Student name:______________________________

CE93 -- Engineering Data Analysis
First Midterm Examination
Wednesday, October 8, 2003

Work on all three problems. Write clearly and state any assumptions you make. The exam is closed books and notes.

The problems have the following weights:

Problem 1 (35 points) ____________________________
Problem 2 (30 points) ____________________________
Problem 3 (35 points) ____________________________

Exam grade (100 points) __________________________

Student name:_________________________________

Problem 1. (10+15+10= 35 points)

Biochemical oxygen demand at 40 stations along a river measured in milligrams/liter are listed in the **ascending** order in the following table.

2.27	2.89	3.19	3.74
2.46	2.93	3.22	3.75
2.49	2.96	3.23	3.83
2.51	3.00	3.24	3.92
2.70	3.04	3.30	4.00
2.73	3.08	3.36	4.03
2.76	3.13	3.37	4.08
2.79	3.16	3.43	4.41
2.82	3.17	3.58	4.64
2.86	3.18	3.66	4.95

a) Determine the median and the first and third quartiles of the data and the inter-quartile range, *IQR*.

b) Sketch a box plot of the data on the attached graph sheet. Clearly identify the key elements of the plot

c) Are there any outliers? If so, identify the data point(s).

Solution:

a) $x_{0.50} = (3.18 + 3.19)/2 = 3.185$ milligrams/liter median and second quartile

$x_{0.25} = (2.86 + 2.89)/2 = 2.875$ milligrams/liter first quartile

$x_{0.75} = (3.66 + 3.74)/2 = 3.700$ milligrams/liter third quartile

$IQR = 3.700 - 2.875 = 0.825$ milligrams/liter inter-quartile range

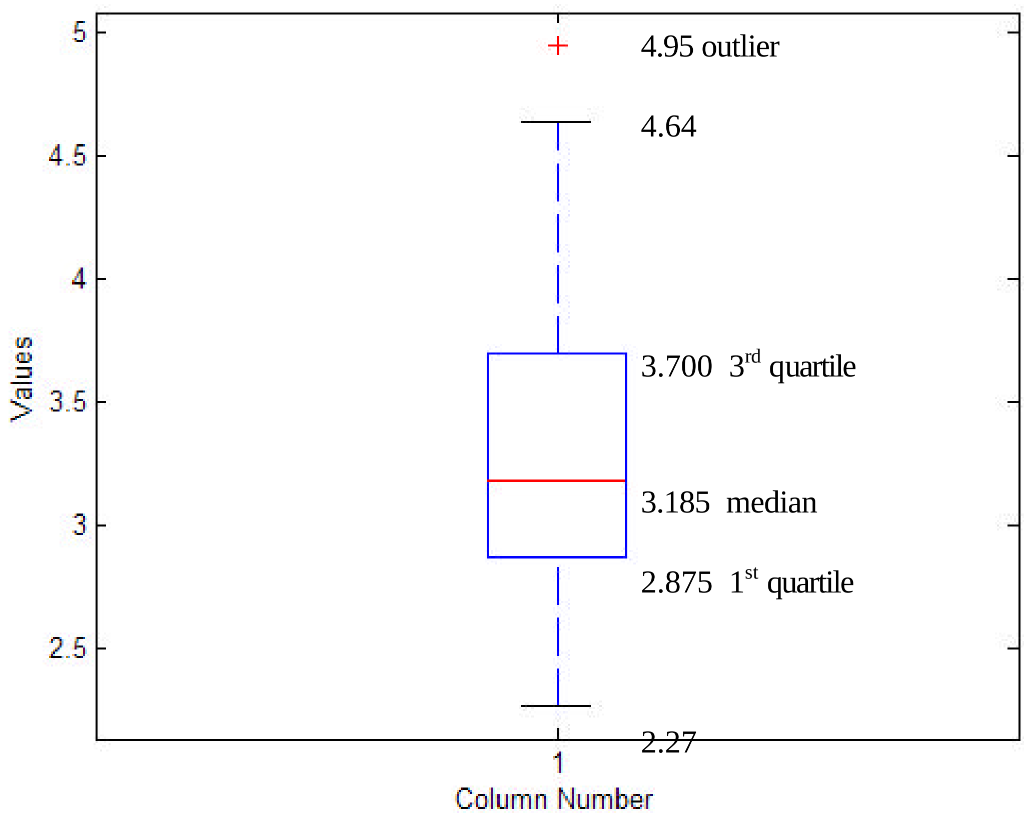
b) $x_{0.25} - 1.5 \times IQR = 2.875 - 1.5 \times 0.825 = 1.638$ milligrams/liter

$x_{0.75} + 1.5 \times IQR = 3.700 + 1.5 \times 0.825 = 4.938$ milligrams/liter

Lower whisker ends at 2.27 (smallest data value within 1.5*IQR* below the 1st quartile).
Upper whisker ends at 4.64 (largest data value within 1.5*IQR* above the 3rd quartile).
See attached page for the box plot.

a) There is one outlier, which is the data point 4.95.

