Chi squared test [68 marks]

1. [Maximum mark: 5] SPM.1.SL.TZ0.6

As part of a study into healthy lifestyles, Jing visited Surrey Hills University. Jing recorded a person's position in the university and how frequently they ate a salad. Results are shown in the table.

	Salad meals per week			
	0	**1 – 2**	**3 – 4**	**> 4**
Students	45	26	18	6
Professors	15	8	5	12
Staff and Administration	16	13	10	6

Jing conducted a χ^2 test for independence at a 5 % level of significance.

(a) State the null hypothesis. [1]

(b) Calculate the p-value for this test. [2]

(c) State, giving a reason, whether the null hypothesis should be accepted. [2]

2. [Maximum mark: 16] EXN.2.SL.TZ0.4

A dice manufacturer claims that for a novelty die he produces the probability of scoring the numbers 1 to 5 are all equal, and the probability of a 6 is two times the probability of scoring any of the other numbers.

(a) Find the probability of scoring a six when rolling the novelty die. [3]

(b) Find the probability of scoring more than 2 sixes when this die is rolled 5 times. [4]

To test the manufacture's claim one of the novelty dice is rolled 350 times and the numbers scored on the die are shown in the table below.

Number scored	Frequency
1	32
2	57
3	47
4	58
5	54
6	102

(c.i) Find the expected frequency for each of the numbers if the manufacturer's claim is true. [2]

A χ^2 goodness of fit test is to be used with a 5% significance level.

(c.ii) Write down the null and alternative hypotheses. [2]

(c.iii) State the degrees of freedom for the test. [1]

(c.iv) Determine the conclusion of the test, clearly justifying your answer. [4]

3. [Maximum mark: 5] 22N.1.SL.TZ0.4

Sergio is interested in whether an adult's favourite breakfast berry depends on their income level. He obtains the following data for 341 adults and decides to carry out a χ^2 test for independence, at the 10% significance level.

		Income level		
		Low	Medium	High
Favourite berry	Strawberry	21	39	30
	Blueberry	39	67	42
	Other berry	32	45	26

(a) Write down the null hypothesis. [1]

(b) Find the value of the χ^2 statistic. [2]

The critical value of this χ^2 test is 7.78.

(c) Write down Sergio's conclusion to the test in context. Justify your answer. [2]

4. [Maximum mark: 9] EXM.1.SL.TZ0.9

Six coins are tossed simultaneously 320 times, with the following results.

0 tail	5 times
1 tail	40 times
2 tails	86 times
3 tails	89 times
4 tails	67 times
5 tails	29 times
6 tails	4 times

At the 5% level of significance, test the hypothesis that all the coins are fair. [9]

5. [Maximum mark: 6] 23M.1.SL.TZ2.6

A company that owns many restaurants wants to determine if there are differences in the quality of the food cooked for three different meals: breakfast, lunch and dinner.

Their quality assurance team randomly selects 500 items of food to inspect. The quality of this food is classified as perfect, satisfactory, or poor. The data is summarized in the following table.

		Quality			
		Perfect	**Satisfactory**	**Poor**	**Total**
Meal	**Breakfast**	101	124	7	**232**
	Lunch	68	81	5	**154**
	Dinner	35	69	10	**114**
	Total	**204**	**274**	**22**	**500**

An item of food is chosen at random from these 500.

(a) Find the probability that its quality is not perfect, given that it is from breakfast. [2]

A χ^2 test at the 5% significance level is carried out to determine if there is significant evidence of a difference in the quality of the food cooked for the three meals.

The critical value for this test is 9.488.

The hypotheses for this test are:

H_0: The quality of the food and the type of meal are independent.
H_1: The quality of the food and the type of meal are not independent.

(b) Find the χ^2 statistic. [2]

(c) State, with justification, the conclusion for this test. [2]

6. [Maximum mark: 9] EXM.1.SL.TZ0.11

A calculator generates a random sequence of digits. A sample of 200 digits is randomly selected from the first 100 000 digits of the sequence. The following table gives the number of times each digit occurs in this sample.

digit	0	1	2	3	4	5	6	7	8	9
frequency	17	21	15	19	25	27	19	23	18	16

It is claimed that all digits have the same probability of appearing in the sequence.

(a) Test this claim at the 5% level of significance. [7]

(b) Explain what is meant by the 5% level of significance. [2]

7. [Maximum mark: 11] EXM.1.SL.TZ0.8

In an effort to study the level of intelligence of students entering college, a psychologist collected data from 4000 students who were given a standard test. The predictive norms for this particular test were computed from a very large population of scores having a normal distribution with mean 100 and standard deviation of 10. The psychologist wishes to determine whether the 4000 test scores he obtained also came from a normal distribution with mean 100 and standard deviation 10. He prepared the following table (expected frequencies are rounded to the nearest integer):

Score	observed frequencies	expected frequencies	Score	observed frequencies	expected frequencies
≤ 70.5	20	6	100.5–110.5	1450	
70.5–80.5	90	96	110.5–120.5	499	507
80.5–90.5	575		120.5–130.5	80	76
90.5–100.5	1282		≥ 130.5	4	4

(a) Copy and complete the table, showing how you arrived at your answers. [5]

(b) Test the hypothesis at the 5% level of significance. [6]

8. [Maximum mark: 7] 22M.1.SL.TZ1.7

Leo is investigating whether a six-sided die is fair. He rolls the die 60 times and records the observed frequencies in the following table:

Number on die	1	2	3	4	5	6
Observed frequency	8	7	6	15	12	12

Leo carries out a χ^2 goodness of fit test at a 5% significance level.

(a) Write down the null and alternative hypotheses. [1]

(b) Write down the degrees of freedom. [1]

(c) Write down the expected frequency of rolling a 1. [1]

(d) Find the p-value for the test. [2]

(e) State the conclusion of the test. Give a reason for your answer. [2]