Probability and Statistics

Lect. dr. Raluca Muresan

West University of Timișoara
Faculty of Mathematics and Computer Science

Table of contents

Two-dimensional discrete random variables (random vectors)

Sometimes, it is necessary to consider two or more random variables defined on the same sample space, at the same time. In the followings, we will present the case of two random variables; the generalization to three or more variables can be achieved with no difficulty.

Two-dimensional discrete random variables (random vectors)

Sometimes, it is necessary to consider two or more random variables defined on the same sample space, at the same time. In the followings, we will present the case of two random variables; the generalization to three or more variables can be achieved with no difficulty.

Example. We consider the experiment consisting of distributing three balls a, b, c in three boxes. Let X be the random variable which represents the number of non-empty boxes:

$$X : \begin{pmatrix} 1 & 2 & 3 \\ \frac{3}{27} & \frac{18}{27} & \frac{6}{27} \end{pmatrix}.$$

Two-dimensional discrete random variables (random vectors)

Sometimes, it is necessary to consider two or more random variables defined on the same sample space, at the same time. In the followings, we will present the case of two random variables; the generalization to three or more variables can be achieved with no difficulty.

Example. We consider the experiment consisting of distributing three balls a, b, c in three boxes. Let X be the random variable which represents the number of non-empty boxes:

$$X : \begin{pmatrix} 1 & 2 & 3 \\ \frac{3}{27} & \frac{18}{27} & \frac{6}{27} \end{pmatrix}.$$

Let Y be the random variable representing the number of balls from the first box:

$$Y : \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{8}{27} & \frac{12}{27} & \frac{6}{27} & \frac{1}{27} \end{pmatrix}.$$

Two-dimensional discrete random variables (random vectors)

Sometimes, it is necessary to consider two or more random variables defined on the same sample space, at the same time. In the followings, we will present the case of two random variables; the generalization to three or more variables can be achieved with no difficulty.

Example. We consider the experiment consisting of distributing three balls a, b, c in three boxes. Let X be the random variable which represents the number of non-empty boxes:

$$X : \begin{pmatrix} 1 & 2 & 3 \\ \frac{3}{27} & \frac{18}{27} & \frac{6}{27} \end{pmatrix}.$$

Let Y be the random variable representing the number of balls from the first box:

$$Y : \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{8}{27} & \frac{12}{27} & \frac{6}{27} & \frac{1}{27} \end{pmatrix}.$$

Two-dimensional discrete random variables (random vectors)

We now consider the random variable Z which represents the pair of numbers (number of non-empty boxes, number of balls from the first box):

$$Z : \begin{pmatrix} (1,0) & (1,3) & (2,0) & (2,1) & (2,2) & (3,1) \\ \frac{2}{27} & \frac{1}{27} & \frac{6}{27} & \frac{6}{27} & \frac{6}{27} & \frac{6}{27} \end{pmatrix}.$$

Two-dimensional discrete random variables

In general, let's consider two random variables X, Y defined on the same sample space $\Omega = \{e_1, e_2, \ldots, e_n\}$. Let $x_1, x_2, ..., x_k$ be the values of the random variable X and $y_1, y_2, ..., y_l$ the values of the random variable Y .

Definition

Using the random variables X, Y we can build up the two-dimensional random vector $Z = (X, Y)$, whose values are the ordered pairs $(x_i; y_j)$ (two dimensional vectors), and the corresponding probabilities are

$$r_{ij} = P(X = x_i \text{ and } Y = y_j),\ 1 \leq i \leq k,\ 1 \leq j \leq l.$$

The probability distribution of Z is given by the following table:

$X \setminus Y$	y_1	y_2	y_3	...	y_j	...	y_l	$P(X = x_i)$
x_1	r_{11}	r_{12}	r_{13}	...	r_{1j}	...	r_{1l}	p_1
x_2	r_{21}	r_{22}	r_{23}	...	r_{2j}	...	r_{2l}	p_2
x_3	r_{31}	r_{32}	r_{33}	...	r_{3j}	...	r_{3l}	p_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	r_{i1}	r_{i2}	r_{i3}	...	r_{ij}	...	r_{il}	p_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_k	r_{k1}	r_{k2}	r_{k3}	...	r_{kj}	...	r_{kl}	p_k
$P(Y = y_j)$	q_1	q_2	q_3	...	q_j	...	q_k	1

