# Exercises (Combinatorics, computer science, statistics)

**Task 10** *Combinatorics*

Permutations can be defined as words of length $n$ over an alphabet with $n$ elements, with the restriction that no letter appears twice in the word.

(a) List systematically all permutations of the $n = 4$ letters A, B, C, D.

(b) Prove for the number $P(n)$ of permutations of $n$ letters:

$$P(n) = n \cdot P(n-1) \quad \text{if } n > 1.$$

**Task 11** *Graphs*

List all undirected, non-isomorphic graphs with 4 vertices and 3 edges. (Multiple edges and loops are not allowed.) How many of them are connected, how many are not?

**Task 12** *Digital storage requirements*

Which data needs more storage capacity:

(a) a DNA sequence of 8 million nucleotides (each being represented as A, T, G or C and requiring 2 bits), or

(b) a colour image, consisting of $1024 \cdot 1024$ pixels, where each pixel carries 24 bits of colour information?

Prove your answer by evaluating both storage requirements in terms of MB (MegaBytes) and comparing them.

**Task 13** *Programming*

(a) Write a mathematical formula which gives the meaning of the Java expression
`x + Math.sqrt(1-x*x)/2`.

(b) What kind of runtime error can easily occur when the above expression is evaluated in a Java programme? Which mathematical condition (for **`x`**) has to be fulfilled to prevent this error?

**Task 14** *Rule-based simulation*

Draw the geometrical structure which is generated after 2 steps of rule application by the following L-system (or XL code), assuming the 3-d object definitions

```
module Bud extends Sphere(1.0);    /* 1.0 is radius */
module Shoot extends F(5.0, 1.0);  /* 5.0 is length of
                                      cylinder, 1.0 is diameter */
```

are in force and the start word is **`Bud`**:

```
Bud ==> Shoot [ RU(45) Bud ] [ RU(-45) Bud ];
```

**Task 15** *Statistics / location and dispersion measures*

Given are heights of trees in m ($n$ = 11 trees):

5.9 7.5 6.6 10.2 7.8 9.4 8.5 7.2 9.1 9.6 9.5

(a) Calculate the $0^{th}$, $25^{th}$, $50^{th}$, $75^{th}$ and $100^{th}$ percentiles (named $Q_0$, $Q_{25}$, $Q_{50}$, $Q_{75}$, $Q_{100}$).

(b) Calculate the median and the mean.

(c) Calculate the variance, the standard deviation and the coefficient of variation.

(d) Calculate the range and interquartile range.

**Task 16** *Statistics / Gaussian distribution*

Given is a field with 100.000 corn plants. The mean height of a plant is $\mu$ = 150 cm and the standard deviation of the population is $\sigma$ = 25 cm. The height is approximately normally distributed. One sample containing $n$ = 25 plants was taken from this population.

(a) What percentage of individual values do you expect to lie between 135 and 180 cm?

(b) What percentage of individual values do you expect to lie between 155 and 175 cm?

(c) What percentage of sample mean values do you expect to lie between 140 and 155 cm?

(d) What percentage of sample mean values do you expect to lie between 140 and 145 cm?

**Task 17** *Statistics / Confidence intervals, one-sided Gauss test*

A sample of size $n$ = 25 was taken from a population of corn plants. The measured average height of the 25 plants was $\bar{x}$ = 155 cm and the standard deviation of the population is $\sigma$ = 25 cm. The height is approximately normally distributed.

(a) Compute a 99% confidence interval for the mean height.

(b) Formulate the null hypothesis and answer the question whether the mean of the population from which the sample was taken from is significantly higher than 145 cm ($\alpha$ = 0.05).

(c) Formulate the null hypothesis and answer the question whether the mean of the population from which the sample was taken from is significantly lower than 160 cm ($\alpha$ = 0.05).

**Task 18** *Statistics / Linear regression and correlation, coefficient of determination*

Investigated was the relationship between the height of trees in m ($y$) and their diameter in cm ($x$).
The estimated regression equation was as follows: $\hat{y} = 1.4 + 0.6\,x$ ,
and the estimate of the coefficient of correlation was $r$ = 0.91.

(a) What height do you expect for a diameter of 10 cm?

(b) What percent of variability of $y$ can be explained by $x$ ?
Give your answer in words.