Question 1 (4 points)

The following information should be used in conjunction with questions 1 to 6.

Dave's Diners, a chain of cafes, is planning to open another one. Dave (the owner) believes that the size of the student population at nearby university campuses is related to sales revenue in the cafes. Dave and his management team collect data from a sample of its diners located near college campuses. The table below summarises their findings.

Student Population (£1000's)	2	6	8	8	12	16	20	20	22	26
Annual Sales (£1000s)	58	105	88	118	117	137	157	169	149	202

Which option gives the least squares fit straight line to this data so that it may be used to predict sales from the size of the student population?

- a. students = 5 + 60 sales
- b. students = 0.12 + 0.2 sales
- c. sales = -9.47 + 0.181 students
- d. none of these.
- e. sales = 12 + student
- f. sales = 60 + 5 students

Question 2 (4 points)

Which option gives the percentage of variation in sales explained by student population size?

- a. none of these.
- b. 94.6
- c. 95.1
- d. 97.0
- e. 90.3
- f. 87.8

Question 3 (2 points)

Using the line found in question 1 which option gives the predicted annual sales of a café with a student population of 11,000?

- a. none of these.
- b. £7485
- c. £115,000
- d. £55, 060

e. £116,000

Question 4 (4 points)

Which option gives the 95% confidence interval for the average annual sales with a student population of 11,000 students?

a. none of these

b. £40,359 to £69,760

c. £81310 to £148,690

d. £13,546 to £14,240

e. £104,150 to £125,850

Question 5 (2 points)

If the student population were to increase by 1000, which option gives the resulting average increase in annual sales?

a. 6

b. 5000

c. 50

d. 6000

e. none of these.

f. 500

Question 6 (4 points)

Which option gives a 95% confidence interval for the increase in average annual sales found in question 5?

a. 3.50 to 6.48

b. none of these.

c. 3.71 to 6.29

d. 3.11 to 6.82

e. 3.86 to 6.14

Question 7 (3 points)

The following information should be used in conjunction with questions 7 to 13.

A soup company markets a number of varieties of 'homemade' soup. The standard-size soup can holds a maximum of 11 ounces, while the label on each can advertises contents of 10.5 ounces. The extra 0.5 ounce is to allow for the possibility of the automatic filling machine placing more soup than the company actually wants in a can. Past experience shows that the number of ounces placed in a can is approximately normally distributed, with a mean of 10.65 ounces and a standard deviation of 0.1 ounce.

Which option gives the proportion of cans which overflow?

a. none of these.

b. 0.0002

- c. 0.9980
- d. 0.9520
- e. 0.0020
- f. 0.0037

Question 8 (3 points)

Which option gives the probability that a can has less soup than advertised?

- a. 0.0548
- b. 0.0062
- c. 0.0655
- d. 0.0808
- e. none of these.
- f. 0.0668

Question 9 (5 points)

If the chance of a can having less soup than advertised is to be no more than 1%, which option gives the minimum mean fill that the automatic filling machine can be set?

- a. none of these.
- b. 10.69
- c. 10.87
- d. 10.76
- e. 10.73
- f. 10.67

Question 10 (1 point)

The following information should be used with questions 10 to 13.

A random sample of 10 cans is selected from the filling process and their contents examined. The following are the individual contents in units of ounces.

10.54 10.62 10.47 10.33 10.60 10.71 10.83 10.34 10.77 10.48

Which option gives the mean fill of these cans?

- a. 10.69
- b. 10.76
- c. none of these.
- d. 10.87
- e. 10.57
- f. 10.73

Question 11 (1 point)

Which option gives the standard deviation of the fill of the 10 cans?

- a. 0.17
- b. none of these.
- c. 0.03
- d. 0.16
- e. 0.15
- f. 0.18

Question 12 (4 points)

Which option gives a 95% confidence interval for the mean fill of the cans?

