Machine Learning:
Assignment Sheet #1

Due on February 15, 2022 at 10:00

Group HB
Henri Sota, Enis Mustafaj

Problem 1.1

The disjoint events A, B, and C are defined in a sample space Ω. Find expressions for the following probabilities in terms of sums of the probabilities $P(\emptyset)$, $P(A)$, $P(B)$, $P(A \cup B)$, $P(A \cup C)$, $P(B \cup C)$, $P(A \cup B \cup C)$:

a) the probability that exactly two of A, B, C occur (*at the same time!*)

The probability that 2 events occur at the same time can be expressed as :

$$P(A \cap B \cap \overline{C}) + P(A \cap \overline{B} \cap C) + P(\overline{A} \cap B \cap C)$$

$$= P((A \cap B) \cap \overline{C}) + P((A \cap C) \cap \overline{B}) + P(\overline{A} \cap (B \cap C))$$

Sine the events are disjoint, the probability can be expressed as:

$$P(\emptyset \cap \overline{C}) + P(\emptyset \cap \overline{B}) + P(\overline{A} \cap \emptyset)$$

$$= P(\emptyset) + P(\emptyset) + P(\emptyset)$$

b) the probability that exactly one of these events occurs

The probability that one event occurs is:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) - P(A \cap B \cap C)$$

Events A, B and C are disjoint, so the probability is:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

Problem 1.2

We consider the chance experiment of four tosses of a fair coin.

a) Give the sample space Ω for this experiment and define a discrete random variable X to describe the number of heads obained in four tosses of a fair coin.

The sample space is:

$$\begin{aligned}\Omega = \{&HHHH, TTTT, HHHT, HHTH, HTHH, THHH, HHTT, HTTH,\\ &TTHH, HTHT, THTH, THHT, HTTT, THTT, TTHT, TTTH\}\end{aligned}$$

The discrete random variable X is defined as:

$$\begin{aligned}X(TTTT) = 0, X(TTTH) = 1, X(TTHT) = 1, X(THTT) = 1\\ X(HTTT) = 1, X(HHTT) = 2, X(HTHT) = 2, X(THHT) = 2\\ X(THTH) = 2, X(TTHH) = 2, X(HTTH) = 2, X(HHHT) = 3\\ X(HHTH) = 3, X(HTHH) = 3, X(THHH) = 3, X(HHHH) = 4\end{aligned}$$

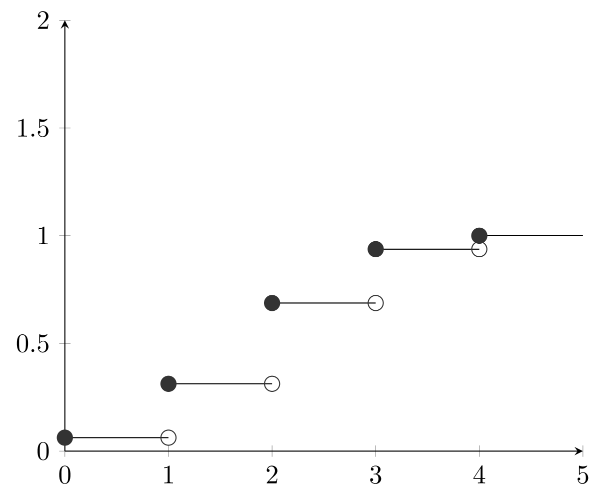
b) Find the PMF and CDF of the defined random variable and plot the CDF.

The PMF of the random variable is:

$$p(X = 0) = \frac{1}{16}, p(X = 1) = \frac{4}{16}, p(X = 2) = \frac{6}{16}, p(X = 3) = \frac{4}{16}, p(X = 4) = \frac{1}{16}$$

The CFD of the random variable is defined as following:

$$F_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{16} & 0 \leq x < 1 \\ \frac{5}{16} & 1 \leq x < 2 \\ \frac{11}{16} & 2 \leq x < 3 \\ \frac{15}{16} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$



c) Compute the probability of the event X IS BIGGER THAN 1.

The probability of the event X IS BIGGER THAN 1 is:

$$F(1 < X \leq 4) = F(4) - F(1) = 1 - \frac{5}{16} = \frac{11}{16}$$

d) Compute the mean of the random variable.

The mean of the random variable is:

$$E(X) = \sum_{x \in R_X} x \cdot P(X = x)$$

$$E(X) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{6}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} = 2$$

Problem 1.3

Let the function ρ with $\rho(x) = \begin{cases} c|x|(1+x)(1-x) & \text{if } -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$ be given.

a) Compute the constant c, such that ρ becomes a density of some random variable.

