STATISTICS: 475.101/102/107/108

Assignment 3, Semester 2, 2000
Due: 4pm Thursday, 14th September

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 right hand corner** of each page with the surname or family name underlined.
3. Attach a white *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 length-wise 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.

Notes:

- Statistics is about summarising, analysing 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 3 will be marked out of 60 marks, 54 marks for questions 1-4 as shown below and 6 marks for communication, style and presentation. Please refer to Section B, Questions 10 to 13, in your Course resource book for examples of how to set out your assignment answers. Your final mark will be converted to a mark out of 10 which will be recorded towards your course work.

475.10x Statistics: Term Test
6.30-7.30pm Tuesday 12th September

The test will be comprised of **26** multiple choice questions covering chapters 1 - 7.

Test rooms for City Campus:

Paper	Surname	Room	Paper	Surname	Room
475.101/102/108	A - C	HSB1	475.101/102/108	Lu - Ngo	LibB10
475.101/102/108	D - Gi	HSB2	475.101/102/108	Ngu - She	LibB15
475.101/102/108	Gj - Ja	ULT	475.101/102/108	Shi - Z	LibB28
475.101/102/108	Je - Lo	LLT			

Test room for Tamaki Campus: All 475.107 students are in room 722.201

You should ensure that you can find your test room well before you have to.

If you have a problem with sitting the test at this time then you must see either David Smith or Christine Miller, in Room 226 of the Maths/Physics building, or Ross Parsonage at the Tamaki campus, before the 6th September.

Question 1. [12 marks]

(a) The following table of probabilities was obtained from *Excel*.

$X \sim$ Normal(μ=18, σ=3)

x	pr($X \le x$)	x	pr($X \le x$)	x	pr($X \le x$)
10	0.0038	15	0.1587	20	0.7475
11	0.0098	16	0.2525	21	0.8413
12	0.0228	17	0.3694	22	0.9088
13	0.0478	18	0.5000	23	0.9522
14	0.0912	19	0.6306	24	0.9772

Use the above table to find the following when $X \sim$ Normal($\mu = 18$, $\sigma = 3$):

(i) pr($X \le 17$) **(ii)** pr($X < 17$)

(iii) pr($X > 21$) **(iv)** pr($12 \le X \le 16$)

(b) The following table of probabilities was obtained from MINITAB:

```
Normal with mean = 23.6000 and standard deviation = 4.25000

 P( X <= x)          x
     0.1000       18.1534
     0.2500       20.7334
     0.5000       23.6000
     0.7500       26.4666
     0.9000       29.0466
```

The amount of time required for a shop owner to organise the day's banking is found to be Normally distributed with a mean of 23.6 minutes and a standard deviation of 4.25 minutes.

(i) What is the minimum amount of time that should be set aside so that the shop owner can organise the day's banking without interrupting other work on 90% of days?

(ii) What is the interquartile range for the amount of time taken to organise the day's banking?

(c) Use either *Excel*, MINITAB or a graphics calculator to solve the following problems where $X \sim$ Normal($\mu = 8.5$, $\sigma = 1.47$):

(i) What is the probability that X is greater than 6?

(ii) What is the probability that X is between 7.7 and 10.6?

(iii) What value of x gives pr($X \le x$) = 0.7?

(d) X has a mean of -2 and a standard deviation of 5 and W has a mean of 4 and a standard deviation of 2. Let X and W be independent random variables and let $Y = 2X - 5W$.

(i) What are the mean and standard deviation of Y?

(ii) What can we say about the shape of the distribution of Y?

NOTES:

- When you need to use the Normal distribution for calculations, you will need to use either *Excel*, MINITAB or a graphics calculator. You will probably need to use the computer in several short sessions. Make sure you are prepared before you get to use the computer.

- Do not hand in any computer output. Simply use the computer package (or graphics calculator) to find the solutions. Report any probabilities to 4 decimal places.

Question 2. [7 marks]

A statistics lecturer and a mathematics lecturer went fishing at a lake during a break at a conference in England. Beforehand they agreed that the one who caught the more impressive fish would have their dinner paid for by the other. The statistics lecturer caught a 40 cm long perch, while the mathematics lecturer caught a 42 cm bream.

From long term studies of the fish populations in the lake, it is known that the length of perch is approximately Normally distributed with a mean of 29.6 cm and a standard deviation of 9.5 cm. The length of bream is approximately Normally distributed with a mean of 38.4 cm and a standard deviation of 4.2 cm.

(a) Which lecturer should pay for dinner? Justify your answer.

(b) The next day the mathematics lecturer went fishing again and caught 1 bream and 1 perch. What is the probability that the bream is longer than the perch? What assumption was made to do this calculation?

Question 3. [16 marks]

A tax inspector is conducting random spot checks on tax returns for small businesses. A sample of small businesses is randomly chosen by computer and then their tax records are checked by several different criteria. If the records fail any of the criteria then the records are forwarded on to be audited in detail.

