Random_variables_part1_handson

October 15, 2025

1 Random Variables and Distribution Simulation

[1]:
```
%%html
<link rel="stylesheet" type="text/css" href="../styles/styles.css">
```

```
<IPython.core.display.HTML object>
```

[2]:
```
import matplotlib.pyplot as plt
import matplotlib.patches as mpatches
import numpy as np
from scipy import stats
import math
```

QUESTION:

A minibus shuttle can accommodate up to 5 passengers per trip. The transportation company accepts at most 6 bookings per trip, with each passenger required to have a booking. Past experience has shown that 25% of people making a booking do not show up for the departure. All passengers are assumed to act independently of each other.

Assume that 6 bookings have been made.

1. What is the probability that at least one passenger with a reservation does not show up for departure?
2. What is the average number of passengers showing up for departure?
3. What is the average number of people transported?

Compare your results with the ones obtained using Python.

QUESTION:

We want to analyze the result of an optimization problem. The solution program has a probability p of converging to the desired value. We denote by X the number of trials necessary to obtain m successes. We assume that the trials are independent.

For all $k \in \mathbb{N}$ determine the probability that $X = k$. What is the average number of trials to perform to obtain m successes?

Hint: You might want to check out a negative binomial distribution a.k.a. *Pascal distribution*. In `scipy.stats` refer to `scipy.stats.nbinom`

QUESTION:

The grade obtained by students on an exam is a normal r.v. $X \sim \mathcal{N}(7, 3^2)$.

1. Calculate the percentage of individuals having more than 10, and the grade below which 10% of students are found.

2. Given these results, we decide to revalue all grades by a linear transformation $Z = aX + b$. What values should be given to a and b so that the previous values become 50% and 7 respectively? (*Hint: calculate* $\mathbb{E}(Z)$ *and* $Var(Z)$ *as functions of* $\mathbb{E}(X)$ *and* $Var(X)$).

QUESTION:

We want to model salary income in European countries. Let X be the salary of individuals and r be the minimum income fixed by the country. We assume that X follows the Pareto distribution $\mathcal{P}(\alpha, r)$, with density

$$f(x) = \begin{cases} \alpha r^{\alpha} x^{-(\alpha+1)} = \frac{\alpha r^{\alpha}}{x^{(\alpha+1)}} \text{ if } x > r \\ 0 \text{ otherwise} \end{cases}$$

1. What is the expected value of X?
2. Give the cumulative distribution function of X, denoted F_X. Show that $U = (\frac{r}{X})^a$ follows a uniform distribution.
3. D_9 corresponds to the minimum income of the richest 10% and D_1 corresponds to the maximum income of the poorest 10%. These quantities are given by $F_X(D_9) = 0.9$ and $F_X(D_1) = 0.1$.

In 2014, the observed ratio D_9/D_1 equals 2.81 in France and 3.56 in the United Kingdom (OECD data on gross earnings, Decile ratios of gross earnings).

Calculate the ratio D_9/D_1 for $r = 1$, as a function of α. For what value of α do we have $D_9/D_1 = 2.8$? For what value of α do we have $D9/D1 = 3.56$? How should the parameter α be interpreted?

QUESTION:

Show that if $X \sim \mathcal{E}(\lambda)$, then for all positive real numbers a and b, we have

$$P(X > a + b \mid X > b) = P(X > a)$$

If X represents for example the lifetime of a printer, what does this property mean?

QUESTION:

Let X be a random variable with values in $\{-1, 0, 1\}$ such that $P(X = -1) = 1/3$, $P(X = 0) = 1/2$ and $P(X = 1) = 1/6$. Propose an algorithm (steps only without implementation) to simulate X. We assume that we know how to simulate a continuous random variable with uniform distribution on $[0, 1]$.