This homework assignment is to be presented on exercise session on **November 12, 2019**. Students should tick in TUWEL problems they have solved and are prepared to present their detailed solutions on blackboard. The solved problems should be ticked by **23:30 on November 11, 2019**.

---

(1) **CPU workload**
The CPU workloads (in %) of a processor were observed eight times and gave

$$25, 13, 7, 9, 44, 3, 2, 33$$

Find all empirical (a) medians, (b) first quar_tiles and (c) 2/3-qua_ntiles.

(2) **Boxplot**
Two novel randomized algorithms (A and B) are to be compared regarding their running time. Both algorithms were executed n times. The running times (in seconds) are stored in the file `algorithms.Rdata`

(a) Set the working directory and load the data using `load()`. Create a boxplot to compare the running times. Color the boxes and add proper notations (axes notations, title etc.). More info via `?boxplot`

(b) Comment on the following statements / questions only using the graphic
  (a) The first quartile of the times in A was about?
  (b) the interquartile range of the times in B is about trice the interquartile range of A
  (c) Is $n = 100$?
  (d) More than half of the running times in B were faster than 3/4 of the running times in A
  (e) At least 50% in A were faster than the 25% slowest in B
  (f) At least 60% in A were faster than the 25% slowest in B

(3) **Histogram**
Set `k <- 100` and generate `x <- rnorm(sample(k:(2*k),1), runif(1,0,k), rexp(1,1/k))`

(a) Explain what is realized in `x`.

(b) Plot a histogram of `x`. Mark its mean in red, its standard deviation in blue and add a legend which explains them both. Helpful commands: `hist()`, `mean()`, `sd()`, `lines()`, `abline()`, `arrows()`, `legend()`

(4) **Unbiasedness of the empirical variance**
Let $n \geq 2$ and $X_1, \ldots, X_n$ be i.i.d. (independent and identically distributed) random variables, with $\sigma^2 := \mathbb{V}ar(X_1) < \infty$. Calculate the expectation of the empirical variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

What would have been the expectation if in S^2 we had scaled with n instead of $n - 1$?

(5) **Postbox 1**

Let $x_1, x_2, \cdots, x_n$ be the locations of n households along a street. At which position p should a postbox be placed such that

$$\sum_{i=1}^{n} |x_i - p|$$

is minimized?

(6) **Postbox 2**

For the setup in **Postbox 1**, which p minimizes

$$\sum_{i=1}^{n} (x_i - p)^2$$

(7) $N(0,1)$**-distribution and neighborhoods**

(a) Plot the density of the $N(0,1)$ distribution.

(b) Which quantiles of $N(0,1)$ mark the neighborhoods of zero that contain first 95%, second 99% and third 99.9% of the probability mass? Which values do they take?

(c) Add these neighborhoods to your plot.

Hint: `plot()`, `qnorm()`