

475.101/102/107/108
STATISTICS

Assignment 1, Semester 1, 2000
Due: 4pm Thursday, 16th March

Instructions for handing in: PLEASE

1. Use (standard) **A4** sized paper.
2. Number each page in the top centre and **print** your name legibly at the **top righthand corner** of each page with the surname or family name underlined.
3. Attach a *Department of Statistics Assignment Cover Sheet* to the front of the assignment. Staple or clip all pages together in the extreme top left hand corner.
4. Fold the paper lengthwise so that the printed side of the cover sheet faces out.
5. **Print** your name, paper number and assignment number on the outside of the cover sheet. **Print** your ID number on the inside of the cover sheet.
6. **Hand the assignment in** to the appropriate hand-in box in the basement of the Maths/Physics building for the city campus students or in the Tamaki Resource Centre for Tamaki campus students.

Notes:

- Statistics is about summarizing, analyzing and communicating information. Communication is an important part of statistics. For this reason you will be expected to write answers which clearly communicate your thoughts. **The mark you receive will be based on your written English** as well as your statistical/technical work.
- Assignment 1 will be marked out of 60 marks, 54 marks for question 1-5 as shown below and 6 marks for communication, style and presentation. Your final mark will be converted to a mark out of 10 which will be recorded towards your coursework.
- **To make your marked assignment easier to find** when they are returned, you could draw a pattern along the edges of the coversheet using coloured pens or put some sort of small sticker on the cover sheet.

Question 1. [13 marks]

- (a) Read §4.6 Course Assessment on page 11 of the Course Information section of the Course Resource Book. The assignment, test and exam marks for four students from this course are shown in the table below. A final mark of at least 50% is ONE of the conditions necessary to pass the course.

	Assignment Total (/15)	Term Test (/20)	Final Exam (/65)
Student 1	4	11	49
Student 2	8	9	35
Student 3	14	10	28
Student 4	12	Did not sit/no excuse	37

- (i) Which students fail to qualify for plussage?
- (ii) Student 2 and Student 3 scored a total of 52% in their coursework and exam (assignments + test + exam). Student 2 passed the paper but Student 3 failed. Why?
- (iii) Which student would have got a final mark of 75% if they had qualified for plussage, but instead got a final mark of 64%?

- (b) Briefly describe how you could use tables of random numbers to generate a random sample of size 50 from a list of 8,216 people.
- (c) The following scheme was proposed to try and find out more information about the customers of a chain of stores:

A questionnaire was designed with questions on employment, socio-economic level, family size and use of the stores. Store personnel would interview every 5th customer between 10am and noon and between 2pm and 4pm (when the stores were not too busy and the staff could be spared).

Discuss **four** potential sources of errors with this sampling scheme. [Note: Simply naming the sources of error is not sufficient to get any marks. You must discuss how or why each source of error is potentially present.]

- (d) A survey of New Zealand high school principals taken after widespread changes to the school system revealed that one in five were under medication for stress and almost half had visited doctors because of the pressure they were under. These figures came from 250 questionnaires returned from 2500 sent out. How reliable were the results? Briefly justify your answer.
- (e) Use your calculator to calculate the mean and standard deviation for the following set of data:
1.3, 3.1, 0.9, -1.0, 6.9, 8.1, 2.6, 3.4, -2.2

Question 2. [12 marks]

Consider the following four studies:

Study 1: A researcher at Melbourne university was interested in whether pain tolerance levels were related to hair colour. 80 people were selected from a group of volunteers, 20 with light blonde hair, 20 with dark blond hair, 20 brunettes and 20 redheads. The subjects underwent a series of tests and pain tolerance scores (on a scale of 0 to 100) were assessed.

Study 2: A technician is interested in the effects of using different baking temperatures on the impact strength of particle board. 20 boards are randomly allocated to 20 different baking temperatures. After the boards are baked, they are sent to a laboratory where the impact strengths are measured.

