CS 107, Probability, Spring 2020
Lecture 29

Michael Poghosyan
mpoghosyan@aua.am

AUA

15 April 2020

Content

- Identically Distributed r.v.s
- Joint Distribution
  - Joint Distribution of any r.v.s by their Joint CDF
  - Joint Distribution of Discrete r.v.s by their Joint PMF

Identically Distributed r.v.s

ID r.v.s

We already met different r.v.s with the same distribution: if X and Y have the same distribution, say, if

$$X, Y \sim \mathit{Unif}([0,1]),$$

that doesn't mean that $X \equiv Y$, i.e., that doesn't mean that $X(\omega) = Y(\omega)$ for any $\omega \in \Omega$. Of course, if $X(\omega) = Y(\omega)$ for any ω, then we will have that X and Y have the same distribution, but the inverse is not true, in general. This is because, when we talk about the r.v.s, then

- we are interested in the **values of** that r.v.;
- we are **not concerned at which point** they are taking that values;
- we are **concerned with which probability** they are taking that values.

ID r.v.s, Example

Example: Let us give examples of r.v.s $X \neq Y$ with the same distribution, say, $X, Y \sim Bernoulli(0.5)$.
We will consider two types of examples:

- Type 1 example, where X and Y are defined on different sample spaces;
- Type 2 example, where X and Y are defined on the same sample space.

ID r.v.s, Example

Type 1 Example: X and Y can be defined on very different experiments.

- Say, we are tossing a fair coin, $X = 0$ if Heads appears and $X = 1$ otherwise. So here $\Omega = \{H, T\}$ and
$$X(H) = 0, \quad X(T) = 1.$$
Also, $\mathbb{P}(X = 0) = 0.5$ and $\mathbb{P}(X = 1) = 0.5$.
- We are picking a number from $\{1, 2, 3, 4\}$ randomly, uniformly (equiprobably). Let $Y = 0$, if the number chosen is less than 3, and $Y = 1$ otherwise. That is, $\Omega = \{1, 2, 3, 4\}$, and
$$Y(1) = 0, \quad Y(2) = 0, \quad Y(3) = 1, \quad \text{and} \quad Y(4) = 1.$$
So $\mathbb{P}(Y = 0) = 0.5$ and $\mathbb{P}(Y = 1) = 0.5$.

ID r.v.s, Example, Type 1 Example, cont'd

Then, clearly, $X, Y \sim Bernoulli(0.5)$, but we cannot even check if $X(\omega) = Y(\omega)$, since this r.v.s (functions) are defined on different sets, different sample spaces. So X and Y are different r.v.s with the same distribution.

ID r.v.s, Example

Type 2 Example: X and Y can be defined on the same experiment. Say, our experiment is rolling a fair die, and we have $\Omega = \{1, 2, 3, 4, 5, 6\}$. We consider the following two r.v.s:

- $X = 0$ if the number shown is Even, and $X = 1$ otherwise. So here

  $$X(2) = X(4) = X(6) = 0 \quad \text{and} \quad X(1) = X(3) = X(5) = 1,$$

  and $\mathbb{P}(X = 0) = 0.5$ and $\mathbb{P}(X = 1) = 0.5$.
- $Y = 0$, if the number chosen is prime, and $Y = 1$ otherwise. That is,

  $$Y(2) = Y(3) = Y(5) = 0 \quad \text{and} \quad Y(1) = Y(4) = Y(6) = 1,$$

  and again, $\mathbb{P}(Y = 0) = 0.5$ and $\mathbb{P}(Y = 1) = 0.5$.

Then, $X, Y \sim Bernoulli(0.5)$, but we do not have $X(\omega) = Y(\omega)$ for any ω: say, $X(4) \neq Y(4)$. So X and Y are different r.v.s (on the same sample space) with the same distribution.

Identically Distributed r.v.s

Now, we give a definition:

Identically Distributed r.v.s

We will say that X and Y are ID (Identically Distributed), if they have the same CDFs, i.e., if

$$F_X(x) = F_Y(x), \qquad \forall x \in \mathbb{R},$$

where F_X and F_Y are the corresponding CDFs.

It can be seen that

- If X and Y are Discrete, then X and Y are ID iff they share the same PMF;
- If X and Y are Continuous, then X and Y are ID iff

$$f_X(x) = f_Y(x) \qquad \textit{for almost all } x \in \mathbb{R}$$

ID r.v.

Example: When we are writing $X, Y \sim \mathcal{N}(0,1)$, this means that X and Y have the same distribution, Standard Normal, so they are ID.

Example: Assume

$$X, Y \sim \begin{pmatrix} 0 & 3 & 7 \\ 0.3 & 0.5 & 0.2 \end{pmatrix}.$$

Then X and Y are identically distributed, they are ID.

Note: Now, assume X and Y are ID r.v.s. Then, for any $A \subset \mathbb{R}$, we will have

$$\mathbb{P}(X \in A) = \mathbb{P}(Y \in A).$$

Note: For the above examples, X and Y can be defined on different sample spaces.

Examples of ID r.v.

Example: Assume $X \sim \mathit{Unif}([0,1])$ and $Y = 1 - X$. Then, it is easy to prove that $Y \sim \mathit{Unif}([0,1])$, so X and Y are ID, although $X \not\equiv Y$.

Example: Assume $X \sim \mathcal{N}(0,1)$ and $Y = -X$. Again, it is easy to prove that $Y \sim \mathcal{N}(0,1)$, so X and Y are ID.

Note: For the above examples, X and Y are defined on the same sample space.

Joint Distribution of R.V.

Joint Distribution of R.V.

Up to this lecture we have considered the Distribution of one r.v. In many cases, we are concerned in the Distribution of 2 or more r.v.s, defined on the same Experiment.

Very important question is how two or more (random) quantities, i.e., r.v. in our terms, are related to each other.

Examples:

Examples:

- How the Height and Weight of a person are related to each other? Here the Experiment is to choose a random person, and r.v.s are the Height and Weight.
- How the AUA Entrance Math Exam grade and the First Year GPA are related? Here the Experiment is to choose an AUA random student, and r.v.s are - well, you can guess.
- How the Education years and Salary are related? Or how the Education years, Probability course grade and Salary are related? Or how the daily FB and GOOG max prices are related? Etc...

During our next lectures, we will talk about the Joint Distribution of several r.v.s, and give methods to describe their Joint Distribution.

Random vectors

We will start by defining and studying the Joint Distribution of 2 r.v.

