Chapter IV

DISCRETE RANDOM VARIABLES

1 Definitions

A random variable (*variabliă aleatoare*) consists in an experiment with a probability measure P defined on a sample space S and a function that assigns a real number to each outcome in a sample space of the experiment.

We denote $\{X = x\} = \{s \in S, X(s) = x\}$.

Examples 1. Let A be the random variable that counts the number of the students asleep in the nest probability lecture.
2. Let B be the random variable that counts the number of phone calls you answer in the next hour.
3. Let C be the random variable that counts the number of minutes you wait until you next answer the phone.

X is **a discrete random variable** if the range of X is a countable set:

$$S_X = \{x_1, x_2, ...\}.$$

X is a **finite random variable** if the range of X is a finite set:

$$S_X = \{x_1, x_2, ..., x_n\}.$$

Examples: 1. The random variable A and B are discrete random variables, M is a continuous random variable.
2. We observe three calls at a telephone switch where voice calls (v) and data calls (d) are equally likely. Let X be the random variable that counts the number of voice calls.

Outcomes:	ddd	ddv	dvd	vdd	vvd	vdv	dvv	vvv
X :	0	1	1	1	2	2	2	3
P :	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

The random variable X is given by the table:

$$X : \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}.$$

2 Probability Mass Function (PMF). Cumulative Distribution Function (CDF)

The probability mass function (PMF) (*distribuţia variabile aleatoare*) of a discrete r.v. X is

$$P_X(x) = P(X = x).$$

Example: PMF of the r.v. X from the above Example is

$$X : \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix},$$

or:

$$P_X(x) = \begin{cases} \dfrac{1}{8}, \; x = 0 \\[2ex] \dfrac{3}{8}, \; x = 1 \\[2ex] \dfrac{3}{8}, \; x = 2 \\[2ex] \dfrac{1}{8}, \; x = 3 \\[2ex] 0, \; otherwise \end{cases}$$

Proposition 2.1. *If X is a discrete r.v. S_X its range, and $P_X(x)$ its PFM, then:*
a) $P_X(x) \geq 0, \;\; \forall x \in S_X$;

b) $\sum_{x \in S_X} P_X(x) = 1$;

c) $P(A) = \sum_{x \in A} P_X(x),\ \forall A \subset S_X$.

The cumulative distribution function (CDF) (*funcţia de repartiţie*) of the r.v. X is

$$F_X(x) = P(X \leq x).$$

Proposition 2.2. *For any discrete r.v.* X, *with the range* $S_X = \{x_1, x_2, ...\}$ *such that* $x_1 < x_2 < ...$, *the following properties hold:*

a) $\lim_{x \to \infty} F_X(x) = 1,\ \lim_{x \to -\infty} F_X(x) = 0$;

b) $\forall x_1 \leq x_2 \Rightarrow F_X(x_1) \leq F_X(x_2)$ *(never decreases)*

c) $F_X(x_i) - F_X(x_i - 0) = P_X(x_i),\ \forall x_i \in S_x$ *(F is continuous from the right and there is a jump at each value of* $x_i \in S_X$*);*

d) $F_X(x) = F_X(x_i),\quad \forall x,\ x_i \leq x < x_{i+i}$ *(between jumps, the graph is an horizontal line);*

e) $F_X(b) - F_X(a) = P(a < X \leq b)$.

Example: For the previous Example, the corresponding CDF is:

$$F_X(x) = \begin{cases} 0,\ x \leq 0 \\ \frac{1}{8},\ x \in (0,1] \\ \frac{1}{8} + \frac{3}{8},\ x \in (1,2] \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8},\ x \in (2,3] \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8},\ x \in (3,\infty) \end{cases} = \begin{cases} 0,\ x \leq 0 \\ \frac{1}{8},\ x \in (0,1] \\ \frac{1}{2},\ x \in (1,2] \\ \frac{7}{8},\ x \in (2,3] \\ 1,\ x \in (3,\infty) \end{cases}$$

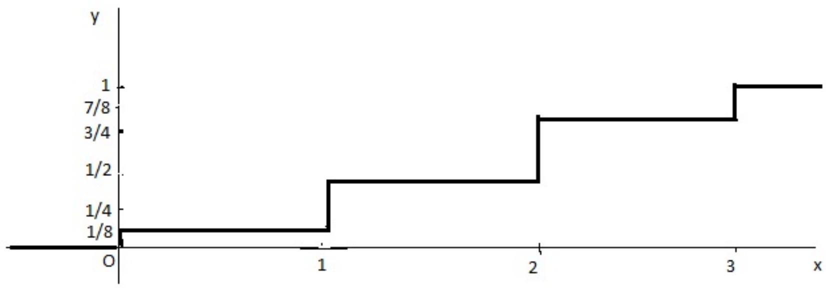

Figure 1

3 Families of Discrete Random Variables

X **is a Bernoulli r.v.** if PMF of X has the form:

$$P_X(x) = \begin{cases} 1-p, \ x = 0 \\ \\ p, \ x = 1 \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p < 1.$$

Example Suppose you test a circuit. With probability p, the circuit is rejected. Let X be the r.v. that counts the number of rejected circuits in one test. Then, we have only two outcomes in the sample space 0 and 1, so the PMF is:

$$P_X(x) = \begin{cases} 1-p, \ x = 0 \\ \\ p, \ x = 1 \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p < 1.$$

X **is a geometric** (p) **r.v.** if the PMF of X has the form:

$$P_X(x) = \begin{cases} p(1-p)^{x-1}, \ x = 1, 2, ... \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p < 1.$$

Example In a test of integrated circuits there is a probability p that each circuit is rejected. Let X be the r.v. that counts the number of tests up to and including the first test that discovers a reject. What is the PMF of X?

