Question 1

A sample of individuals were randomly assigned one of two diet plans. Over a 6-week period these individuals followed their assigned plan. Their weight loss was recorded at the end of the 6-week period and the results were as follows:

	Plan 1	Plan 2
sample size	42	50
mean weight loss	7.1 lbs	5.2 lbs
variance	10.1 lbs^2	16.1 lbs^2

We wish to test the hypothesis that there is no difference between diet plans.

(a) State the null and alternative hypotheses.   **(b)** Calculate the test statistic.
(c) Calculate the p-value.   **(d)** What is your conclusion?

Question 2

A company claims that it pays men and women equally. The salaries for some randomly selected employees (in thousands) were recorded and the results were as follows:

	Male	Female
sample size	5	3
mean salary	30.2	28.4
standard deviation	1.7	1.9

(a) The F-test was carried out and a p-value of 0.7297 was obtained. What does this mean?
(b) We wish to test the hypothesis that there is no difference between salaries - what are the null and alternative hypotheses?   **(c)** If testing at the 10% level, what is the rejection region? (note your answer to part (a)).   **(d)** Is there evidence to suggest gender inequality?

Question 3

A sample of students from two universities was randomly selected. Each student had to complete the same programming task and the time to completion was recorded in each case. The results were as follows:

	University A	University B
sample size	15	15
mean	12.5 hrs	11.1 hrs
variance	3 hrs^2	1.5 hrs^2

We wish to test the hypothesis that there is no difference between universities at the 5% level.

(a) State the null and alternative hypotheses.   **(b)** If we do *not* assume equal variances, what are the critical values?   **(c)** Calculate the test statistic and, hence, provide your conclusion.
(d) Between what two values does the p-value lie? (note: the p-value cannot be calculated exactly using the t-tables)

Question 4

The government wish to know if there is a difference in the proportions of people living in rural and urban areas in support of a new policy. From a sample of 38 people in rural areas, it was found that 20 support the policy and from a sample of 116 individuals in urban areas, it was found that 70 support the policy.

(a) Test the hypothesis that there is no difference in proportions at the 5% level.

Question 5

A sample of 100 individuals were asked which product they prefer and the results were as follows:

	Product 1	Product 2	Product 3	Product 4	Product 5	$\sum$
Frequency	19	24	24	14	19	100

(a) If there was no difference between products, what would the expected frequencies be?
(b) Test the hypothesis that the observed matches the expected (use $\alpha = 0.05$).

Question 6

We wish to test the hypothesis that the following sample has come from a normal distribution:

x	< 5	$5-7$	$7-9$	$9-11$	$11-13$	$13-15$	$15-17$	> 17	$\sum$
Frequency	3	10	23	62	39	14	6	3	160

where the mean and standard deviation were also calculated: $\bar{x} = 10.4$ and $s = 2.5$. With these estimates for μ and σ it is easy to calculate the following probabilities (normal tables):

x	< 5	$5-7$	$7-9$	$9-11$	$11-13$	$13-15$	$15-17$	> 17	$\sum$
p_i	0.015	0.072	0.201	0.307	0.256	0.116	0.029	0.004	1.00

(a) Calculate the expected frequencies (group classes where $e_i < 5$). **(b)** Calculate the test statistic. **(c)** Can we reject the hypothesis that the 5% level? **(c)** Between what two values does the p-value lie?

Question 7

One hundred computer science graduates from each of three different universities were asked how many programming languages they are competent in. The results were as follows:

		Languages 1	2	3	4+	$\sum$
University	A	16	38	39	7	100
	B	18	29	41	12	100
	C	28	31	38	3	100
	$\sum$	62	98	118	22	300

(a) Calculate the expected frequencies assuming independence of the two variables.
(b) Calculate the χ^2 statistic and, hence, a range within which the p-value lies.
(c) What is your conclusion? **(d)** Calculate the raw difference scores $(o_i - e_i)$ and comment.