- a. 10.47 to 10.67
- b. 10.44 to 10.70
- c. 10.45 to 10.69
- d. 10.46 to 10.67
- e. none of these.

Question 13 (3 points)

If a test of the hypothesis that the mean fill is not 10.65 ounces is made using the answer to question 12, which option gives the most appropriate answer?

- a. As the confidence interval does not contain the origin there is no evidence to reject the null hypothesis.
- b. As the sample mean is not equal to 10.65 the population mean is different and so we reject Ho.
- c. none of these.
- d. As 10.65 lies inside the confidence interval there is evidence to think that the mean fill has changed.
- e. As 10.65 lies inside the confidence interval there is no evidence to think that the mean fill has changed.

Question 14 (1 point)

The following information should be used in conjunction with questions 14 to 22.

Two assembly lines are designed on the assumption that there should be no difference between mean assembly times of a particular household appliance. Independent tests for the two assembly operations show the following results.

Line 1	Line 2
$n_1=10$	$n_2=15$
$\bar{x}_1=14.8$ minutes	$\bar{x}_2=14.0$ minutes
$s_1=0.8$ minutes	$s_2=0.6$ minutes

The hypothesis that there is no difference between the mean assembly times for the two processes is going to be tested at the 0.02 significance level.

Which option gives the appropriate test procedure?

- a. two-sample z-test.
- b. none of these.
- c. Chi-square test.
- d. paired t-test.
- e. single sample t.
- f. two-sample t-test

Question 15 (1 point)

Which option gives the value of the number of degrees of freedom for this test?

- a. 23
- b. 24
- c. 25
- d. 22
- e. 5
- f. none of these.

Question 16 (5 points)

Which option gives the observed value of the test statistic?

- a. 2.86
- b. 1.99
- c. 1.65
- d. 2.57
- e. none fo these.
- f. 2.92

Question 17 (3 points)

Which option gives the most appropriate conclusion for this test?

- a. As the observed value of the test statistic is less than the critical value, there is evidence that the two assembly lines work at different average speeds.
- b. As the observed value of the test statistic is greater than the critical value, there is evidence that the two assembly lines work at different average speeds.
- c. As the observed value of the test statistic is less than the critical value, there is no evidence that the two assembly lines work at different average speeds.
- d. As the observed value of the test statisic is greater than the critical value, there is no evidence that the two assembly lines work at different average speeds.
- e. none of these.

Question 18 (2 points)

The following information should be used in conjunction with questions 18 to 22.

The leader of the workers union argues that there is a difference between the two lines, and that operators on Line 1 should be compensated for the extra time it takes to assemble the household appliance. Management argue that there is no difference in the lines and any perceived difference is due to the fact that some operators are faster than others. To see if there is any truth in what the union and management are saying, the management decide to set up an experiment. A group of 10 workers are randomly selected (five from each line) and placed at random on either Line 1 or Line 2. After performing the operations each of the 10 workers are moved across from the line they were placed on initially to the other one. The two sets of timings (in minutes) for the 10 workers are as follows:

Worker	1	2	3	4	5	6	7	8	9	10
Line 1	13.6	10.1	11.0	14.9	12.7	10.8	15.1	9.9	9.9	15.3
Line 2	13.5	9.9	10.9	15.0	11.5	10.8	14.7	10.2	9.6	14.8

Which option gives the observed value of the most appropriate test statistic if the hypothesis that there is no difference between the two assembly lines is being considered.

- a. two-sample z-test
- b. none of these.
- c. single sample t
- d. paired t-test
- e. Chi-square test
- f. two-sample t-test

Question 19 (1 point)

Which option gives the average difference between the two sets of timings?

- a. 0.19
- b. -0.21
- c. 0.24
- d. -0.19
- e. -0.24
- f. 0.21
- g. none of these.

Question 20 (1 point)

Which option gives the standard deviation of the differences between the two timings?