Assume that the time taken to carry out a spot check on a small businesses' tax return is approximately Normally distributed with a mean of 15.8 minutes and a standard deviations of 3.4 minutes.

(a) The following question was posed to a group of students: *What are the expected value and standard deviation for the total time taken to spot check 17 randomly selected businesses' accounts?* One of the students answered as follows:

Let C be the time taken to spot check a single businesses accounts.
$C \sim$ Normal(μ =15.8 minutes, σ = 3.4 minutes)
Let T be the total time taken to spot check 17 businesses accounts.
$T = 17 \times C$. $E[T] = 17 \times E[C] = 17 \times 15.8 = 268.6$. $sd[T] = 17 \times sd[C] = 17 \times 3.4 = 57.8$.

What incorrect assumption underlies this working? Correct the calculations.

(b) The inspector sets aside 5 hours of each day to do spot checks.

(i) On what proportion of days could the inspector complete 17 spot checks in the allocated time?

(ii) Let R be the time remaining from the five hours after carrying out 17 spot checks. Give the distribution (with values of parameters) for R.

(iii) The inspector decides that after carrying out 17 spot checks, if there is more than 12 minutes time remaining, she will start another spot check, otherwise she will start on other work. Using this strategy, what is the probability of the inspector starting an 18^{th} spot check?

(c) From experience, the inspector has found that 10% of returns that are spot-checked are forwarded on for auditing. Assume that the inspector carries out exactly 17 spot checks each day.

(i) Let A be the number of accounts forwarded on for auditing on a particular day. Give the distribution (with values of parameters) for A. Justify your choice of distribution.

(ii) What is the expected number of accounts forwarded on for auditing on a particular day? What is the standard deviation of the number of accounts forwarded on for auditing on a particular day?

(iii) Give the distribution (with values of parameters) for the daily average number of accounts forwarded on for auditing from 50 working days. Justify your choice of distribution.

(iv) Is the distribution in (iii) exact or approximate? Justify your answer.

Question 4. [19 marks]

The manager of an importing company purchased a new machine for packaging rice. The specifications of the machine claim that the amount of rice put in each package will be Normally distributed with an average amount of rice as specified and a standard deviation of 2.7 grams. The machine is set to fill packets with 506 grams of rice.

The manager requires the machine to produce packets containing rice weighing within the range 500-512 grams for at least 95% of the packets.

(a) Assuming that the specifications for the machine are accurate, calculate the probability that a packet of rice contains the desired amount of rice. Will the managers requirements be met?

(b) Let $\overline{X}$ be the mean weight from a sample of 25 packets of rice. Assume that the specifications for the machine are accurate.

(i) Give the distribution and parameters for $\overline{X}$.

(ii) Was the central limit theorem needed to answer (i)? Briefly justify your answer.

(iii) Calculate the interval within which the central 95% of values of $\overline{X}$ should fall.

(c) Let $\hat{P}$ be the proportion of packets of rice outside the 500 – 512 gram weight range from a sample of 800 packets of rice. Assume that the specifications for the machine are accurate.

(i) Give the distribution and parameters for $\hat{P}$.

(ii) Was the central limit theorem needed to answer (i)?

(iii) Calculate the interval within which the central 95% of values of $\hat{P}$ should fall.

(d) A sample of 25 packets of rice was filled and weighed. The resulting weights were as follows:

502.8, 507.6, 515.0, 499.2, 511.9, 503.4, 506.3, 505.5, 502.6, 501.9,
500.8, 502.9, 503.9, 510.3, 502.4, 502.9, 505.4, 510.4, 501.6, 501.4,
498.3, 504.0, 504.4, 504.5, 503.9.

sample size = 25, sample mean = 504.5 grams, sample standard deviation = 3.93 grams.

(i) Create a stem-and-leaf plot of the data (either by hand or by computer).

(ii) Does the sample of weights appear to have an exactly Normal distribution? Briefly justify your answer.

(iii) Are there any features of the stem-and-leaf plot that suggest major departures from the Normal distribution? Briefly justify your answer.

(iv) Based on the above answers, is it <u>likely</u> that the sample data has come from a Normal distribution?

(e) Assuming that the specifications of the machine are correct, would it be unusual to get a sample of 25 packets of rice with a mean weight of 504.5 grams Does getting a sample of 25 packets of rice with a mean weight of 504.5 grams cast any doubts on the specifications of the machine? [Hint: refer to **(b)**.]

(f) A sample of 800 packets of rice are produced. Of these, 26 were found to be outside the 500 – 512 gram weight range.

Assuming that the specifications of the machine are correct, would it be unusual to get a sample of 800 packets of rice with 26 of the packets falling outside the 500 – 512 gram weight range. Does getting a sample of 800 packets of rice with 26 of the packets falling outside the 500 – 512 gram weight range cast any doubts on the specifications of the machine? [Hint: refer to **(c)**.]

(g) Write a **brief** report (no more than 4 sentences) discussing whether or not the given specifications of the machine appear to correct.