Assume X and Y are two r.v. defined on the same Experiment (same Probability Space). We will say that (X, Y) is a (2D) random vector defined on that Experiment.

For two (or more) r.v.s the main problems we want to consider are:

- Given a set $A \subset \mathbb{R}^2$, calculate the probability
$$\mathbb{P}\Big((X, Y) \in A\Big);$$
- Describe how X and Y are related to each other.

Describing the Joint Distribution

Recall, that for the 1D case, i.e., for the case of one r.v. X, we were describing the distribution of X through its:

- CDF $F_X(x)$ (this works for any r.v., i.e. discrete, continuous, singular);
- PMF (if X is discrete) or PDF (if X is continuous).

Now, in the analogy, we will describe the Joint Distribution of (X, Y) through their:

- Joint CDF $F_{X,Y}(x, y)$ (this will work for any pair of r.v.s, i.e. both discrete, both continuous, one is discrete and the other one is continuous,...);
- Joint PMF (if X and Y are both discrete) or Joint PDF (if X and Y are both continuous).

Joint CDF of 2 RVs, aka bivariate case

Recall that, for one r.v. X, the CDF of X was defined by

$$F_X(x) = \mathbb{P}(X \leq x), \qquad x \in \mathbb{R}.$$

We are defining the Joint CDF in the similar way:

Joint CDF of 2 r.v.

The **Joint CDF of random variables** X **and** Y or the **CDF of a random vector** (X, Y) is the function

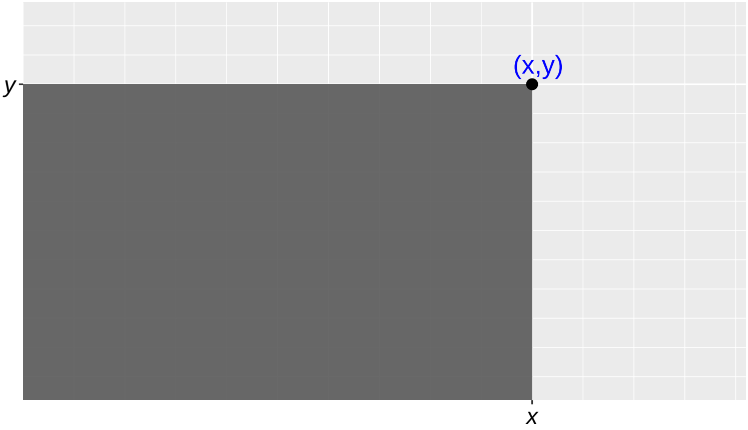
$$F(x, y) = F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y), \qquad \forall (x, y) \in \mathbb{R}^2.$$

Here, when writing $\mathbb{P}(X \leq x, Y \leq y)$ we mean the probability that both $X \leq x$ and $Y \leq y$, i.e.,

$$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x \text{ and } Y \leq y).$$

Joint CDF, geometrically

Geometrically, the Joint CDF of X, Y, $F(x,y)$ at the point (x,y) is showing the probability that (X,Y) will fall into the region $(-\infty, x] \times (-\infty, y]$:



Properties of a Joint CDF

If $F(x, y)$ is the Joint CDF of X and Y, then:

- $0 \leq F(x, y) \leq 1$, for all $x, y \in \mathbb{R}$;
- $F(+\infty, +\infty) = 1$;
- $F(x, -\infty) = F(-\infty, y) = 0$ for any $x, y \in \mathbb{R}$
- For a fixed y, $F(x, y)$ is increasing and right-continuous function of x; and for fixed x, $F(x, y)$ is increasing and right-continuous function wrt y;
- For any $a \leq b$ and $c \leq d$,

$$F(b, d) - F(a, d) - F(b, c) + F(a, c) \geq 0.$$

In fact, if F satisfies all above properties, then it is a Joint CDF of some random vector (X, Y).

Calculation of Probabilities through the Joint CDF

Recall that, for 1D case, if $F(x) = \mathbb{P}(X \leq x)$ is the CDF of the r.v. X, then

$$\mathbb{P}(a < X \leq b) = F(b) - F(a).$$

Now assume that the Joint CDF of X and Y is $F(x, y)$. Then, by definition,

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y),$$

and we can obtain then

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Note: This was the sum from the above properties. That's why it is non-negative!

