

Statistical Foundations of Data Science
Past Homework Distributions

The content of this assignment is based on the lectures on "How to lie with statistics" and "Combinatorics"

1. Determine which of the following are distributions. For those which are not distributions explain why (you shouldn't need more than a sentence for the explanation)

(a)

$$p(X) = \begin{cases} \frac{1}{3} & x \in \{1, 6, 20\} \\ 0 & elsewhere \end{cases}$$

[2]

(b)

$$p(X) = \begin{cases} \frac{x^2}{10} & x \in \{-2, -1, 1, 2\} \\ 0 & elsewhere \end{cases}$$

[2]

(c)

$$p(X) = \begin{cases} \frac{3}{5} & x = 2 \\ \frac{3}{5} & x = 3 \\ \frac{-1}{5} & x = 4 \\ 0 & elsewhere \end{cases}$$

[2]

(d)

$$p(X) = \begin{cases} \frac{3}{4^x} & x \in \{1, 2, 3, \dots\} \\ 0 & elsewhere \end{cases}$$

[2]

(e)

$$p(X) = \begin{cases} \frac{3}{4^x} & x \in \{0, 1, 2, 3, \dots\} \\ 0 & elsewhere \end{cases}$$

[2]

2. One of the two functions given below is a probability density function (i.e. a distribution).

A

$$f(x) = \begin{cases} 3e^{-3x} & x > 0 \\ 0 & elsewhere \end{cases}$$

B

$$f(x) = \begin{cases} 3e^{-4x} & x > 0 \\ 0 & elsewhere \end{cases}$$

- (a) Determine which one is a distribution and state why the other one isn't. [3]
 - (b) Compute the mean of the distribution[3]
 - (c) Compute the variance of the distribution[4]
3. Let X be distributed continuous uniform between 0 and 10. Compute
- a $\mathbb{P}(X > 7)$ [2]
 - b $\mathbb{P}(X^2 > 49)$ [3]
 - c $\mathbb{E}[X]$ [2]
 - d $V(X)$ [3]

4. Consider the following distribution

$$p(X) = \begin{cases} \frac{1}{5} & x \in \{1, 2, 3\} \\ c & x = 4 \\ 0 & elsewhere \end{cases}$$

- a Find c [3]
 - b Compute $\mathbb{E}[X]$ [3]
 - c Compute $V(X)$ [4]
5. A fair coins is tossed 12 times and the number of heads is recorded as X . Compute:
- a Compute $\mathbb{E}[X]$ [2]
 - b Compute $V(X)$ [4]
 - c $\mathbb{P}(X \geq 10)$ [4]