Student name:____________________________________



Problem 2. (15+15 = 30 points)

A contractor is estimating the probability of completing a construction job on time through winter months. She estimates the probability of on-time completion to be 95% if the weather remains good (G), 70% if it rains (R), and 50% if it snows (S). Long time forecast suggests the probabilities of these weather conditions to be 0.4, 0.4 and 0.2, respectively.

a) compute the probability of on-time completion of the construction job.

b) Suppose you are told that the job was not completed on time. Determine the probability that it snowed.

Solution:

Let C = one-time completion of construction job. We are given:

$$P(C \mid G) = 0.95, \quad P(C \mid R) = 0.70, \quad P(C \mid S) = 0.50$$

$$P(G) = 0.4, \quad P(R) = 0.4, \quad P(S) = 0.2$$

a)

$$\begin{aligned} P(C) &= P(C \mid G)P(G) + P(C \mid R)P(R) + P(C \mid S)P(S) \\ &= (0.95)(0.4) + (0.70)(0.4) + (0.50)(0.20) \\ &= 0.76 \end{aligned}$$

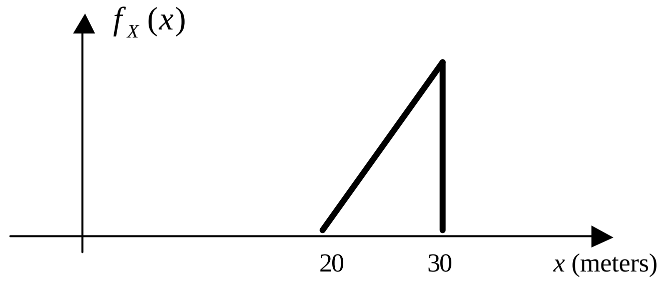
b)

$$\begin{aligned} P(S \mid \overline{C}) &= \frac{P(\overline{C} \mid S)}{P(\overline{C})} P(S) \\ &= \frac{1 - P(C \mid S)}{1 - P(C)} P(S) \\ &= \frac{1 - 0.50}{1 - 0.76} (0.20) \\ &= 0.417 \end{aligned}$$

Student name:________________________________

Problem 3. (15+5+10+5 = 35 points)

A soil engineer estimates the depth X to the bedrock below the foundation of a building to be somewhere between 20 and 30 meters, but more likely towards the higher values. On this basis, he assigns the triangular probability density function shown below to express his uncertainty about the depth.



a) Determine the mean, the median, the mode, the standard deviation, and the coefficient of variation (c.o.v.) of X.

b) What do you guess is the skewness coefficient of the distribution? (Do not make any calculations.)

c) Compute and plot the cumulative distribution of X.

d) What is the probability that the depth will be between 25 and 27 meters?

Solution:

a) Determine the expression for PDF with unit area underneath:

$$f_X(x) = 0.02(x-20) \quad \text{for} \quad 20 < x < 30$$
$$= 0 \quad \text{elsewhere}$$

$$\mu_X = \int_{20}^{30} 0.02x(x-20)dx = 0.02\left(\frac{x^3}{3} - 10x^2\right)_{20}^{30} = 0.02\left(\frac{(30)^3}{3} - 10(30)^2 - \frac{(20)^3}{3} + 10(20)^2\right)$$
$$= 26.667\text{m} \quad \text{mean}$$

Find point that halves the area:

$$\frac{1}{2}(x_{0.50} - 20)\left(\frac{x-20}{10/0.2}\right) = 0.5 \quad \Rightarrow \quad (x_{0.50} - 20)^2 = 50$$

$$x_{0.50} = 20 + \sqrt{50} = 27.071\text{m} \quad \text{median}$$

$$\tilde{x} = 30\text{m} \quad \text{mode}$$

$$\text{E}[X^2] = \int_{20}^{30} 0.02x^2(x-20)dx = 0.02\left(\frac{x^4}{4} - \frac{20}{3}x^3\right)_{20}^{30} = 0.02\left(\frac{(30)^4}{4} - \frac{20(30)^3}{3} - \frac{(20)^4}{4} + \frac{20(20)^3}{3}\right)$$
$$= 716.667\text{m}^2 \quad \text{mean-square}$$

$$\sigma_X = \sqrt{716.667 - 26.667^2} = 2.357 \text{ m} \quad \text{standard deviation}$$

$$\delta_X = \frac{2.357}{26.667} = 0.0884 \quad \text{c.o.v.}$$

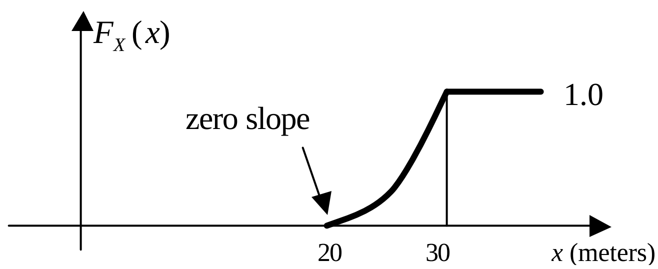
b)

The distribution is strongly skewed to the left. My guess is that the skewness coefficient is around -2.

c)

$$F_X(x) = \int_{20}^{x} 0.02(x-20)dx = 0.01(x-20)^2 \quad 20 < x < 30$$

$$= 1 \quad 30 \le x$$



d)

$$P(25 < X < 27) = F_X(27) - F_X(25)$$
$$= 0.01(27-20)^2 - 0.01(25-20)^2$$
$$= 0.24$$