

# Exercise Sheet 1 – Probability and Statistics

INF161 Autumn Semester 2020

**Exercise 1** *Compute the probability of rolling a 5 on a fair dice.*

**Exercise 2** *A family has two children. You met one of them, Leo, and he is a boy. What is the probability that the other child is also a boy?*

**Exercise 3** *A building is examined by policemen with four dogs that are trained to detect the scent of explosives. If there are explosives in a certain building, and each dog detects them with probability 0.6, independently of the other dogs, what is the probability that the explosives will be detected by at least one dog?*

**Exercise 4** *The probability that a particular hard drive will crash is 0.01. Therefore, it has two backups, each having a 0.02 probability of crashing. All hard drives crash independently of each other. What is the probability that all three hard drives will crash?*

**Exercise 5** *Suppose that a shuttle's launch depends on three key devices that operate independently of each other and malfunction with probabilities 0.01, 0.02, and 0.02, respectively. The launch is postponed if any of the key devices malfunction. What is the probability that the shuttle launches on schedule?*

**Exercise 6** *Ninety percent of flights depart on time. Eighty percent of flights arrive on time. Seventy-five percent of flights depart on time and arrive on time.*

- 1. You are meeting a flight that departed on time. What is the probability that it will arrive on time?*
- 2. You have met a flight that arrived on time. What is the probability that it departed on time?*
- 3. Are the events departing on time and arriving on time independent?*

**Exercise 7** *As part of a business strategy, 20% of new internet service subscribers receive a special promotion from the provider. A group of 10 neighbors sign up for the service. What is the probability that at least 4 of them get the promotion?*

**Exercise 8** *An exciting computer game is released. Sixty percent of players complete all levels. Of those players, thirty percent then buy an extended version of the game. Players who do not complete all levels never buy the extended version. Among 15 randomly selected players, what is the expected number of people that will buy the extended version? What is the probability that at least two players will buy it?*

**Exercise 9** Suppose that the average household income in some country is 900 coins, and the standard deviation is 200 coins. Assuming that income follows the Normal distribution, compute the proportion of “the middle class”, whose income is between 600 and 1200 coins.

**Exercise 10** A new computer program consists of two modules. The first module contains an error with probability 0.2. The second module is more complex; it has a probability of 0.4 to contain an error, independently of the first module. An error in the first module alone causes the program to crash with probability 0.5. For the second module, this probability is 0.8. If there are errors in both modules, the program crashes with probability 0.9. Suppose the program crashed. What is the probability that there are errors in both modules?

**Exercise 11** Let a probability distribution  $P$  be defined on the sample space  $\Omega$  and let  $A, B \subseteq \Omega$ . Prove that if  $A \subseteq B$ , then  $P(A) \leq P(B)$ .

**Exercise 12** Let a probability distribution  $P$  be defined on the sample space  $\Omega$  and let  $A \subseteq \Omega$ . Prove that  $P(A) = 1 - P(A^c)$  ( $A^c$  is the complement of  $A$ ).

**Exercise 13** Look at the sample in Figure 1 and compute the mean, median, mode, 75% quantile, variance, and standard deviation. Then, calculate the 95% and 90% confidence interval of the mean.

70	36	43	69	82	48	34	62	35	15
59	139	46	37	42	30	55	56	36	82
38	89	54	25	35	24	22	9	56	19

Figure 1: Random sample of CPU time for 30 jobs (in seconds)

**Exercise 14** A research study has been conducted to determine the effectiveness of a drug over time. The table below shows the results of the experiment.

Time (in years)	1	2	3	4	5
Effectiveness (%)	96	84	70	58	52

1. Fit a linear regression model to this data.
2. According to the linear model, when will the effectiveness be 80%? When will the drug have lost all effectiveness?

Answer these questions with a python implementation.

**Exercise 15** The script `visualize.py` generates a sample drawn from a binomial distribution. Run the python script `visualize.py` with different instances of the variable “num\_samples”. Explain what happens to the distribution of the data in different instances of the variable “num\_samples”.