-value for this test is 0.0726. State the conclusion that can be drawn and justify your answer.

[2 marks]

Question 6a

It is claimed that women from Japan are taller on average than women from India. The heights, in cm, of 11 women from each country have been collated in the table below.

Japan	India
173.0	155.2
158.2	157.8
148.5	156.0
150.6	142.7
168.7	149.6
149.8	150.1
158.8	152.6
155.3	148.2
159.2	151.3
158.9	147.6
166.0	168.0

(a) Write down the type of test that can be used to compare the means of two sets of data.

[2 marks]

Question 6b

(b) State the null and alternative hypotheses.

[2 marks]

Question 6c

(c) Perform the appropriate test at the 5% significance level.

[2 marks]

Question 6d

(d) State whether or not the initial claim is justified. Give a reason for your answer.

[2 marks]

Question 7a

The average weight of a newborn baby born at 38 weeks is expected to be less than the average weight of a newborn born at full term (40 weeks). The weights of several babies, in kg, born at 38 weeks and 40 weeks in one hospital are recorded.

38 weeks	3.12	2.87	3.53	3.08	2.86	3.15	3.03	2.99		
40 weeks	3.08	3.59	3.49	3.61	2.99	3.58	3.42	3.55	3.66	3.58

A *t*-test is to be performed at a significance level of 10%.

(a) State the null and alternative hypotheses.

[2 marks]

Question 7b

(b) State whether a one-tailed or a two-tailed test should be used.

[1 mark]

Question 7c

(c) Calculate the *p*-value statistic.

[2 marks]

Question 7d

(d) State whether the initial expectation is confirmed by the test result. Justify your answer.

[2 marks]

Question 8a

A survey was conducted to establish whether particular colours were favoured by people in different age groups. A group of 280 children, teenagers and adults were asked to pick their favourite colour from a choice of red, yellow, blue, green and pink.

The observed values are recorded in the table below.

	Red	**Yellow**	**Blue**	**Green**	**Pink**	**Total**
Children	20	11	18	8	15	72
Teenagers	22	14	23	20	6	85
Adults	26	21	30	26	20	123
Total	68	46	71	54	41	280

A chi-squared test is to be performed at the 5% significance level. The critical value for the test is 15.507.

(a) Complete the contingency table for the expected values below.

	Red	**Yellow**	**Blue**	**Green**	**Pink**	**Total**
Children						
Teenagers						
Adults						
Total						

[3 marks]

Question 8b

(b) State the null and alternative hypotheses.

[2 marks]

Question 8c

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

[1 mark]

Question 8d

(d) Write down the chi-squared test statistic for this data.

[2 marks]

Question 8e

(e) Comment on your results within the context of the question.

[2 marks]

Question 9a

The heights of giraffes are normally distributed with a mean of 3.8 m and a standard deviation of 0.7 m. As part of a conservation project in Kenya, the heights of 350 giraffes are measured and the results of the survey are seen in the table below. A chi-squared test at a significance level of 10% is to be performed to determine if the surveyed giraffes fit the normal distribution stated.

Height (cm)	Frequency
$h < 3$	50
$3 \leq h < 4$	160
$4 \leq h < 5$	119
$5 \leq h < 6$	21

(a) Complete the following table of expected heights.

Height (cm)	Probability	Expected frequency
$h < 3$		
$3 \leq h < 4$		
$4 \leq h < 5$		
$5 \leq h < 6$		

[3 marks]

Question 9b

(b) Write down the null and alternative hypotheses.

[2 marks]

Question 9c

(c) State the number of degrees of freedom.

[1 mark]

Question 9d

(d) Calculate the p-value.

[2 marks]

Question 9e

(e) State whether the results of the chi-squared test support the null hypothesis. Justify your answer.

[2 marks]