

OLLSCOIL NA hÉIREANN, GAILLIMH
NATIONAL UNIVERSITY OF IRELAND, GALWAY

SEMESTER ONE EXAMINATIONS, 2003 – 2004

SECOND COMMERCE (EVENING) EXAMINATION

APPLIED QUANTITATIVE METHODS [MA210]

Dr. Trevor C. Bailey
Paul Wilson, M.A., H.D.E.

Time allowed: **Three** Hours.

Answer any *four* questions.

All questions, but not necessarily parts therein, carry equal marks.

Standard Normal Distribution Tables and a table of the critical values of the Mann-Whitney
U-test are attached.

Tables of the critical values of the t and χ^2 distributions are available in the supplied
Mathematical ("log") tables.

A Formula sheet is attached.

1. (a) i. As part of a study of the distances commuted to work by its employees, the following data were collected by a company.

28	17	25	20	09	36	34	51	04	17
12	06	14	79	23	07	31	39	04	46
70	05	27	89	17	09	46	77	04	20
31	02	17	42	15	12	38	72	53	100
28	94	31	49	61	59	26	25	04	56
41	22	32	02	55	11	18	45	76	51

- A. Are the above data continuous, ordinal or nominal?
- B. Illustrate the data with a histogram with intervals 0–20, 20–40, 40–60 and 60–100.
- (b) i. Give an example of a set of data that does not have a mode.
- ii. Give an example of a set of data whose mode is equal to zero.
- (c) Two groups, *A* and *B*, of students are given a multiple choice exam. 100 questions are asked, and there are four alternative answers to each question. The students in group *A* have studied diligently for the test, whereas those in group *B* have not studied at all. The results of the test are as follows.

	<i>A</i>	<i>B</i>
<i>Mean</i>	80	20
<i>Standard Deviation</i>	30	5

- i. Find the Coefficient of Variation of each of the two sets of marks.
- ii. With reference to your answer to the above question, briefly explain why the Coefficient of Variation is a useful statistic to employ if one wishes to compare the variability of two sets of data whose means differ by a large amount.
- (d) Find the mean, median, mode and population standard deviation of
- $$\{2, 6, 4, 4, 0, 1, 4, 8, 9\}$$
- (e) It is wished to survey the opinions of nurses employed in Irish nursing homes about the recent legislation on smoking. A list of all such nursing homes, and the nurses working in them is available. Explain how you would obtain a sample of 200 such nurses using (i) Simple Random Sampling, (ii) Cluster Sampling, (iii) Sequential sampling.

Question Two is on The Next Page

2. (a) The probability distribution for damage claims paid by a car insurance company is as follows:

Payment (Euro)	0	1000	2000	3000	4000	5000
Probability	0.80	0.14	0.03	0.01	0.01	0.01

What is the mean value of a damage claim?

- (b) The number of night-calls that a given doctor receives is Poisson distributed with a mean of $\lambda = 1.3$ per night. What is the probability that
- he receives exactly two call-outs in a given night,
 - he is not called out in a given night,
 - he is not called out for any given two nights in a row,
 - he is not called out two nights in a row, but is called out exactly three times on the third night?
- (c) A coin is weighted so that when tossed it will come up "heads" four out of five times.
- What is the sample space of this experiment.
 - If this coin is tossed ten times, what is the probability that it will come up heads:
 - exactly twice,
 - at least once?
- (d) It is known that the amount of petrol sold daily at a given filling station is normally distributed with a mean of $\mu = 5,000$ litres, and a standard deviation of 600 litres. What is the probability that on any given day the amount of petrol sold is:
- between 5,000 and 6,200 litres,
 - between 5,800 and 6,900 litres,
 - less than 4,340 litres?

Question Three is on The Next Page

3. (a) A factory manager believes that the mean weight of "one kilogram gram bags of sugar" being produced is in fact greater than 1,000 grams. He takes a random sample of 100 such bags, and finds that the bags in the sample have mean weight of $\mu = 1,001$ grammes with a standard deviation of 4 grams.
- State a suitable null (H_0) and alternative (H_1) hypothesis.
 - May the null hypothesis be rejected at a significance level of $\alpha = 0.05$?
 - If a *type one error* was made above, what, in terms of this specific case, would this mean?
 - If the factory manager had decided to reject the null hypothesis if the mean weight of the bags in the sample had been greater than 1,000.5 grammes, what would have been the probability of a type-one error?
- (b) i. State two conditions that should hold for an independent samples t-test to be valid.
- ii. A researcher believes that the ratio of current assets to current liabilities is greater for "healthy" firms than for failed firms. Random samples of such firms are taken, and the following summary statistics are obtained.