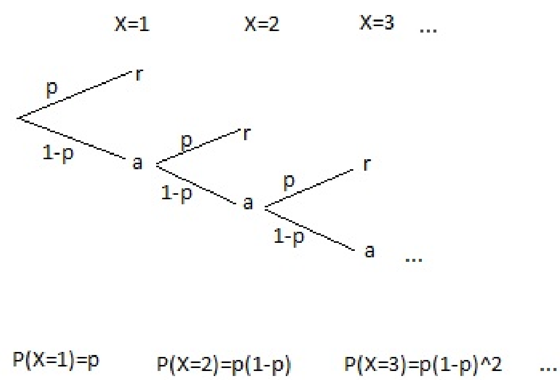

Figure 2

so

$$P_X(x) = \begin{cases} p(1-p)^{x-1}, \ x = 1, 2, ... \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p < 1.$$

X **is a binomial** (n, p) **r.v.** if the PMF has the form

$$P_X(x) = C_n^x \cdot p^x \cdot (1-p)^x, \ \ 0 < p < 1, n \geq 1, n \in \mathbb{N}$$

Example Suppose we test n circuits and each circuit is rejected with probability p independent of the results of other tests. Let k be the number of rejects in n tests. Then

$$P_k(x) = C_n^x \cdot p^x \cdot (1-p)^{n-x}.$$

X **is a Pascal** (k, p) **r.v.** if the PMF of X has the form

$$P_X(x) = C_{x-1}^{k-1} p^k (1-p)^{x-k}, \ \ 0 < p < 1, k \in \mathbb{N}^*.$$

Example Suppose you test circuits until you find k rejects. Let X be the r.v. that counts the number of tests. Then the PMF of X is

$$P_X(x) = C_{x-1}^{k-1} p^k (1-p)^{x-k}, \ \ 0 < p < 1, k \in \mathbb{N}^*.$$

X **is a discrete uniform** (k, l) **r.v.** if the PMF OF X has the form

$$P_X(x) = \begin{cases} \dfrac{1}{l-k+1}, \ x = k, k+1, k+2, ..., l \\ \\ 0, \ otherwise \end{cases}, \ \ k, l \in \mathbb{Z}, k < l.$$

Example Roll a fair die. Let X be the r.v. that provides the number of the spots that appears on the side facing up. Then, the PMF of X is:

$$P_X(x) = \begin{cases} \dfrac{1}{6}, \ x = 1, 2, ..., 6 \\ \\ 0, \ otherwise \end{cases}$$

X is a Poisson (α) **r.v.** if the PMF of X has the form

$$P_X(x) = \begin{cases} \dfrac{\alpha^x \cdot e^{-\alpha}}{x!}, \ x = 0, 1, 2, ... \\ \\ 0, \ otherwise \end{cases}, \ \ \alpha > 0.$$

Remark 3.1. To describe a Poisson r.v., we will call the occurrence of the phenomenon of interest **an arrival**. A Poisson model often specifies **an average rate** λ arrivals per second, and a time interval T seconds. In this time interval, the number of arrivals X has a Poisson PMF with $\alpha = \lambda T$.

Example The number of hits at a Web site in any time interval is a Poisson r.v. A particular site has an average $\lambda = 2$ hits per second.

a) What is the probability that there are no hits in an interval of 0.25 seconds?
b) What is the probability that there are no more than 2 hits in an interval

of one second?

Solution a) In an interval of 0.25 seconds, the number of hits X is a Poisson r.v with $\alpha = \lambda \cdot T = 2$ hits/sec$\cdot$0.25sec= 0.5hits, so:

$$P_X(x) = \begin{cases} \dfrac{0.5^x \cdot e^{-0.5}}{x!}, \ x = 0, 1, 2, ... \\ \\ 0, \ otherwise \end{cases}$$

The probability of the number of hits is:

$$P(X = 0) = P_X(0) = \frac{1 \cdot e^{-0.5}}{0!} = e^{-0.5}.$$

b) In an interval of 1 seconds: $\alpha = \lambda \cdot T = 2$ hits/sec$\cdot$1sec= 2hits, so:

$$P_Y(x) = \begin{cases} \dfrac{2^x \cdot e^{-2}}{x!}, \ x = 0, 1, 2, ... \\ \\ 0, \ otherwise \end{cases}$$

Therefore

$$\begin{aligned} P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(L = 2) \\ &= P_Y(0) + P_Y(1) + P_Y(2) \\ &= e^{-2} + 2\dot{e}^{-2} + \frac{2^2 \cdot e^{-2}}{2} = 0.677 \end{aligned}$$