Two-dimensional discrete random variables

As the events $(X = x_i, Y = y_j)$ form a partition of the sample space , the sum of the probabilities from this table must be equal to 1:

$$\sum_{i=1}^{k} \sum_{j=1}^{l} r_{ij} = 1.$$

The first column (line) together with the last column (line) of the table form the marginal probability distribution of the random variable X (Y).

Two-dimensional discrete random variables

Definition

We call distribution function of the random vector (X, Y) the function defined by

$$F(x,y) = P(X \leq x \text{ and } Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P(X = x_i \text{ and } Y = y_j) =$$

$$= \sum_{x_i \leq x} \sum_{y_j \leq y} r_{ij}.$$

Two-dimensional discrete random variables

Proposition

The distribution function of the random vector (X, Y) satisfies the following properties:

1. $F(x_1, y) \leq F(x_2, y)$ *if* $x_1 < x_2$;
   $F(x, y_1) \leq F(x, y_2)$ *if* $y_1 < y_2$.
2. $F(x, -\infty) = F(-\infty, y) = 0$ *and* $F(x, \infty) = F(\infty, y) = 1$.
3. $F(x, \infty)$ *is the distribution function of the random variable* X,
   $F(\infty, y)$ *is the distribution function of the random variable* Y.

Definition

We say that the random variables X and Y are independent if for every pair (i, j) we have

$$r_{ij} = p_i \cdot q_j.$$

Two-dimensional discrete random variables

Definition

We say that the random variables X and Y are independent if for every pair (i,j) we have

$$r_{ij} = p_i \cdot q_j.$$

Proposition

If the random variables X, Y are independent, then:

1. *the conditional distributions are the same as the marginal distributions:*

$$P(x_i|y_j) = \frac{r_{ij}}{q_j} = \frac{p_i \cdot q_j}{q_j} = p_i,$$

$$P(y_j|x_i) = \frac{r_{ij}}{p_i} = \frac{p_i \cdot q_j}{p_i} = q_j.$$

2. $F(x,y) = \sum_{x_i \le x} \sum_{y_j \le y} r_{ij} = \sum_{x_i \le x} \sum_{y_j \le y} p_i \cdot q_j = \left(\sum_{x_i \le x} p_i\right) \left(\sum_{y_j \le y} q_j\right) = F(x,\infty)F(\infty,y).$

Covariance. Correlation coefficient

Definition

If X and Y are random variables defined on the same sample space Ω, we call **covariance** *of the variables X and Y, the number*

$$Cov(X, Y) = E([X - E(X)][Y - E(Y)]).$$

Covariance. Correlation coefficient

Definition

If X and Y are random variables defined on the same sample space Ω, we call **covariance** *of the variables X and Y, the number*

$$Cov(X, Y) = E([X - E(X)][Y - E(Y)]).$$

Proposition

The following equality takes place:

$$Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y).$$

Covariance. Correlation coefficient

Definition

If X and Y are random variables defined on the same sample space Ω, we call **covariance** *of the variables X and Y, the number*

$$Cov(X, Y) = E([X - E(X)][Y - E(Y)]).$$

Proposition

The following equality takes place:

$$Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y).$$

Proposition

If X and Y are independent random variables, then:

$$Cov(X, Y) = 0.$$

Covariance. Correlation coefficient

Definition

If X and Y are two variables defined on the same sample space Ω, we call **correlation coefficient** *of the variables X and Y, the number*

$$\rho(X, Y) = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{V(X) \cdot V(Y)}} = \frac{Cov(X, Y)}{\sigma(x) \cdot \sigma(Y)}.$$

Convergence of sequences of random variables.

We consider a sequence of random variables $X_1, X_2, \ldots, X_n, \ldots$ defined on the same sample space Ω.
In the probability theory we can find different concepts of convergence for the sequences of random variables $(X_n)_n$.

