F71SB2 Statistics 2 Assignment 2

NAME (BLOCK LETTERS).....
I affirm that this assignment is my sole and original work.
SIGNATURE.....
TUTORIAL CLASS: Day.....Time.....Room.....

Due: 25 February 2008 at 4.30pm (use the AMS course-work box, outside EM 1.25).

INSTRUCTIONS: Attempt ALL FOUR questions. To receive full credit you must **show your work** and **explain your answers**. There are 40 marks available.

1. Let X be a continuous random variable whose probability density function (*pdf*) is

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & \text{for } 1 < x < \infty; \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find an expression for the cumulative distribution function (*cdf*) $F_X(x)$. [3]

(b) Hence, or otherwise, calculate the probability $P(X > 5)$. [2]

(c) Consider now a transformation of X given by $Y = g(X) = X^{-1}$.

i) Determine the range of the random variable Y. [1]

ii) Derive the *pdf* of Y. [4]

ii) Identify fully the distribution of Y using your answers to i) and ii). [2]

2. Test scores (out of 100) in a large class of students are thought to be normally distributed with mean $\mu = 62$ and standard deviation $\sigma = 12$.

   (a) Calculate the probability that the score of a randomly selected student from this class will be greater than 40. [4]

   (b) Find the score which will be exceeded by the score of a randomly selected student with probability 0.05. [4]

3. The height (in cm) of 30 randomly selected plants of a certain kind are given below (in ascending order):

25.1	25.5	27.2	27.9	28.8	28.9	29.1	29.7	29.9	30.1
30.3	30.6	30.6	31.1	31.3	31.7	31.7	31.9	32.2	32.4
32.6	32.9	33.3	33.5	33.8	34.2	34.5	34.9	35.3	36.3

(For the above data $\sum x_i = 937.3$ and $\sum x_i^2 = 29504.21$)

(a) Calculate the median, first quartile and third quartile of these data. [3]

(b) Calculate the interquartile range (IQR) and explain briefly its meaning. [2]

(c) Construct a stem-and-leaf diagram for the data. Does the plot support the suggestion that the distribution of heights is Normal? Give reasons for your answer. [4]

(d) Calculate the sample mean and sample standard deviation of the data. [3]

(e) What percentage of these data lie within (i) one sample standard deviation of the sample mean; (ii) two sample standard deviations of the sample mean? Do these percentages support the suggestion that the distribution of heights is Normal? [3]

4. A fair coin is tossed independently 100 times. Use the Central Limit Theorem to calculate the approximate probability that the number of times that a 'head' is achieved will exceed 40. [5]

BLANK (*use this space to continue your answers if needed*)