Firms	n	\bar{x}	s
Healthy	8	1.7256	0.6393
Failed	12	0.8236	0.4811

- Perform an independent samples t-test to test the hypothesis that the mean asset/liability ratio is greater for healthy firms than for failed firms. Interpret your findings.
 - Construct a 95% confidence interval for the ratio of current assets to current liabilities for healthy firms.
- (c) A physical therapist wishes to learn whether an exercise programme increases flexibility. She measures the flexibility (in centimetres) of ten randomly selected subjects, both before and after an intensive eight week training program, and obtains the following data:

Before	42	51	40	50	49	30	38	60	36	46
After	46	54	44	56	49	50	48	62	34	49

Use the Mann-Whitney U-test to test the claim that the median flexibility after the programme differs from the median flexibility before.

Question Four is on The Next Page

4. (a) It is wished to investigate if the ages of mature students are uniformly distributed over the intervals 20–30, 30–40, 40–50 and 50–60.

A sample of 100 mature students is taken. The age profile of the students as follows:

20–30	30–40	40–50	50–60
30	30	22	18

May we reject H_0 : Distribution of all such students is in the ratio 1:1:1:1 at $\alpha = 0.05$?

- (b) i. A. What is meant by saying that two variables are *independent*?
 B. The event “Tommy goes swimming in the sea” and “Jean goes swimming in the sea” are independent. If the probability that Tommy goes swimming in the sea is 0.4 and the probability that Jean goes swimming in the sea is 0.5, what is the probability that neither of them go swimming in the sea?
 ii. Blood is classified as A, B, AB, or O. In addition, blood can be classified as Rh^+ or Rh^- . In a survey of 500 randomly selected individuals, a phlebotomist obtained the results shown below.

	A	B	AB	O
Rh^+	176	28	22	198
Rh^-	30	12	4	30

Use the χ^2 test for independence of two variables to test whether blood type and Rh-level are independent at the $\alpha = 0.05$ level of significance.

Interpret your answer.

- iii. What is meant by an “expected count”? (N.B. You are to explain what the term means, *not* how it is calculated). What condition concerning expected counts should be satisfied for the results of a χ^2 -test to be valid? Briefly explain how, when this condition is not met, it may be possible to modify the null hypothesis so that the condition is met.

Question Five is on The Next Page

5. (a) i. What is meant by the term *least squares regression line*?
- ii. Explain what is meant by saying that two variables are *weakly, positively, linearly* correlated.
- iii. Does the fact that two variables are strongly correlated necessarily imply that one is causing the other? Illustrate your answer with an example.
- (b) Sketch scatter diagrams that illustrate two variables that are:
- i. strongly positively correlated,
- ii. weakly negatively correlated.
- (c) It is wished to investigate if the share price of "Dalton and Company Ltd." may be used to predict the share price of "Shirley and Sons Ltd." to a reasonable degree of accuracy. The share prices for both companies over a 12 day period are recorded. These share prices are given in the table below.

<i>Dalton</i>	<i>Shirley</i>
200	300
210	305
210	310
195	315
170	270
160	260
160	265
175	295
180	300
195	310

- i. Using a calculator, or otherwise, estimate the correlation coefficient of the two variables.
- ii. Using a calculator, or otherwise, find the equation of the least squares regression line.
- iii. What is the predicted mean value of the Shirley and Sons Ltd. shares when Dalton and Company Ltd. shares are trading at 180 cent?
- iv. In the table above the value of the Dalton and Company shares when the Shirley and Sons shares are trading at 180 cent is 300 cent. Why does this value differ from the answer you obtained in the previous part of this question?
- v. Show briefly how it is possible for two variables that are (linearly) uncorrelated in the population to appear correlated due to the samples selected.
- vi. It may be shown that the significance of the (unstandardised) coefficient of "Dalton and Company Ltd." is zero. Explain what is meant by this.

End of Question Five