Definition

We say that the sequence of random variables (X_n) **converges surely or everywhere** *towards* X *if*

$$\lim_{n\to\infty} X_n(e) = X(e), \ \forall e \in \Omega.$$

Convergence of sequences of random variables.

We consider a sequence of random variables $X_1, X_2, \ldots, X_n, \ldots$ defined on the same sample space Ω.
In the probability theory we can find different concepts of convergence for the sequences of random variables $(X_n)_n$.

Definition

We say that the sequence of random variables (X_n) **converges surely or everywhere** *towards X if*

$$\lim_{n\to\infty} X_n(e) = X(e),\ \forall e \in \Omega.$$

Definition

We say that the sequence of random variables (X_n) converges towards X **in probability**, *if*

$$\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1,\ \forall \varepsilon > 0.$$

Convergence of sequences of random variables.

Definition

We say that the sequence of random variables (X_n) *converges* **almost surely** *towards the random variable* X, *if*

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

Convergence of sequences of random variables.

Definition

We say that the sequence of random variables (X_n) converges **almost surely** *towards the random variable X, if*

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

Definition

Let $F_n(x)$ be the distribution function of the variable X_n, $(n = 1, 2, \ldots)$ and $F(x)$ the distribution function of the variable X. The sequence X_n converges towards X **in distribution** *if*

$$\lim_{n\to\infty} F_n(x) = F(x).$$

Convergence of sequences of random variables.

Definition

If

$$\lim_{n\to\infty} D^2(X_n - X) = 0$$

we say that the sequence X_n **converges in mean square** *to* X.

Convergence of sequences of random variables.

Definition

If

$$\lim_{n\to\infty} D^2(X_n - X) = 0$$

we say that the sequence X_n **converges in mean square** *to X.*

Proposition

If the sequence X_n converges in mean square to X, then X_n converges to X in probability.

Convergence of sequences of random variables.

Definition

If

$$\lim_{n\to\infty} D^2(X_n - X) = 0$$

we say that the sequence X_n **converges in mean square** *to X.*

Proposition

If the sequence X_n converges in mean square to X, then X_n converges to X in probability.

Proposition

If the sequence Xn converges almost surely to X, then X_n converges to X in probability.

Law of large numbers

Theorem

(Chebyshev)
Let (X_n) *be a sequence of random variables defined on a sample space* Ω. *If the random variables are independent and* $V(X_n) \le c$, $\forall n$, *then for all* $\varepsilon > 0$ *we have*

$$\lim_{n\to\infty} P(|\overline{X}_n - E(\overline{X}_n)| < \varepsilon) = 1,$$

where $\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$.

Law of large numbers

Theorem

(Chebyshev)
Let (X_n) *be a sequence of random variables defined on a sample space* Ω*. If the random variables are independent and* $V(X_n) \leq c$*,* $\forall n$*, then for all* $\varepsilon > 0$ *we have*

$$\lim_{n\to\infty} P(|\overline{X}_n - E(\overline{X}_n)| < \varepsilon) = 1,$$

where $\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$*.*

Chebyshev's Theorem shows that even if the independent random variables can take values far away from their expected values, the arithmetic mean of a sufficiently large number of such random variables takes, with a large probability, values in the neighborhood of the constant $\frac{1}{n}\sum_{j=1}^{n} E(X_j)$.

Law of large numbers

So, there is a big difference between the behavior of the random variables and their arithmetic mean. In the case of the random variables we cannot predict their value with a large probability, while, in the case of their arithmetic mean we can give its value with a probability close to 1.

Law of large numbers

So, there is a big difference between the behavior of the random variables and their arithmetic mean. In the case of the random variables we cannot predict their value with a large probability, while, in the case of their arithmetic mean we can give its value with a probability close to 1.

Theorem

(Bernoulli)
We suppose we make n independent experiences, in each experience the probability of the event A being p, and let ν be the number of times the event A is accomplished during the n experiences. For each ε we have

$$\lim_{n\to\infty} P\left(\left|\frac{\nu}{n} - p\right| < \varepsilon\right) = 1.$$