Study 3: The manager of an automotive workshop is interested in whether using a new diagnostic machine will speed up the regular servicing of cars. There are two mechanics working on regular services, one with 8 years experience while the other mechanic had only 2 years experience. One mechanic was told to use the diagnostic machine on the next 10 cars she serviced, but not use it on the following 10 cars. The other mechanic was told not to use the diagnostic machine on the next 10 cars he serviced, but to use it on the following 10 cars. Each mechanic recorded the time it took to complete the services for each of these jobs.

Study 4: A sociologist is interested in comparing the exam results for male and female students doing 10 different subjects at bursary for a particular year. The New Zealand Qualifications Authority was contacted and provided numbers of male and female students who sat the exam and the numbers of male and female students who got each of the grades A, B, C and D.

- (a) For each study, describe what "treatment" is being compared and what "response" is being measured to compare the treatments.
- (b) Which of the studies would be described as experiments and which would be described as observational studies?
- (c) For the studies that are observational, could an experiment have been carried out instead? If not, briefly explain why not.
- (d) For the studies that are experiments, briefly discuss what forms of blinding would be possible to be used.
- (e) In which of the studies has blocking been used? Briefly describe what was blocked and why it was blocked.

Question 3. [10 marks]

The following data represent the daily number of parking tickets given out in a downtown region of a small city over a period of 29 working days.

42, 47, 46, 35, 43, 39, 38, 40, 50, 37, 68, 37, 47, 44, 49
41, 34, 38, 41, 36, 42, 38, 38, 58, 34, 32, 42, 49, 52.

- (a) By hand, construct a stem-and-leaf plot of the data using an appropriate scale.
- (b) Calculate the five-number summary for the data.
- (c) By hand, draw a box plot for this set of data. Show your working.
- (d) Using your plots, in plain English, **briefly** comment on the data in terms of the original story.

Question 4. [15 marks]

Backpacks are commonly seen in many places, especially university campuses, schools, shopping malls and airplanes. In 1997, a US consumer magazine conducted a study on backpacks. Some of the data from the study is reproduced in the table below. The price of the backpacks is to the nearest US dollar, volume is measured in cubic inches and number of books is the number of 5-inch by 7-inch by half-inch books that the backpack easily holds.

Backpack Code	PM	MM	BL	DS	CM	MR	MP	RP	JC	CB	EW	EA
Price(\$)	38	40	50	48	35	35	45	50	35	95	40	40
Volume	1586	1810	2200	1874	1810	1950	1670	1500	1635	1102	1700	1500
Number of Books	47	44	48	50	50	49	49	49	40	73	38	49

- (a) Calculate the average volume used per book (i.e. the volume divided by the number of books) for each backpack.
- (b) Following the guidelines for tabular presentation (Section 2.2.2 of the text), redraw the above table with the information from (a) added to it.
Note:
 - Assume that we are primarily interested in the price of the backpacks.
 - Add averages to the table as appropriate.
 - You may use *Excel* to draw the table.
- (c) Use either *Excel* or *Minitab* to create three separate scatter plots – volume vs price, number of books vs price and number of books vs volume. Print out the three plots.
Note:
 - Each axis of the plots should be appropriately labelled.
 - If possible, try to print the three plots on one sheet of paper.
 - To use *Excel* to draw scatter plots is similar to using *Excel* to draw dot plots (see page 7 of *Excel* notes in coursebook).
- (d) Write a short report (from 150 to 250 words). Your report should include a discussion on the data (comments on trends, problems or interesting features in the data and what other data could be collected to help consumers compare backpacks) and a recommendation for consumers (which backpack(s) do you recommend buying and why).

Question 5. [4 marks]

For **each** of the four studies described in question 2:

- (i) State whether the resulting (treatment and response) variables are both quantitative, both qualitative or one quantitative and one qualitative.
- (ii) Describe an appropriate tool for looking at the relevant relationship.