Sample Paper

BUSINESS STATISTICS

Level 3

Subject Code: 3009 SAMPLE

Time allowed: **3 hours**

INSTRUCTIONS FOR CANDIDATES

- Answer **5** questions.
- All questions carry equal marks.
- There are statistical tables and a list of formulae at the end of the paper.
- Graph paper is provided within the **answer book**.
- Write your answers in blue or black ink/ballpoint. Pencil may be used only for graphs, charts, diagrams, etc.
- All answers must be correctly numbered but need not be in numerical order.
- Workings must be shown.
- Your work should be accurate and neat.
- You may use a calculator, provided the calculator gives no printout, has no word display facilities, is silent and cordless. The provision of batteries and their condition is your responsibility.

QUESTION 1

(a) Describe the main characteristics of the normal distribution. (4 marks)

The weights of containers used by a company have a normal distribution with mean 50 grams and standard deviation 5 grams.

(b) The containers are supplied in batches of 100. Calculate the probability that a randomly selected batch will have a total weight of less than 5120 grams

(8 marks)

The weight of the contents put in the containers is also normally distributed with mean 250 grams and standard deviation 6 grams. The weight of the container and contents are independent of each other.

(c) Calculate the probability that the weight of a randomly selected **filled** container will be more than 318 grams.

(8 marks)

(Total 20 marks)

QUESTION 2

(a) Explain what is meant by the sampling distribution of the sample proportion. (4 marks)

The manufacturer of a long life battery is testing a new additive to the process. Before the new additive was used a random sample of 12 items was selected and tested to destruction. Following the change in additive a second random sample, of 10 items, was tested to destruction. The results were as follows:

	Life in hours											
Before new additive	25	28	36	29	27	26	31	40	25	36	28	22
After new additive	29	36	28	31	25	26	32	27	38	25		

(b) Test whether the new additive has increased the life of the batteries. (12 marks)

(c) Explain what is meant by a *type 2 error*. State whether a *type 2 error* may have been committed in your conclusions to part (b). Explain your answer.

(4 marks)

(Total 20 marks)

QUESTION 3

(a) Explain the factors which should be taken into consideration when constructing a general price index such as the Retail Price Index or Consumer Price Index.

(10 marks)

(b) Explain **two** sampling methods you might use to choose the representative sample of items to include in the consumer price index. For each sampling method give an advantage and a disadvantage for using this method compared with other sampling methods.

(10 marks)

(Total 20 marks)

QUESTION 4

A company records its electricity consumption against the temperature for 10 days.

Electricity Consumption (Kilowatt hours 000)	**Temperature ^{0}C**
37	12
42	19
46	14
35	17
49	15
58	10
29	22
63	11
76	8
45	14

(a) Calculate and state the regression equation for electricity consumption based on temperature.

(10 marks)

The calculated value of the Coefficient of Determination is 68.69%.

(b) (i) Explain what the Coefficient of Determination measures and interpret the above figure.

(4 marks)

The critical value of the correlation coefficient at the 0.05 significance level for a sample of 10 items is 0.4733.

(ii) Use the result for the Coefficient of Determination to calculate the correlation coefficient and test whether the correlation coefficient differs significantly from zero.

(6 marks)

(Total 20 marks)

QUESTION 5

(a) Explain what is meant by and give a business example of:

(i) mutually exclusive events

(ii) conditional probability

(4 marks)

A company decides to advertise its service on 2 or 3 television channels which have national coverage. Viewer statistics show that 48% of the population view channel A, 30% view channel B and 25% view channel C. 12% view channels A and B, 4% view channels A and C and 2% view channels B and C. 1% view all three channels.

(b) If an advertisement was placed on all three channels, using a Venn diagram or otherwise, find the percentage of the population that would be able to:

(i) have the opportunity to see the advertisement

(ii) see the advert on only one channel

(7 marks)

(c) If the company only had sufficient funds available to place the advert on 2 channels, assuming the intention is to maximise the audience, which 2 should it choose and what percentage of the population would be able to see the advertisement.

(3 marks)

All three channels show a free phone number where viewers can contact the firm for more information. The number called identifies the channel that the viewer was watching and records identify whether the caller bought the service. This analysis shows channel A calls represented 40% of respondents and 20% of these bought the service. For channel B the percentages were 35% and 30% respectively and for channel C 25% and 25% respectively.

(d) (i) Of those that called what percentage bought the service? (3 marks)

(ii) If a person bought the service what is the probability it was a person calling the channel A number?

(3 marks)

(Total 20 marks)

QUESTION 6

(a) Explain when you would use the Chi-squared test of significance. (4 marks)

A random sample of consumers was carried out to investigate if there was any relationship between the type of television bought and the social class of the purchaser. The results were as follows:

	Type of Television		
Social Class	CRT	LCD	Plasma
A	60	45	25
B	65	70	45
C	155	120	130
D & E	120	85	80

(b) Test whether there is any association between the type of television bought and the social class of the purchaser.

(12 marks)

(c) By combining the data for the number of people who bought plasma televisions estimate the 99% confidence interval for the population proportion for plasma televisions bought.

(4 marks)

(Total 20 marks)

QUESTION 7

(a) Explain the difference between a 1 tailed and a 2 tailed test of significance. (4 marks)

A company has two forms of packaging, A and B for its product. It conducted two random samples of the customers who had returned the product as damaged identified by type of packaging.

	Packaging A	**Packaging B**
Items returned	27	22
Sample size	1000	750

(b) Test whether the proportion returned differs due to the type of packaging used.