Calculation of Probabilities through the Joint CDF

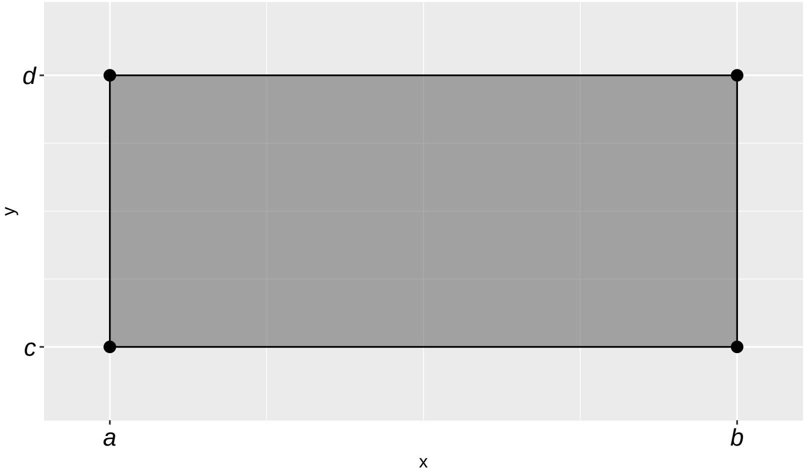


Figure: We calculate $\mathbb{P}\big((X, Y) \in (a, b] \times (c, d]\big)$

Calculation of Probabilities through the Joint CDF

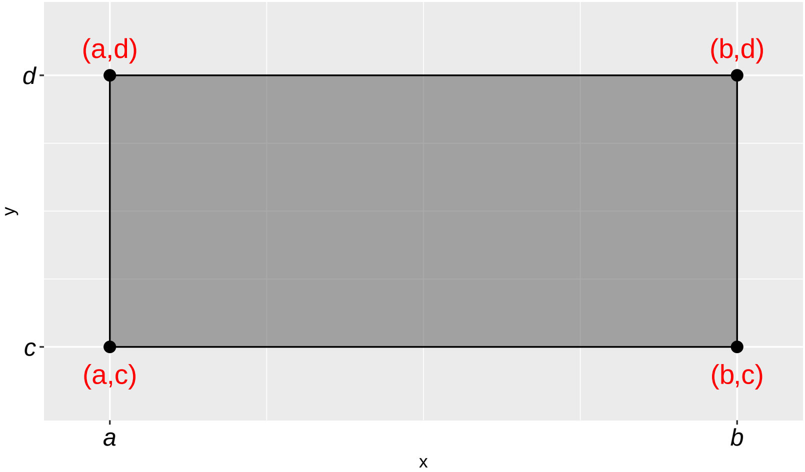


Figure: We calculate $\mathbb{P}\big((X, Y) \in (a, b] \times (c, d]\big)$

Calculation of Probabilities through the Joint CDF

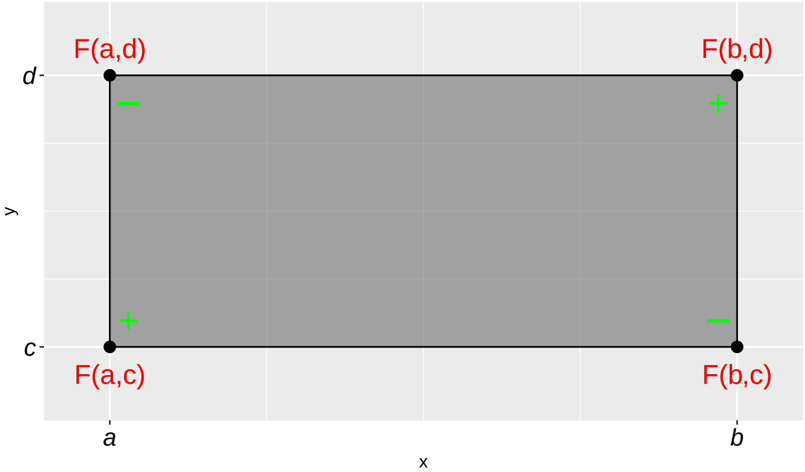


Figure: We calculate $\mathbb{P}\big((X,Y) \in (a,b] \times (c,d]\big)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

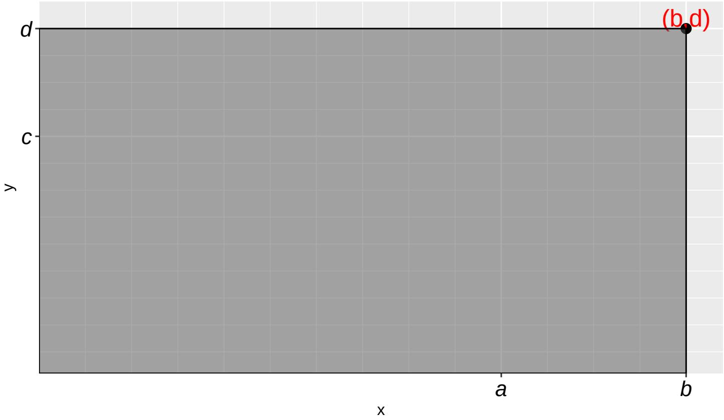


Figure: We calculate first $F(b,d)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

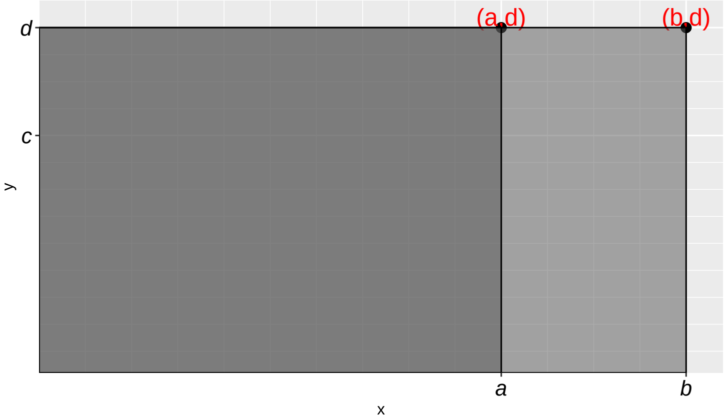


Figure: Now we calculate $F(b,d) - F(a,d)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

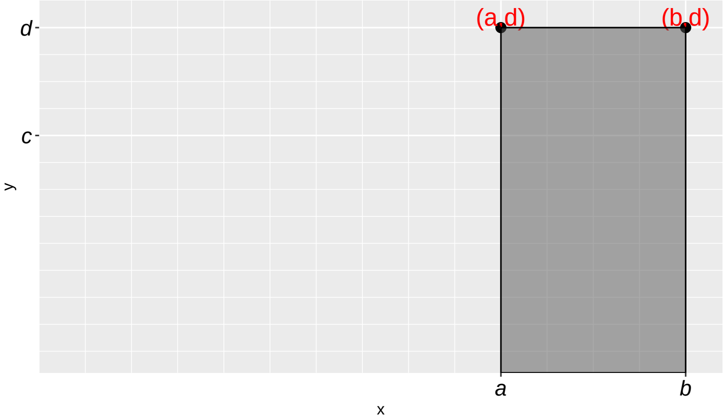


Figure: Now we calculate $F(b, d) - F(a, d)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

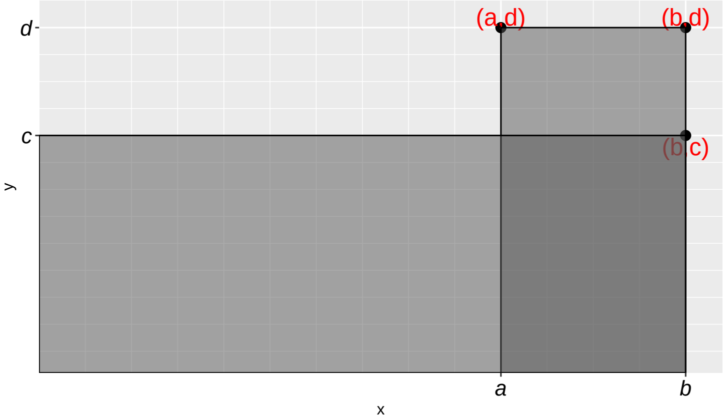


Figure: Now we calculate $F(b, d) - F(a, d) - F(b, c)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