- a. 0.1524
- b. 0.4115
- c. 0.1693
- d. 0.3904
- e. none of these.

Question 21 (3 points)

Which option gives a 95% confidence interval for the mean difference between the two sets of timings?

- a. -0.637 to 0.064
- b. -0.054 to 0.534
- c. 0.054 to 0.534
- d. -0.534 to -0.054
- e. none of these.
- f. -0.534 to 0.054
- g. -0.064 to 0.637

Question 22 (3 points)

Whose argument is supported by your results, management or union? Which option gives the most appropriate answer?

- a. none of these.
- b. As the confidence interval does not contain the origin, there is eveidence of a difference and so the union's argument is supported.
- c. As the confidence interval contains the origin, there is no evidence of a difference and so the management's arguement is supported.
- d. As the confidence interval contains the origin, there is evidence of a difference, and so the union's argument is supported.
- e. As the confidence interval does not contain the origin, there is no evidence of a difference, and so the management's argument is supported.

Question 23 (2 points)

From the option list provided select two examples of nominal data?

weight, names, position in race, IQ, temperature, height, colours, preference score

Answer:

1.

2.

Question 24 (2 points)

From the option list provided select two examples of ordinal data.

weight, names, position in race, IQ, temperature, height, colours, preference score

Answer:

1.

2.

Question 25 (2 points)

From the option list provided select two examples of interval scale data.

weight, names, position in race, IQ, temperature, height, colours, preference scores

Answer:

1.

2.

Question 26 (2 points)

From the option list provided select two examples of ratio scale data.

weight, names, position in race, IQ, temperature, height, colours, preference score

Answer:

1.

2.

Question 27 (2 points)

The following information should be used with questions 27 to 32.

An employment agency provided the following data as an examle of selection among 40 male and 40 female applicants for 12 positions.

Applicant	Selected	Not Selected	Total
Male	7	33	40
Female	5	35	40

The chi-square test of independence was suggested as a way of determining if the decision to hire seven males and five females should be interpreted as having a selection bias in favour of males. Let the population of male and female applicants selectd be denoted π_M and π_F respectively.

Which option gives the null and alternative hypothesis for this chi-square test?

a. $H_0 : \pi_M \neq \pi_F \quad v \quad H_1 : \pi_M > \pi_F$

b. $H_0 : \pi_M = \pi_F \quad v \quad H_1 : \pi_M \neq \pi_F$

c. $H_0 : \pi_M > \pi_F \quad v \quad H_1 : \pi_M = \pi_F$

d. $H_0 : \pi_M = \pi_F \quad v \quad H_1 : \pi_M > \pi_F$

e. none of these.

f. $H_0 : \pi_M < \pi_F \quad v \quad H_1 : \pi_M > \pi_F$

Question 28 (1 point)

How many degrees of freedom are there for this chi-square test?

a. 1

b. 3

c. none of these.

d. 2

e. 5

f. 4

Question 29 (1 point)

Which option gives the critical chi-square value if we use a 10% significance value?

- a. 7.88
- b. 2.71
- c. 0.02
- d. 3.84
- e. none of these.
- f. 4.61

Question 30 (4 points)

Which option gives the observed value of the chi-square statistic for the above test?

- a. 0.467
- b. 0.392
- c. 0.201
- d. none of these.
- e. 3.920

Question 31 (2 points)

Which option gives the most appropriate conclusions of your hypothesis test?

- a. There is no evidence of discrimination.
- b. There is evidence of discrimination.
- c. Proportion of females selected greater than the number of males.
- d. none of these.
- e. Row and column classifications are dependent.
- f. Proportion of males selected greater than the number of females.

Question 32 (6 points)

Which option gives a 95% confidence interval for the difference between the proportion of males and females selected?

- a. -0.160 to 0.260
- b. -0.260 to 0.160
- c. -0.106 to 0.206
- d. none of these.
- e. -0.260 to -0.160
- f. -0.206 to 0.106