For ρ to become a density of some random variable, the function must fulfill the requirement that the integral over the domain should equal to 1.

$$
\begin{aligned}
\int_{-\infty}^{\infty} \rho(x)\,dx &= \int_{-\infty}^{\infty} c|x|(1+x)(1-x)\,dx \\
&= \int_{-1}^{1} c|x|(1+x)(1-x)\,dx \\
&= -\int_{-1}^{0} cx(1+x)(1-x)\,dx + \int_{0}^{1} cx(1+x)(1-x)\,dx \\
&= c\left(-\int_{-1}^{0} x(1+x)(1-x)\,dx + \int_{0}^{1} x(1+x)(1-x)\,dx\right)
\end{aligned}
$$

To solve the inner integrals, we use the substitution $u = (1+x)(1-x)$:

$$
\begin{aligned}
\int x(1+x)(1-x)\,dx &= -\frac{1}{2}\int u\,du \\
&= -\frac{u^2}{4} + C \\
&= -\frac{(1+x)^2(1-x)^2}{4} + C
\end{aligned}
$$

Using the result of the integral to find the value of c:

$$
\begin{aligned}
\int_{-\infty}^{\infty} \rho(x)\,dx &= c\left(-\int_{-1}^{0} x(1+x)(1-x)\,dx + \int_{0}^{1} x(1+x)(1-x)\,dx\right) \\
&= c\left(-\left[-\frac{(1+x)^2(1-x)^2}{4}\right]_{-1}^{0} + \left[-\frac{(1+x)^2(1-x)^2}{4}\right]_{0}^{1}\right) \\
&= c\left(\frac{1}{4} + \frac{1}{4}\right) \\
&= \frac{c}{2}
\end{aligned}
$$

Equating this result to 1, gives the value of c to be 2.

b) Consider a random variable X with the just computed density ρ. Find its CDF and plot it.

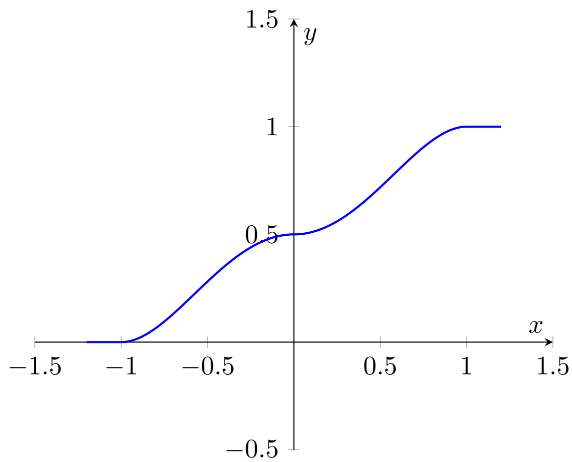
To find the CDF of the random variable X, one can use the property that the derivative of the CDF is the density function. Combining this with the boundaries of the range of the CDF, the CDF can be calculated in the following way:

$$
F_X(x) = \begin{cases} 0 & x < -1 \\ \int_{-\infty}^{x} \rho(x)\,dx & -1 \le x \le 1 \\ 1 & x > 1 \end{cases}
$$

$$\begin{aligned}
\int \rho(x)\,dx &= \int 2|x|(1+x)(1-x)\,dx \\
&= 2\frac{x}{|x|}\int \frac{|x|}{x}|x|(1+x)(1-x)\,dx \\
&= 2\frac{x}{|x|}\int x(1+x)(1-x)\,dx \\
&= -2\frac{x}{|x|}\frac{x^4-2x^2}{4} + C \\
&= -\frac{x(x^4-2x^2)}{2|x|} + C
\end{aligned}$$

Using the boundary definitions, the value of the constant is $\frac{1}{2}$. We arrive at the definition of the CDF and its plot:

$$F_X(x) = \begin{cases} 0 & x < -1 \\ -\frac{x(x^4-2x^2)}{2|x|} + \frac{1}{2} & -1 \le x \le 1 \\ 1 & x > 1 \end{cases}$$



c) Use the CDF to compute

$$P(X < -0.5), \quad P(X > 0.5), \quad P(-0.5 < X < 0.5)$$

$$P(X < -0.5) = F(-0.5) - P(X = -0.5) = F(-0.5) = \frac{9}{32} = 0.28125$$

$$P(X > 0.5) = 1 - P(X <= 0.5) = 1 - F(0.5) = 1 - \frac{23}{32} = 1 - 0.71875 = 0.28125$$

$$P(-0.5 < X < 0.5) = F(0.5) - P(X = 0.5) - F(-0.5) = F(0.5) - F(-0.5) = 0.71875 - 0.28125 - 0.4375$$

Programming Problem 1.1

In this first programming exercise, we would like to familiarize us with typical linear algebra computing tasks. To this end, we define the vectors / matrices

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}, \quad M = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

a) Compute and print the inner product between **a** and **b**.

```
np.dot(a.T, b)
```

b) Compute and print the matrix-vector product between M and **b**.

```
M.dot(b)
```

c) Compute and print the l_2 norm of **b**.

```
np.linalg.norm(b)
```

d) Compute and print the solution **x** of the linear system of equations $M\mathbf{x} = \mathbf{b}$ using the LU factorization.

```
np.linalg.solve(M, b)
```

e) Compute and print the solution **x** of the linear system of equations $M\mathbf{x} = \mathbf{b}$ using the Cholesky factorization.

```
L = np.linalg.cholesky(M)
y = np.linalg.solve(L, b)
x = np.linalg.solve(L.T.conj(), y)
```

The solutions to these exercises are in programming_exercises.py.