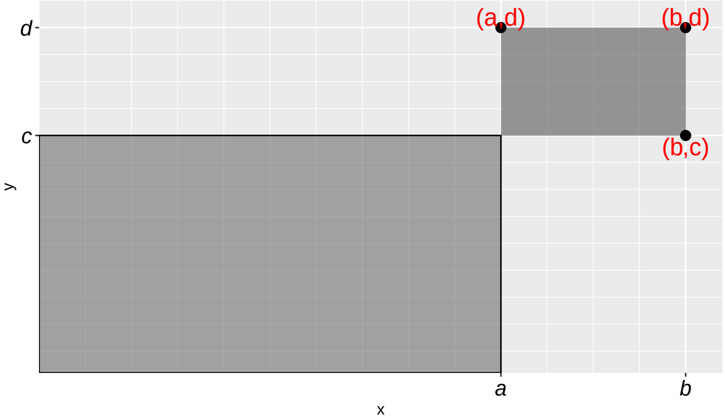


Figure: Now we calculate $F(b, d) - F(a, d) - F(b, c)$

Geometric Proof: Prob of falling into the rectangle

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

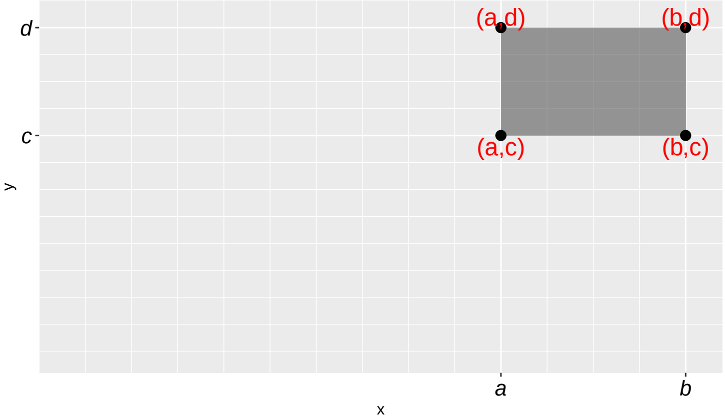


Figure: And this is for $F(b, d) - F(a, d) - F(b, c) + F(a, c)$

Expressing 1D Distributions through the Joint one

It is a remarkable fact that, having the joint distribution of (X, Y), we can find the individual distributions of X and Y, and even more: we will have also the relationship between X and Y. We will talk about this later. But first, let us see how we can find the individual distributions of X and Y, having their joint distribution.

So we consider the following problem:

Find the CDF $F_X(x)$ of X and the CDF $F_Y(x)$ of Y through the Joint CDF $F(x, y)$.

Expressing 1D Distributions through the Joint one

By the definition of the Joint CDF $F(x, y)$ of X and Y:

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y), \qquad \forall x, y \in \mathbb{R}.$$

Now, we want to calculate

$$F_X(x) = \mathbb{P}(X \leq x) \qquad \text{and} \qquad F_Y(x) = \mathbb{P}(Y \leq x), \quad \forall x \in \mathbb{R}.$$

This is not so hard to do:

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y \text{ is arbitrary}) =$$

$$= \mathbb{P}(X \leq x, Y \leq +\infty) = F(x, +\infty).$$

Similarly,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq +\infty, Y \leq y) = F(+\infty, y).$$

So we have obtained the following result:

Marginal CDFs through the Joint CDF

- $F_X(x) = F(x, +\infty)$ for any $x \in \mathbb{R}$
- $F_Y(y) = F(+\infty, y)$ for any $y \in \mathbb{R}$

Given the Joint CDF $F(x, y)$ of X and Y, the CDFs F_X and F_Y are called **Marginal CDFs of X and Y**.

