Home assignment 1

Exercise 1.1

Hand calculations: The number of coliforms (on a log-transformed scale) were counted in milk samples.

3.9
5.3
6.1
4.9
9.1
2.8
3.5
3.2
2.6
5.9

    a) Compute the mean value and the variance for the sample
    b) Compute the median in the sample

Exercise 1.2

Computer exercise: On the course homepage you will find the file `cordblood.txt`, which contains data from serological blood samples from the umbilical cord at 369 births in Sweden in 2007.
The main purpose of the study was to investigate the immunity against four common diseases in Sweden: measles, parotitis, rubella and chicken pox.

The variables in the file are the following

Hospital	The hospital where the sample was taken.
Age	The age of the mother.
Sex	The sex assigned to the child.
Measles	The level of antibodies against measles in the cord blood.
Parotitis	The level of antibodies against parotitis in the cord blood.
Rubella	The level of antibodies against rubella in the cord blood.
Chickenpox	The level of antibodies against chicken pox in the cord blood.

There is a mixture of categorical and numerical variables in the data set.

a) Read the file into R as you did with the cats data in the computer exercise.

b) Provide the following graphs and computations with some comments/interpretation on what they show

1. Produce a histogram and a boxplot of the four antibody variables (Measles, Parotitis, Rubella or Chickenpox). (using the code `par(mfrow=c(4,2))` or `c(2,4)` allows you to plot 8 graphs in the same window)
2. Calculate mean, median, variance and interquartile range for all four antibody variables. Is it more appropriate to use mean and variance or median and interquartile range in this case? Why?
3. Calculate the central measurement and dispersion you deemed more appropriate for the four variables, grouped by the assigned sex of the child.
4. Make a scatter plot with two antibody variables against each other.
5. Make a scatter plot with an antibody variable against the age of the mother, show the results for the different sexes.
6. Produce boxplots of an antibody variable but also use the categorical variable hospital to split the data.

(Do not forget to include your comments for each graph or computation).