(12 marks)

(c) If a 1 tailed test had been carried out, would the conclusion you have reached in part (b) remain the same? Explain your answer.

(4 marks)

(Total 20 marks)

QUESTION 8

(a) Explain the circumstance in which *a t test* is used in preference to *a normal distribution based z test,* when testing for the mean of one sample.

(4 marks)

The mean weight of a random sample of 12 bags of sugar is 0.988 kg with standard deviation 0.015 kg.

(b) Test, at the 0.01 significance level, whether the mean weight of the sample is significantly less than the specified mean weight of the bags of sugar given as 1.0 kg.

(8 marks)

(c) If the company introduces a new production technique which reduces the standard deviation to 0.09 kg, and assuming the sample size remains 12, what weight should the company set as the target weight so that 99% of the time the sample mean weight of the bags exceeds 1 kg?

(4 marks)

(d) Explain the impact of a larger sample size upon the value of the standard error

(4 marks)

(Total 20 marks)

FORMULAE

Mean $\bar{x} = \frac{\sum fx}{\sum f}$

Standard deviation

$$s = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

Pearson measure of skewness

$$\frac{3(\bar{x} - \text{Median})}{s}$$

Coefficient of variation

$$\frac{s}{\bar{x}} \times 100$$

Product moment correlation coefficient

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left(n\sum x^2 - (\sum x)^2\right)\left(n\sum y^2 - (\sum y)^2\right)}}$$

Spearman's rank correlation coefficient

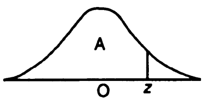
$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Least squares regression line $\hat{y} = a + bx$

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

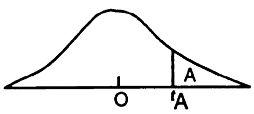
$$a = \frac{\sum y}{n} - \frac{b\sum x}{n}$$

TABLE 1 – NORMAL DISTRIBUTION
A is the area to the left of the given value of *z*



z	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
A	.500	.540	.580	.618	.655	.692	.726	.758	.788	.816	
z	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	
A	.841	.864	.885	.903	.919	.933	.945	.955	.964	.971	
z	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
A	.977	.982	.986	.989	.992	.994	.995	.996	.997	.998	.999

TABLE 2 – t DISTRIBUTION
t_A is the value of the t statistic with ν degrees of freedom with area A to the right of it



ν	1	2	3	4	5	6	7	8
$t_{0.05}$	6.31	2.92	2.35	2.13	2.02	1.94	1.90	1.86
$t_{0.025}$	12.71	4.30	3.18	2.78	2.57	2.45	2.37	2.31
$t_{0.01}$	31.82	6.97	4.54	3.75	3.37	3.14	3.00	2.90
$t_{0.005}$	63.66	9.93	5.84	4.60	4.03	3.71	3.50	3.36
ν	9	10	11	12	13	14	15	16
$t_{0.05}$	1.83	1.81	1.80	1.78	1.77	1.76	1.75	1.75
$t_{0.025}$	2.26	2.23	2.20	2.18	2.16	2.15	2.13	2.12
$t_{0.01}$	2.82	2.76	2.72	2.68	2.65	2.62	2.60	2.58
$t_{0.005}$	3.25	3.17	3.11	3.05	3.01	2.98	2.95	2.92
ν	17	18	19	20	21	22	23	24
$t_{0.05}$	1.74	1.73	1.73	1.73	1.72	1.72	1.71	1.71
$t_{0.025}$	2.11	2.10	2.09	2.09	2.08	2.07	2.07	2.06
$t_{0.01}$	2.57	2.55	2.54	2.53	2.52	2.51	2.50	2.49
$t_{0.005}$	2.90	2.88	2.86	2.85	2.83	2.82	2.81	2.80

One sample *z* test

Mean

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Proportion

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$$

Two sample *z* test

Mean

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Proportion

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

	Price	Quantity
Laspeyres index	$\frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$	$\frac{\sum p_0 q_1}{\sum p_0 q_0} \times 100$
Paasche index	$\frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$	$\frac{\sum p_1 q_1}{\sum p_1 q_0} \times 100$

Weighted index $\frac{\sum WI}{\sum W}$

One sample t test

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad \text{where} \quad s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

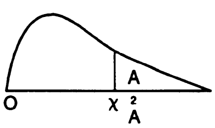
Independent samples t test

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n} + \frac{1}{m}}} \quad \text{where} \quad s = \sqrt{\frac{\sum(x - \bar{x})^2 + \sum(y - \bar{y})^2}{n + m - 2}}$$

Chi-square test $\chi^2 = \sum \frac{(O - E)^2}{E}$

Test for $\rho = 0$ $\quad t = \frac{r\sqrt{n - 2}}{\sqrt{1 - r^2}}$

TABLE 3 – CHI SQUARED DISTRIBUTION
χ^2_A is the value of the χ^2 statistic with ν degrees of freedom with area A to the right of it



ν	1	2	3	4	5	6
$\chi^2_{0.05}$	3.84	5.99	7.81	9.49	11.07	12.59
$\chi^2_{0.01}$	6.63	9.21	11.34	13.28	15.09	16.81
ν	7	8	9	10	11	12
$\chi^2_{0.05}$	14.07	15.51	16.92	18.31	19.68	21.03
$\chi^2_{0.01}$	18.48	20.09	21.67	23.21	24.73	26.22