BIO 300 ASSIGNMENT #1

NOTE: This assignment is due Friday, September 24th.

1. The following 40 measurements are glucose utilization rates for a random sample of carp (*Cyprinus carpio*).

7.8	8.7	9.3	9.1	8.5	6.3	6.8	7.8	7.2	7.6
8.0	7.8	8.6	7.3	8.1	7.4	8.8	7.0	7.3	8.2
7.5	7.5	8.1	7.9	8.3	6.9	7.3	7.1	8.6	9.6
7.7	6.9	7.3	7.3	6.7	8.4	8.9	8.2	7.1	7.8

   (a) Construct a frequency table for the following 40 measurements. Give the cumulative frequency in addition. (*Note: Your first problem will be to decide on the number of width classes. In general, samples of fewer than 40 or 50 shouldn't be given more than 10 classes, otherwise too few frequencies per class will result. Samples of a few thousand can use more than 20 classes*.)
   (b) Draw a histogram for these measurements labelling edges of bins not the midpoints.
   (c) For these same data, calculate the sample mean. (*Note: While computing, retain as many decimals as your calculator can tolerate, but print statistics with only one more significant digit than in the data provided).*
   (d) For these same data, calculate the sample median. (Ignore Zar's complex formula - stick with that used in class).
   (e) Write down the formula for the sample standard deviation. Compute the standard deviation for the above data set.
   (f) What fraction of the values in the above data set lie within one standard deviation of the sample mean? Within two standard deviations?

2. A paleontologist measured the width (in mm) of the last upper molar in 36 specimens of the extinct mammal *Acropithecus rigidus*. The results were as follows:

6.1	5.7	6.0	6.5	6.0	5.7
6.1	5.8	5.9	6.1	6.2	6.0
6.3	6.2	6.1	6.2	6.0	5.7
6.2	5.8	5.7	6.3	6.2	5.7
6.2	6.1	5.9	6.5	5.4	6.7
5.9	6.1	5.9	5.9	6.1	6.1

   (a) Construct a frequency table.
   (b) Draw the histogram for these measurements.
   (c) Calculate the sample mean, median, standard deviation and variance.

3. Distinguish, in words, between:
   (a) Bar graph and histogram.
   (b) Mean, median and mode.
   (c) Population and sample.
   (d) Random and non-random sample.
   (e) Standard deviation and standard error.

(f) Parameter and statistic.

4. A student assessed the effect of parasitism by the cestode *Schistocephalus solidus* on the feeding behaviour of adult sticklebacks (*Gasterosteus aculeatus*). Feeding strikes on *Artemia nauplii* by parasitised and non-parasitised fish are recorded below:

Feeding Strikes Per Minute

Parasitised	Non-parasitised
4.7	18.6
4.8	9.2
7.2	16.2
1.6	28.2
6.8	17.1
5.2	12.8
4.2	19.1
6.2	23.0
7.8	29.0
3.1	9.5

(a) Compute the sample mean and standard deviation for parasitised and non-parasitised fish. Which group has the higher mean feeding rate? Which group of fish is more variable?
(b) Compute the standard error of the mean for each group. Which group's $\overline{X}$ would you consider as the more reliable measure of population mean feeding rate? Why?
(c) Calculate the coefficient of variation for both groups of fish. Do parasitised and non-parasitised fish differ in their variability relative to the mean?

5. A population of sockeye salmon was sampled on two successive days for PCB concentration. The data are summarized below:

	Day 1	Day 2
Sample size	32	45
Mean	42.5 ppb	35.4 ppb
Standard deviation	8.1	6.9

(a) Calculate the grand mean for the combined sample of 77 salmon measurements.
(b) Calculate the pooled sample variance.
(c) Can you calculate the variance for the combined sample? (*Note: this is not the same as the pooled sample variance*).

6. What effect would adding the constant 5.0 to every measurement in Q1 have on:

(a) the sample mean?
(b) the sample standard deviation?
(c) The coefficient of variation?
(d) The standard error?
(e) What effect would doubling all values have on the sample variance?