Joint CDF and Marginal CDFs

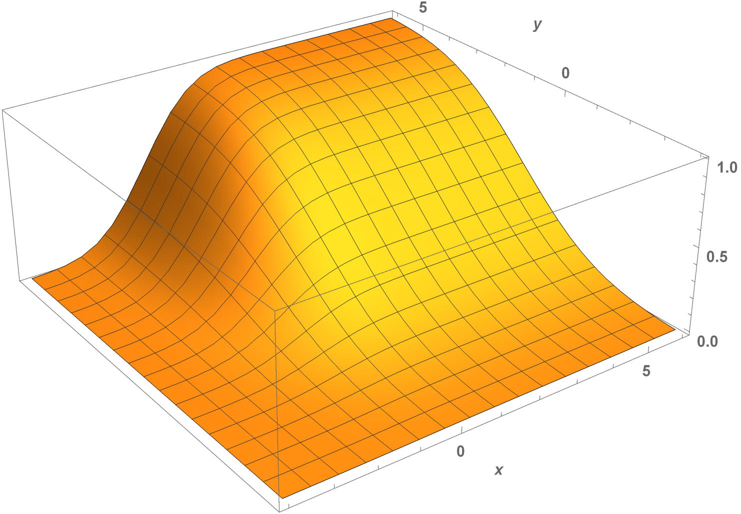


Figure: Joint CDF (surface), $F_{X,Y}(x, y)$

Joint CDF and Marginal CDFs

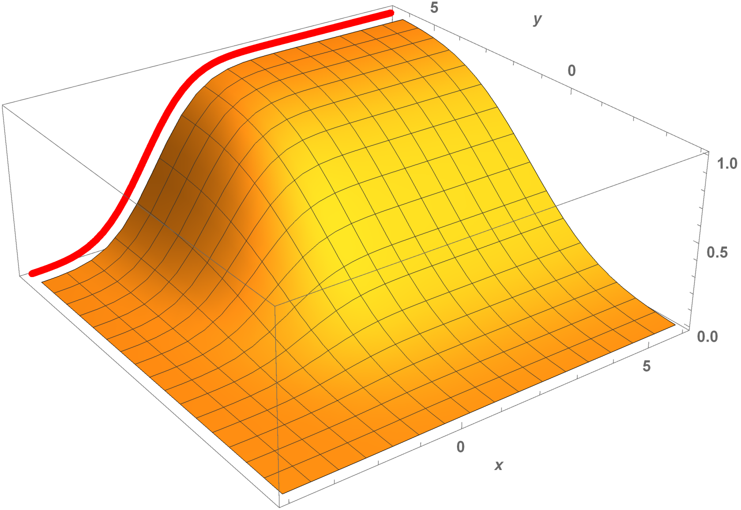


Figure: Joint CDF $F_{X,Y}(x,y)$ and the CDF of X, $F_X(x)$

Joint CDF and Marginal CDFs

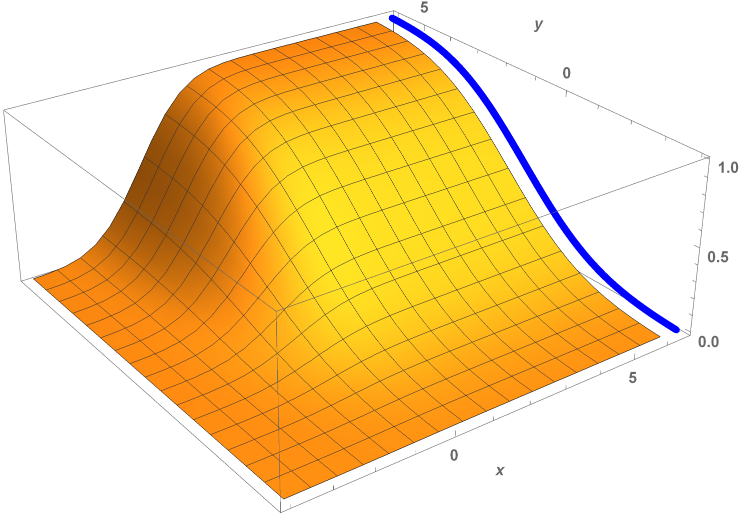


Figure: Joint CDF $F_{X,Y}(x, y)$ and the CDF of Y, $F_Y(y)$

Joint CDF and Marginal CDFs

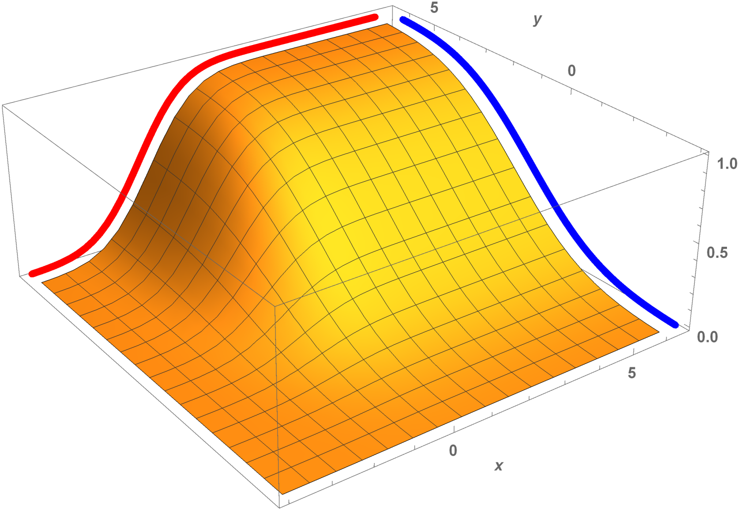


Figure: Joint CDF (surface), and Marginal CDFs (red and blue)

Expressing 1D Distributions through the Joint one

Note: Having the Joint CDF $F(x, y)$, we were able to find the CDFs (Distributions) of X and Y easily. Unfortunately, the inverse is not true, in general: having the individual CDFs of X and Y, we **cannot** find the Joint CDF of X and Y. **This is because F_X and F_Y do not give any info about the relationship between X and Y, which is a very important information**.

Joint CDF of n RVs

Above we described the joint distribution of 2 r.v.s, X and Y, through their Joint CDF:

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y), \qquad \forall (x, y) \in \mathbb{R}^2.$$

Now, assume we have n r.v.s, $X_1, X_2, ..., X_n$, defined on the same Sample Space Ω. Then we denote

$$\mathbf{X} = [X_1, X_2, ..., X_n]^T$$

and call it an n-dim random vector with coordinates X_k. And we describe the distribution of $\mathbf{X}$ through its Joint CDF defined in the following way: for any $\mathbf{x} = [x_1, ..., x_n]^T \in \mathbb{R}^n$,

$$F_{\mathbf{X}}(\mathbf{x}) = F(x_1, x_2, .., x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n).$$

Joint CDF of n RVs

Having the Joint CDF of r.v.s $X_1, ..., X_n$

$$F_{\mathbf{X}}(\mathbf{x}) = F(x_1, x_2, .., x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n),$$

we can find:

- Individual CDF of X_k:

$$F_{X_k}(x) = F_{\mathbf{X}}(+\infty, ..., +\infty, x, +\infty, ..., +\infty), \qquad \forall x \in \mathbb{R},$$

  where x is in the k-th place;
- the Joint CDF of any subgroup of r.v.s. Say, the Joint CDF of X_1 and X_3 will be

$$F_{X_1,X_3}(x_1, x_3) = F_{\mathbf{X}}(x_1, +\infty, x_3, +\infty, ..., +\infty), \quad \forall x_1, x_3 \in \mathbb{R}.$$

Joint distribution of 2 r.vs

Although Joint CDFs completely describe the distribution of 2 (or more) r.v.s, but, unfortunately, CDFs are not so useful for calculating probabilities of the form

$$\mathbb{P}\big((X, Y) \in A\big), \qquad A \subset \mathbb{R}^2,$$

except the case when A is a rectangle (or a union of disjoint rectangles). Also, they are not so handy to find the relationship between X and Y. So we will develop the theory for Joint distributions by using Joint PMFs (for discrete random vectors) and Joint PDFs (for continuous random vectors).

Joint distribution of 2 r.vs

So again let us talk about the 2D case, about the distribution of 2 r.v.s X and Y. Then the following cases can happen:

- X and Y are discrete r.v.s - so they are Jointly Discrete;
- X and Y are Jointly Continuous r.v.s;
- One of X and Y is Discrete, the other one is Continuous;
- Other

In our course, we will consider only the first two cases.

Discrete random vectors

Discrete R. Vectors, Joint PMF

Assume that the r.v.s X and Y, defined on the same Experiment, are both discrete. Then we will say that the random vector (X, Y) is discrete.

Let $x_1, x_2, ..., x_n, ...$ be the values of X and $y_1, y_2, ..., y_m, ...$ be the values of Y (not necessary of the same size: say, the range of X can be finite, and the range of Y can be infinite).

Then, the vector (X, Y) can take only the values (x_i, y_j). We introduce the probabilities

$$p_{i,j} = \mathbb{P}\Big((X, Y) = (x_i, y_j)\Big) = \mathbb{P}(X = x_i, Y = y_j), \qquad i, j = 1, 2, ...$$

and call the table of probabilities $\mathbb{P}(X = x_i, Y = y_j)$ **the Joint PMF** of X and Y (or, of the Random Vector (X, Y)).

