

# Probability and statistics

## Task 1.

E-shop customers choose the delivery of goods by one of the three carriers DA, DB and DC. It was found that out of the total number of sent packages, 1/2 was delivered by DA, 1/3 by DB and the rest by DC. Unfortunately, compared to the announced delivery time, shipments are delayed, namely 5% of shipments by DA, 7% of shipments by DB and 10% of shipments by DC.

1. Determine the probability that the shipment will be delivered late to the customer. (5 points)
2. If the customer complains about the delay in the shipment, what is the probability that the shipment was delivered to him by the DB carrier? (5 points)

## Task 2.

The seller in the online store offers SD memory cards with a capacity of 256 GB. It has the last 7 pieces in stock, 3 of which are of poor quality (they have only half of the capacity compared to the capacity stated by the manufacturer). The customer orders and pays for 4 cards. The seller selects 4 cards at random at his disposal and sends them to the customer.

1. Determine the probability and distribution function of a random variable  $X$ , which expresses the number of poor quality cards among the cards sent to the customer. Write down with tables. (3 points)
2. Determine the mean and standard deviation of the random variable  $X$ . (2 points)
3. The customer complains about any SD card that does not have the appropriate capacity. The seller's loss associated with a complaint about each SD card is CZK 200. Determine the probability function of the total loss of the seller when sending 4 SD cards. (2 points)
4. Determine the mean, standard deviation, and total loss mode when shipping 4 SD cards. (3 points)

## Task 3.

The length of a call to the Customer Support Center (in minutes) is 10 minutes and can be modelled by an exponential distribution.

1. Sketch the probability density of the call duration. (1 point)
2. What length of a call is exceeded by maximum of 11% of the calls? Record the result in the probability density sketch. (2 points)
3. What is the probability that the call will last longer than 12 minutes if it already lasted 8 minutes? (3 points)
4. What is the probability that the average time of 40 calls will not exceed 13 minutes? (4 points)

Text on examples 4 and 5 In the research study, physical activity was monitored in university students (technical, humanity, economics and sport fields). The survey ran from 2019 to 2021. For each of the participating student, data of time (in minutes) spent during the day with moderate to high-intensity physical activity (PA) were recorded based on accelerometer measurements, during present learning and after 6 months of distance learning. The average daily values of PA in minutes (pa\_norm (PA at the time of present learning), pa\_dist (PA at the

time of 6 months of distance learning)) and information about the type of field of study (field) are available in the assigned data file.

**Task 4.**

Analyze PA students during distance learning. Remember to verify the assumptions for the use of statistical induction methods.

1. For students of technical fields, determine point and 95% left-side interval estimates of the mean, or median, of PA at the time of distance learning. Interpret the result. (3 points)
2. For students of technical fields, determine whether the observed mean PA or median PA, at the time of distance learning statistically significantly (at the level of significance of 5%) exceeds 30 minutes. Use an interval estimate and an appropriate pure significance test to verify. (3 points)
3. At the significance level of 5%, determine whether the mean value or median, PA at the time of distance learning for students of sports, is statistically significantly higher than for students of technical fields. Use the appropriate pure significance test for verification. (4 points)

**Task 5.**

At the significance level of 5%, decide whether the mean value or median of PA at the time of distance learning is statistically significantly different between the fields of study. If so, find out whether some fields of study can be marked in terms of average value, or median of PA at the time of distance learning as homogeneous and sort the fields of study according to the monitored parameter and the results of the post-hoc analysis in descending order. Be sure to verify the prerequisites for using the selected test. (10 points)