This is the fourth homework assignment. The problems are to be presented on exercise session on **November 5, 2019**. Students should tick in TUWEL problems they have solved and are prepared to present their detailed solutions on blackboard. The problems should be ticked by **23:30 on November 4, 2019**.

---

(1) **Human resource testing**
Some human resource deprtments administer standard IQ tests to all employees. The Stanford-Binet test scores are well modeled by a Normal model with expectation 100 and standard deviation 16.

(a) If the applicant pool is well modeled by this distribution, a randomly selected applicant would have what probability of scoring in the region between 84 and 116?

(b) For the IQ test administered by human resources, what cutoff value would separate the middle 95%?

(2) **Coin throws**
An unfair coin is thrown 600 times. The probability of geting a tail in each throw is $\frac{1}{4}$.

(a) Use a Binomial distribution to compute the probability that the number of heads obtained does not differ more than 10 from 250.

(b) Use a Normal approximation without a continuity correction to calculate the probability in (a). How does the result change if the approimation is provided with a continuity correction?

(3) **Cars arrivals**
Suppose cars arrive at a parking lot at a rate of 50 per hour. Assume that the process is modeled by a Poisson random variable with $\lambda = 50$.

(a) Compute the probability that in the next hour the number of cars that arrive at this parking lot will be between 54 and 62.

(b) Compare the value obtained in (a) with the probability calculated by using a Normal approximation.

(4) **Coffee and doughnuts**
At a certain coffee shop, all the customers can buy a cup of coffee and also a doughnut. The shop owner believes that the number of cups he sells each day is normally distributed with an expectation of 320 cups and a standard deviation of 20 cups. He also believes that the number of doughnuts he sells each day is independent of the coffee sales and is normally distributed with an expectation of 150 doughnuts and a standard deviation of 12.

(a) The shop is open every day but Sunday. Assuming day-to-day sales are independent, what is the probability he will sell more than 2000 cups of coffee in a week?

(b) If he makes a profit of 50 cents on each cup of coffee and 40 cents on each doughnut, can he reasonably expect to have a day's profit over 300 euro? Justify your answer.

(5) **Sum and average**
Let X be a random variable with $\mathcal{N}(5, 2^2)$. Let $X_1, X_2, \ldots, X_{50}$ be independent identically distributed copies of X. Let S be their sum and $\bar{X}$ their average, i.e.

$$S = X_1 + \cdots + X_{50} \qquad \text{and} \qquad \bar{X} = \frac{1}{50}(X_1 + \cdots + X_{50}).$$

(a) Plot the density and the cumulative distribution function for X.

(b) What are the expectation and the standard deviation of S and of $\bar{X}$?

(c) Generate a sample of 50 numbers from $\mathcal{N}(5, 2^2)$. Plot the histogram for this sample. Do the same for a sample of 500 numbers from $\mathcal{N}(5, 2^2)$.

(6) **Simulations**

(a) By applying the *R*-function `replicate()` generate a sample $X_1, \ldots, X_{10}$ of size 10 from an exponential distribution with a rate parameter 0.2 and sum up its elements. Do this sum 10 000 times and make a histogram of the simulation. Can you say something about the shape of distribution?

(b) Use *R* to simulate 50 tosses of a fair coin (0 and 1). We call a *run* a sequence of all 1's or all 0's. Estimate the average length of the longest run in 10000 trials and report the result.

Hint: Use the commands `rbinom` and `rle`. The command `rle`() stands for run length encoding. For example,

```
rle(rbinom(10, 1, 0.5))$lengths
```

is a vector of the lengths of all the different runs in trial of 10 flips of a fair coin.

(7) **Histograms of averages of** $\exp(1)$

(a) Generate a frequency histogram of 1000 samples from an $\exp(1)$ random variable.

(b) Generate a density histogram for the average of 2 independent $\exp(1)$ random variables.

(c) Using `rexp`(), `matrix`() and `colMeans`() generate a density histogram for the average of 50 independent $\exp(1)$ random variables. Make 10000 sample averages and use a binwidth of 0.1 for this. Look at the spread of the histogram.

(d) Add a graph of the pdf of $\mathcal{N}(1, 1/50)$ on your plot in problem (c).