Joint PMF, Properties, and the Table form

Clearly,

$$\sum_{i,j} p_{i,j} = \sum_{i,j} \mathbb{P}(X = x_i,\, Y = y_j) =$$

$$= \mathbb{P}(X \text{ is arbitrary, } Y \text{ is arbitrary}) = \mathbb{P}(\Omega) = 1.$$

So, for any Joint PMF, we need to have

$$p_{i,j} \geq 0 \qquad \text{and} \qquad \sum_{i,j} p_{i,j} = 1.$$

Joint PMF in the Table form

Usually we write the Joint PMF of X and Y in the table form:

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1} = \mathbb{P}(X = x_1, Y = y_1)$	$p_{2,1} = \mathbb{P}(X = x_2, Y = y_1)$	$\dots$
y_2	$p_{1,2} = \mathbb{P}(X = x_1, Y = y_2)$	$p_{2,2} = \mathbb{P}(X = x_2, Y = y_2)$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Or, for short,

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Here, the sum of all $p_{i,j}$-s is 1.

Joint PMF Plot

We can draw the Joint PMF plot in the following way: at the point (x_i, y_j), we can draw a vertical line with a height $p_{i,j}$. Say, for the following Joint PMF,

$Y \backslash X$	0	3	4
-1	0.3	0.1	0.05
1	0.1	0.2	0.25

the plot will be:

Joint PMF Plot

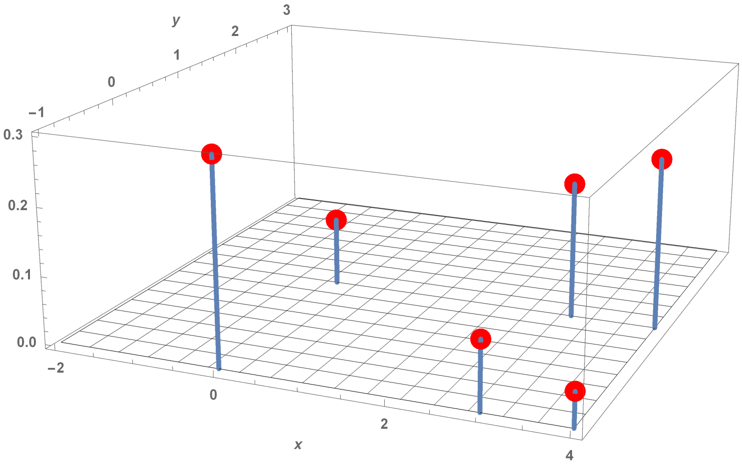


Figure: Joint PMF

Joint CDF Plot, Discrete case

For 2D discrete random vectors, the Joint CDF, as in 1D case, will be a piecewise constant function (of two variables), and the graph of the Joint CDF will look like the following:

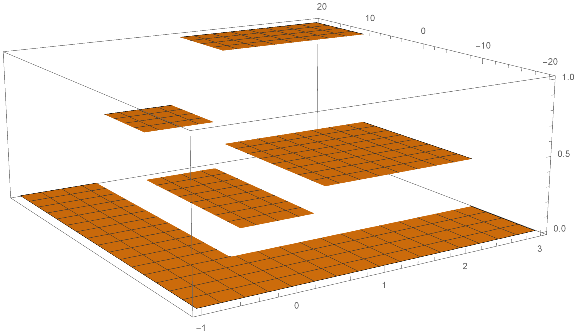


Figure: Joint CDF, Discrete r.v.s

Joint PMF, Calculation of Probabilities

Having the Joint PMF, we can calculate all probabilities we want: if $A \subset \mathbb{R}^2$, then

$$\mathbb{P}\Big((X, Y) \in A\Big) = \sum_{(x_i, y_j) \in A} p_{i,j}.$$

Say,

$$\mathbb{P}(a \le X \le b, c \le Y \le d) = \mathbb{P}((X, Y) \in [a, b] \times [c, d]) =$$

