2. Random variable and probability distribution

Discrete random variable

Probability mass function f_X of a discrete random variable X is defined by:

$$f_X(x_i) \equiv P(X = x_i) = p(x_i), \quad S_X = \{x_i\}$$

and has the following properties: $0 \leq f_X(x_i) \leq 1$ in $\sum_{x_i \in S_X} f_X(x_i) = 1$.

Cumulative distribution function F_X of a discrete random variable X is defined by:

$$F_X(x_i) \equiv P(X \leq x_i) = \sum_{x_j \leq x_i} f_X(x_j)$$

and has the following properties:

$$\lim_{x_i \to -\infty} F_X(x_i) = \lim_{x_i \to -\infty} P(X \leq x_i) = 0,$$

$$\lim_{x_i \to +\infty} F_X(x_i) = \lim_{x_i \to +\infty} P(X \leq x_i) = 1,$$

$$x_i \geq x_j \Rightarrow F_X(x_i) \geq F_X(x_j).$$

Probability of an interval: $P(a \leq X \leq b) = \sum_{x_i \in [a,b]} f_X(x_i) = F_X(b) - F_X(a).$

Continuous random variable

Cumulative distribution function F_X of a continuous random variable X is defined by:

$$F_X(x) \equiv P(X \leq x) \quad \text{za} \quad -\infty < x < \infty$$

and has the following properties:

$$\lim_{x \to -\infty} F_X(x) = \lim_{x \to -\infty} P(X \leq x) = 0,$$

$$\lim_{x \to +\infty} F_X(x) = \lim_{x \to +\infty} P(X \leq x) = 1,$$

$$x_2 \geq x_1 \Rightarrow F_X(x_2) \geq F_X(x_1).$$

Probability density function f_X of a continuous random variable X is defined by:

$$f_X(x) \equiv \lim_{\Delta x \to 0} \frac{\Delta P}{\Delta x} = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta F_X(x)}{\Delta x} = \frac{dF_X(x)}{dx} = F_X'(x)$$

and has the following properties: $f_X(x) \geq 0$ in $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.

Probability of an interval: $P(a \leq X \leq b) = \int_a^b f_X(x)\, dx = F_X(b) - F_X(a).$

Useful connections: $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)\, du \quad \text{za} \quad -\infty < x < \infty,$

$$\frac{d}{dx} F_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f_X(u)\, du = f_X(x).$$

Mean and variance

	X discrete	X continuous
Mean $m_X = \mathrm{E}[X]$	$m_X = \sum_{x_i \in S_X} x_i\, f_X(x_i)$	$m_X = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$
Variance $\mathrm{Var}[X] = \sigma_X^2 = \mathrm{E}[(X - m_X)^2]$	$\sigma_X^2 = \sum_{x_i \in S_X} (x_i - m_X)^2\, f_X(x_i)$	$\sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2\, f_X(x)\, dx$

2. Random variable and probability distribution – problems

1. The sample space of a random experiment is $\{a, b, c, d, e, f\}$, and each outcome is equally likely. A random variable X is defined as follows:

elementary event (=outcome)	a	b	c	d	e	f
x_i	0	0	1,5	1,5	2	3

   Determine the probability mass function f_X. Use it to determine the following probabilities: a) $P(X = 1{,}5)$, b) $P(0{,}5 < X < 2{,}7)$, c) $P(X > 3)$, d) $P(0 \leq X < 2)$, e) $P(X = 0 \text{ or } X = 2)$. Determine the corresponding cumulative probability function.
   R: $f(0) = 2/6$; $f(1{,}5) = 2/6$; $f(2) = 1/6$; $f(3) = 1/6$; a) $2/6$, b) $3/6$, c) 0, d) $4/6$, e) $3/6$

2. Check that the tabulated function is a probability mass function and determine the following probabilities.

x_i	-2	-1	0	1	2
$f_X(x_i)$	1/8	2/8	2/8	2/8	1/8

   a) $P(X \leq 2)$, b) $P(X > -2)$, c) $P(-1 \leq X \leq 1)$, d) $P(X \leq 1 \text{ or } X = 2)$. Determine the corresponding cumulative probability function.
   R: It is. a) 1, b) 7/8, c) 6/8, d) 1; $F(X < -2) = 0$; $F(-2 \leq X < -1) = 1/8$; $F(-1 \leq X < 0) = 3/8$; $F(0 \leq X < 1) = 5/8$; $F(1 \leq X < 2) = 7/8$; $F(2 \leq X) = 1$

3. In a semiconductor manufacturing process, three wafers from a lot are tested. Each wafer is classified as pass or fail. Assume that the probability that a wafer passes the test is 0.8 and that wafers are independent. Determine the probability mass function of the number of wafers from a lot that pass the test. Determine the corresponding cumulative probability function.
   R: $f(0) = 0{,}008$; $f(1) = 0{,}096$; $f(2) = 0{,}384$; $f(3) = 0{,}512$; $F(X < 0) = 0$; $F(0 \leq X < 1) = 0{,}008$; $F(1 \leq X < 2) = 0{,}104$; $F(2 \leq X < 3) = 0{,}488$; $F(3 \leq X) = 1$

4. The probability density function of the time to failure of an electronic component in a copier (in hours) is $f_X(x) = \mathrm{e}^{-\frac{x}{1000}}/1000$ for $x > 0$. Determine the corresponding cumulative probability function. Determine the probability that:
   a. A component lasts more than 3000 hours before failure.
   b. A component fails in the interval from 1000 to 2000 hours.
   c. A component fails before 1000 hours.
   d. Determine the number of hours at which 10 % of all components have failed.

   R: $F_X(x) = 1 - \mathrm{e}^{-\frac{x}{1000}}$; a) 0,0498, b) 0,233, c) 0,632, d) 105

5. Determine the mean and variance of random variables in Exercises 1 and 2.

6. Determine the probability density functions that correspond to the following cumulative probability functions:
   a. $F_X(x) = 1 - \mathrm{e}^{-2x}$ for $x > 0$. R: $f_X(x) = 2\mathrm{e}^{-2x}$
   b. $F_X(x) = \begin{cases} 0; & x < 0 \\ 0{,}2x; & 0 \leq x < 4 \\ 0{,}04x + 0{,}64; & 4 \leq x < 9 \\ 1; & 9 \leq x \end{cases}$ R: $f_X(x) = \begin{cases} 0; & x < 0 \\ 0{,}2; & 0 \leq x < 4 \\ 0{,}04; & 4 \leq x < 9 \\ 0; & 9 \leq x \end{cases}$