Answer **all** questions in the answer booklet provided. Please start each question on a new page. You are advised to show all working, where possible. Where an answer is incorrect, some marks may be given for a correct method, provided this is shown by written working. Solutions found from a graphic display calculator should be supported by suitable working, for example, if graphs are used to find a solution, you should sketch these as part of your answer.

1. [Maximum mark: 15]

Sila High School has 110 students. They each take exactly one language class from a choice of English, Spanish or Chinese. The following table shows the number of female and male students in the three different language classes.

		Language Choice			
		English	**Spanish**	**Chinese**	**Total**
Student Gender	**Female**	25	8	10	43
	Male	21	14	32	67
	Total	46	22	42	110

A χ^2 test was carried out at the 5% significance level to analyse the relationship between gender and student choice of language class.

(a) Write down the null hypothesis, H_0, for this test. [1]

(b) State the number of degrees of freedom. [1]

(c) Use your graphic display calculator to write down

(i) the expected frequency of female students who chose to take the Chinese class;

(ii) the χ^2 statistic. [3]

The critical value at the 5% significance level for this test is 5.99.

(d) State whether or not H_0 should be rejected. Justify your statement. [2]

(e) One student is chosen at random from this school.

(i) Find the probability that the student does not take the Spanish class.

Another student is chosen at random from this school.

(ii) Find the probability that neither of the two students take the Spanish class.

(iii) Find the probability that at least one of the two students is female. [8]

1. A survey of 400 people is carried out by a market research organization in two different cities, Buenos Aires and Montevideo. The people are asked which brand of cereal they prefer out of Chocos, Zucos or Fruti. The table below summarizes their responses.

	Chocos	Zucos	Fruti	Total
Buenos Aires	43	85	62	190
Montevideo	57	35	118	210
Total	100	120	180	400

(a) One person is chosen at random from those surveyed. Find the probability that this person

(i) does not prefer Zucos;

(ii) prefers Chocos, given that they live in Montevideo.

(4)

(b) Two people are chosen at random from those surveyed. Find the probability that they both prefer Fruti.

(3)

The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed.

(c) State the null hypothesis.

(1)

(d) State the number of degrees of freedom.

(1)

(e) Show that the expected frequency for the number of people who live in Montevideo and prefer Zucos is 63.

(2)

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

(2)

(g) State whether the market research organization would accept the null hypothesis. Clearly justify your answer.

(2)

(Total 15 marks)

2. Manuel conducts a survey on a random sample of 751 people to see which television programme type they watch most from the following: Drama, Comedy, Film, News. The results are as follows.

	Drama	Comedy	Film	News
Males under 25	22	65	90	35
Males 25 and over	36	54	67	17
Females under 25	22	59	82	15
Females 25 and over	64	39	38	46

Manuel decides to ignore the ages and to test at the 5% level of significance whether the most watched programme type is independent of **gender**.

(a) Draw a table with 2 rows and 4 columns of data so that Manuel can perform a chi-squared test.

(3)

(b) State Manuel's null hypothesis and alternative hypothesis.

(1)

(c) Find the expected frequency for the number of females who had "Comedy" as their most-watched programme type. Give your answer to the nearest whole number.

(2)

(d) Using your graphic display calculator, or otherwise, find the chi-squared statistic for Manuel's data.

(3)

(e) (i) State the number of degrees of freedom available for this calculation.

(ii) State the critical value for Manuel's test.

(iii) State his conclusion.

(3)

(Total 12 marks)

3. (a) For his Mathematical Studies project, Marty set out to discover if stress was related to the amount of time that students spent travelling to or from school. The results of one of his surveys are shown in the table below.

Travel time (t mins) ↓	Number of students: high stress	moderate stress	low stress
$t \leq 15$	9	5	18
$15 < t \leq 30$	17	8	28
$30 < t$	18	6	7

He used a χ^2 test at the 5% level of significance to find out if there was any relationship between student stress and travel time.

(i) Write down the null and alternative hypotheses for this test.

(2)

(ii) Write down the table of expected values. Give values to the nearest integer.

(3)

(iii) Show that there are 4 degrees of freedom.

(1)

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

(2)

The χ^2 critical value for 4 degrees of freedom at the 5% level of significance is 9.488.

(v) What conclusion can Marty draw from this test? Give a reason for your answer.

(2)

(b) Marty asked some of his classmates to rate their level of stress out of 10, with 10 being very high. He also asked them to measure the number of minutes it took them to get from home to school. A random selection of his results is listed below.

Travel time (x)	13	24	22	18	36	16	14	20	6	12
Stress rating (y)	3	7	5	4	8	8	4	8	2	6

(i) Write down the value of the (linear) coefficient of correlation for this information.

(1)

(ii) Explain what a positive value for the coefficient of correlation indicates.

(1)

(iii) Write down the linear regression equation of y on x in the form $y = ax + b$

(2)

(iv) Use your equation in part (iii) to determine the stress rating for a student who takes three quarters of an hour to travel to school.

(2)

(v) Can your answer in part (iv) be considered reliable? Give a reason for your answer.

(2)

(Total 18 marks)

4. In a competition the number of males and females taking part in different swimming races is given in the table of observed values below.

	Backstroke (100 m)	Freestyle (100 m)	Butterfly (100 m)	Breaststroke (100 m)	Relay (4 × 100 m)
Male	30	90	31	29	20
Female	28	63	20	37	12

The Swimming Committee decides to perform a χ^2 test at the 5% significance level in order to test if the number of entries for the various strokes is related to gender.

(a) State the null hypothesis.

(1)

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

(1)

(c) Write down the critical value of χ^2.

(1)

The expected values are given in the table below:

	Backstroke (100 m)	Freestyle (100 m)	Butterfly (100 m)	Breaststroke (100 m)	Relay (4 × 100 m)
Male	32	a	28	37	18
Female	26	68	23	b	14

(d) Calculate the values of a and b.

(2)

(e) Calculate the χ^2 value.

(3)

(f) State whether or not you accept your null hypothesis and give a reason for your answer.

(2)

(Total 10 marks)

5. The veterinarian has gathered the following data about the weight of dogs and the weight of their puppies.

		Dog		**Total**
		Heavy	Light	
	Heavy	36	27	63
Puppy	Light	22	35	57
	Total	58	62	120

The veterinarian wishes to test the following hypotheses.

H_0: A puppy's weight is independent of its parent's weight.
H_1: A puppy's weight is related to the weight of its parent.

(a) The table below sets out the elements required to calculate the χ^2 value for this data.

	f_o	f_e	$f_e - f_o$	$(f_e - f_o)^2$	$(f_e - f_o)^2 / f_e$
heavy/heavy	36	30.45	–5.55	30.8025	1.012
heavy/light	27	32.55	5.55	30.8025	0.946
light/heavy	22	27.55	5.55	30.8025	1.118
light/light	35	*a*	*b*	*c*	*d*

(i) Write down the values of *a*, *b*, *c*, and *d*.

(4)

(ii) What is the value of χ^2_{calc} for this data?

(1)

(iii) How many degrees of freedom exist for the contingency table?

(1)

(iv) Write down the critical value of χ^2 for the 5% significance level.

(1)

(b) Should H_0 be accepted? Explain why.

(2)

(Total 9 marks)