$$= \sum_{\substack{a \le x_i \le b \\ c \le y_j \le d}} p_{i,j}.$$

Joint PMF, Calculation of Probabilities

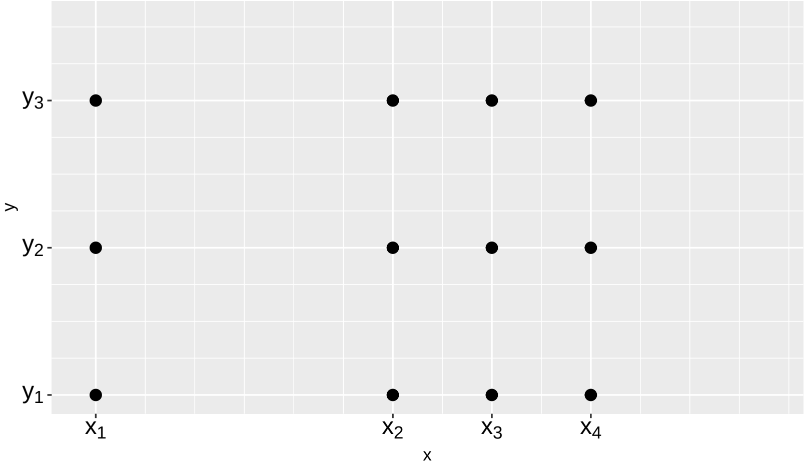


Figure: Joint PMF support

Joint PMF, Calculation of Probabilities

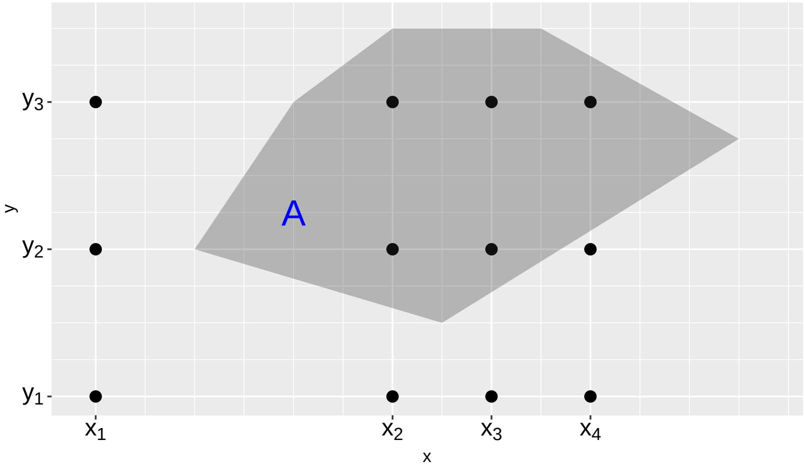


Figure: Joint PMF support and the region A

Joint PMF, Calculation of Probabilities

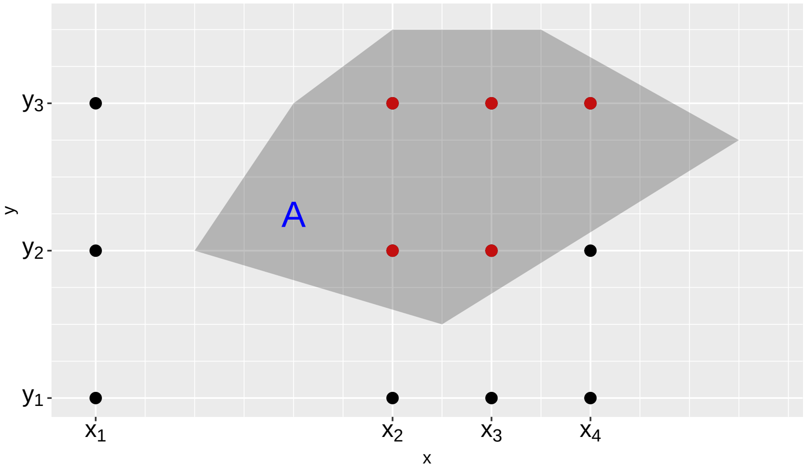


Figure: Points $(x_i, y_j) \in A$

Example:

Example 29.1: Assume X and Y are discrete r.v.s with the following Joint PMF:

$Y \backslash X$	-1	0	2
2	0.3		0.1
5	0	0.4	0.1

a. Fill in the table;

b. Plot the Joint PMF;

c. Calculate $\mathbb{P}(0 \leq X < 3,\ Y \leq 3.5)$;

d. Calculate $\mathbb{P}(X^2 + Y \leq 6)$.

Example:

Example 29.2: Assume we are rolling a fair die. Let $X = -\sqrt{2}$, if the number shown is even, $X = 0$, if 1 appears, and $X = 4$ otherwise. Also, let $Y = 0$, if 2 or 5 appears, and $Y = 1$ otherwise.

a. Construct the Joint PMF of (X, Y);
b. Calculate $\mathbb{P}(X \in [0, 4.5], Y \in [-1, 0.5])$;
c. Calculate $\mathbb{P}(X^2 + Y^2 > 4)$;
d. Calculate $\mathbb{P}(X = Y)$;
e. Calculate $F(3, 2.4)$, where F is the Joint CDF of X and Y.

Example:

Example 29.3: Assume

$$X \sim \begin{pmatrix} 0 & 2 & 5 \\ 0.1 & 0.7 & 0.2 \end{pmatrix}$$

and assume $Y = \sqrt{X}$. Find the PMF of (X, Y).

Example:

Example 29.4: Assume X and Y are discrete r.v.s. Assume we know that $Y \sim Bernoulli(0.4)$, and we know the following: if $Y = 0$, then $X \sim Geom(0.8)$, and if $Y = 1$, then $X \sim Pois(2)$. Construct the Joint PMF of X and Y, and find the distribution of $X - Y$.

Exercise

Exercise: Given a Joint PMF of a random vector (X, Y), write a code to generate a random sample from that distribution.

Marginal PMFs

Assume that the Joint PMF of Discrete r.v.s X and Y is given by:

$Y \setminus X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

We want to construct the individual PMFs of X and Y. The process described below is called **Marginalization**.
To find the PMF of X, we need to calculate $\mathbb{P}(X = x_i)$. And we need to use the Probabilities $\mathbb{P}(X = x_i, Y = y_j)$:

Marginal PMFs

Assume that the Joint PMF of Discrete r.v.s X and Y is given by:

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

We write

$$\mathbb{P}(X = x_i) = \mathbb{P}(X = x_i,\, Y \text{ is arbitrary}) = \sum_j \mathbb{P}(X = x_i,\, Y = y_j),$$

so

$$\mathbb{P}(X = x_i) = \text{The sum of the column under } x_i = \sum_j p_{i,j}$$

Marginal PMFs

Assume that the Joint PMF of Discrete r.v.s X and Y is given by:

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Similarly, for the PMF of Y,

$$\mathbb{P}(Y = y_j) = \text{The sum of the row right to } y_j = \sum_i p_{i,j}$$

Marginal PMFs

Again, if the Joint Distribution of X and Y is

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

then

- $\mathbb{P}(X = x_i) = \sum_j \mathbb{P}(X = x_i, Y = y_j)$
- $\mathbb{P}(Y = y_j) = \sum_i \mathbb{P}(X = x_i, Y = y_j)$

The description is, say, for the first formula, that we are "summing out Y" (or "integrating out Y" in the continuous case).

Marginal PMFs

We are adding these to out Joint PMF table, in the **Margins**:

$Y \backslash X$	x_1	x_2	$\dots$	$\mathbb{P}(Y = y_j)$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$	$p_{1,1} + p_{2,1} + \dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$	$p_{1,2} + p_{2,2} + \dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbb{P}(X = x_i)$	$p_{1,1} + p_{1,2} + \vdots$	$p_{2,1} + p_{2,2} + \vdots$	$\vdots$	

The red-colored parts are the PMF's of X and Y, respectively.

Marginal PMFs

So, having the Joint PMF of X and Y,

$Y \backslash X$	x_1	x_2	$\dots$	$\mathbb{P}(Y = y_j)$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$	$p_{1,1} + p_{2,1} + \dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$	$p_{1,2} + p_{2,2} + \dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbb{P}(X = x_i)$	$p_{1,1} + p_{1,2} + \vdots$	$p_{2,1} + p_{2,2} + \vdots$	$\vdots$	

We will obtain the PMF of, say, X:

Values of X	x_1	x_2	$\dots$
$\mathbb{P}(X = x)$	$p_{1,1} + p_{1,2} + \dots$	$p_{2,1} + p_{2,2} + \dots$	$\dots$

Marginal PMFs: Example

Example 29.5: Assume that the Joint PMF of X and Y is given by:

$Y \backslash X$	-2	0	1
$1/3$	0.1	0.15	0.2
$\sqrt{3}$	0.3	0.1	0.15

a. Check that this is a legitimate (valid) Joint PMF;
b. Find the (Marginal) Distributions of X and Y;
c. Calculate $\mathbb{P}(X > -1)$;
d. Calculate $\mathbb{P}(X = 1 \,|\, Y = \sqrt{3})$;
e. Find the conditional distribution of X given Y.

