THE UNIVERSITY OF SYDNEY
MATH1905 Statistics Advanced

Semester 2	**Assignment 1 – Problems**	2012

This assignment is due by Tuesday, 2nd October, 4pm. Your assignment should be posted in the locked collection boxes at the western end of the verandah (closest to Eastern Avenue) on Carslaw Level 3. Please do not post your assignment before the due date since the boxes are also sed for the collection of assignments in other units. Your assignment should be stapled inside a manila folder, on the front of which you should write the initial of your family name as a LARGE letter, and a cover sheet signed and attached. The cover sheet may be downloaded from the school's MATH1905 website. Assignments without a cover sheet may not be marked. This assignment is worth 10 per cent of your assessment for MATH1905. There are five questions worth a total of 50 marks, and the breakdown of marks is indicated.

1. (Problem sheet 2 question 2) [worth 13 marks in total] The following list gives the number of days with rain from 1977 - 1990 for Wollongong for July, August and December.

July	2	8	6	7	6	12	8	15	7	9	11	6	12	12
August	5	7	4	4	7	3	12	6	10	9	16	10	9	9
December	8	19	7	12	13	12	18	10	16	9	16	19	15	13

Use R or do the following by hand:

(a) Provide for each month the five number summary.

(b) Calculate the coefficient of correlation between the July and August figures and between the July and December figures. Comment on any difference.

(c) Assume you had the number of days with rain in July of an additional year, i.e. your new July data is

$$2,\ 8,\ 6,\ 7,\ 6,\ 12,\ 8,\ 15,\ 7,\ 9,\ 11,\ 6,\ 12,\ 12,\ x_{15}$$

Determine the range of x_{15} such that this new observation would appear as a potential outlier in the boxplot.

2. (Problem sheet 2 question 4) [worth 10 marks in total] Use the R function `sample` to simulate the following experiments

(a) throw a fair coin 100 times (1=head, 0=tail)

(b) throw a coin 100 times where P(Head) = 3/10

(c) throw a coin 100 times where P(Head) = 9/10

(d) throw a 6 faced-die 100 times, the die is loaded in such a way that it is twice more likely to get an even number than an odd number.

(e) Toss a fair coin 50 times. Note: you can sample from a vector whose components are letters: `c('H','T')`

(f) Pick 6 of 54 lottery numbers (without replacement)

(g) Pick 5 `cards=paste(rep(c('A',2:10,'J','Q','K'),4),c('H','D','S','C'))`

(h) Roll 2 dice 5 times. (Hint: `dice=as.vector(outer(1:6,1:6,paste))`)

3. (Problem sheet 7 question 5) [worth 10 marks in total] Don tosses a fair coin 11 times while Giulia tosses a fair coin 10 times. What is the probability that Don obtained more heads than Giulia?

4. (Problem sheet 7 question 7) [worth 7 marks in total] It is known from clinical trials that 75% of patients with vascular migraine obtain relief within one hour with a new (expensive!) migraine tablet. The tablet is prescribed by a medical centre for 120 unrelated patients who suffer from this form of migraine.

(a) What is the probability that at least 80 of these patients will obtain the relief promised?

(b) What is the probability that more than 96 obtain the relief promised?

(Use the normal approximation with correction for continuity.)

5. (Problem sheet 8 question 2) [worth 10 marks in total] It has been claimed that at least 60% of all purchasers of a certain computer program will call the manufacturer's hotline within one month of purchase. A random sample of 12 purchasers of this software is drawn and 3 of those in the sample had contacted the hotline within one month of purchase. Does this provide evidence that the claim of a 60% contact rate is an overestimate? Let p be the true proportion of all purchasers who contact the hotline.

(a) Set up appropriate hypotheses to perform a statistical test.

(b) Why is a 1-sided test appropriate here?

(c) Calculate an exact P-value based on these data. Interpret your findings.