Marginal PMFs

Now assume X and Y are Jointly discrete r.v.s, and $A \subset \mathbb{R}$. Assume we want to calculate

$$\mathbb{P}(X \in A).$$

We can do this in two ways:

I: First calculate the Marginal PMF of X, then calculate that probability:

$$\mathbb{P}(X \in A) = \sum_{x_i \in A} \mathbb{P}(X = x_i) = \sum_{x_i \in A} p_i;$$

II: Do it in one step:

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in A,\, Y \text{ is arbitrary}) =$$

$$= \sum_{\substack{x_i \in A \\ \text{any } y_j}} \mathbb{P}(X = x_i,\, Y = y_j) = \sum_{\substack{x_i \in A \\ \text{any } y_j}} p_{i,j}.$$

Marginal PMFs: Example

Example 29.6: Assume that the Joint PMF of X and Y is given by:

$Y \backslash X$	-1	2
1	0.1	0.4
2	$0.1 \cdot 0.5$	$0.4 \cdot 0.5$
3	$0.1 \cdot 0.5^2$	$0.4 \cdot 0.5^2$
4	$0.1 \cdot 0.5^3$	$0.4 \cdot 0.5^3$
$\vdots$	$\vdots$	$\vdots$

Calculate $\mathbb{P}(2 \leq Y < 4)$.

Marginal PMFs: Example

Example 29.7: Assume that the Joint PMF of X and Y is given by:

$Y \backslash X$	0	2	4	$\mathbb{P}(Y = y_j)$
-1	0.1		0.1	0.4
0	0.3			0.5
1				
$\mathbb{P}(X = x_i)$	0.5	0.3		

Calculate $\mathbb{P}(X + Y > 2)$.

Conditional Distribution of r.v.s

Now assume X and Y are Jointly discrete r.v.s, with the PMF

$Y \backslash X$	x_1	x_2	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Now, for any j, we can find the distribution of $X \,|\, Y = y_j$, i.e., we can find the distribution of X, given that $Y = y_j$:

$$\mathbb{P}(X = x_i \,|\, Y = y_j) = \frac{\mathbb{P}(X = x_i,\, Y = y_j)}{\mathbb{P}(Y = y_j)} = \frac{p_{i,j}}{\sum_k p_{k,j}}.$$

Conditional Distribution of r.v.s

The idea is the following: say, we want to find the conditional distribution of X, given that $Y = y_2$. Then we look at the row in Joint PMF, corresponding to $Y = y_2$:

$Y \backslash X$	x_1	x_2	x_3	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$p_{3,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$p_{3,2}$	$\dots$
y_3	$p_{1,3}$	$p_{2,3}$	$p_{3,3}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

The red row is our new, reduced Sample Space. Now, we calculate the proportion of each $p_{i,2}$ in that row, i.e. we normalize that row to have in the sum 1 - we divide all elements in that row to the sum of all elements:

Conditional Distribution of r.v.s

The idea is the following: say, we want to find the conditional distribution of X, given that $Y = y_2$. Then we look at the row in Joint PMF, corresponding to $Y = y_2$:

$Y \backslash X$	x_1	x_2	x_3	$\dots$
y_1	$p_{1,1}$	$p_{2,1}$	$p_{3,1}$	$\dots$
y_2	$p_{1,2}$	$p_{2,2}$	$p_{3,2}$	$\dots$
y_3	$p_{1,3}$	$p_{2,3}$	$p_{3,3}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

We obtain:

$$\frac{p_{1,2}}{p_{1,2}+p_{2,2}+p_{3,2}+\dots}, \quad \frac{p_{2,2}}{p_{1,2}+p_{2,2}+p_{3,2}+\dots}, \quad \dots$$

And these are, exactly, the probabilities $\mathbb{P}(X = x_i \,|\, Y = y_2)$, $i = 1, 2, \dots$

Conditional PMFs: Example

Example 29.8: Let X and Y be Jointly discrete with the following Joint PMF:

$Y \backslash X$	-10	0	7
4	0.4	0	0.1
5	0.1	0.1	0.3

a. Find the Conditional distributions $X|Y$;
b. Find the Conditional distribution $Y|X = 7$ and $Y|X = 0$.

Conditional PMFs: Example

Example 29.9: Let X and Y be Jointly discrete with the following Joint PMF:

$Y \backslash X$	-10	0	7
4	0.02	0.06	0.12
5	0.08	0.24	0.48

a. Find the Conditional distributions $X|Y$;
b. Explain/Interpret.

Conditional PMFs: Example

Example 29.10: Assume

$$Y \sim \begin{pmatrix} 1.5 & 5 \\ 0.3 & 0.7 \end{pmatrix}$$

and we have the following Conditional PMFs:

$X \mid Y = 1.5$	3	5	7
$\mathbb{P}(X = x \mid Y = 1.5)$	0.3	0.6	0.1

$X \mid Y = 5$	3	5	7
$\mathbb{P}(X = x \mid Y = 5)$	0.4	0.4	0.2

Construct the Joint PMF of X and Y.

Generalization for more than 2 r.v.s

The notions above can be generalized to more than 2 r.v.s. Say, if X, Y, Z are r.v. in the same Experiment, then we have already defined their Joint CDF:

$$F(x, y, z) = \mathbb{P}(X \leq x, Y \leq y, Z \leq z), \qquad \forall (x, y, z) \in \mathbb{R}^3;$$

Now, if X, Y, Z are Discrete, then their Joint PMF is given by

$$p_{i,j,k} = \mathbb{P}(X = x_i, Y = y_j, Z = z_k),$$

where x_i, y_j, z_k run over all possible values of X, Y, Z, respectively.

In this case, having a subset $A \subset \mathbb{R}^3$, we can calculate

$$\mathbb{P}\big((X, Y, Z) \in A\big) = \sum_{(x_i, y_j, z_k) \in A} p_{i,j,k}.$$

Generalization for more than 2 r.v.s

Also, we can find individual (Marginal) distribution of, say, X: the PMF will be

$$\mathbb{P}(X = x_i) = \sum_{j,k} \mathbb{P}(X = x_i,\, Y = y_j, Z = z_k),$$

and also we can find Joint PMFs of any pair of X, Y, Z. Say, the Joint PMF of X, Z will be

$$\mathbb{P}(X = x_i, Z = z_k) = \sum_{j} \mathbb{P}(X = x_i,\, Y = y_j, Z = z_k).$$