Lecture note 6: Jointly distributed random variables

Foundation of Probability Theory/STA 203

Zhuosong ZHANG

Department of Statistics and Data Science, SUSTech

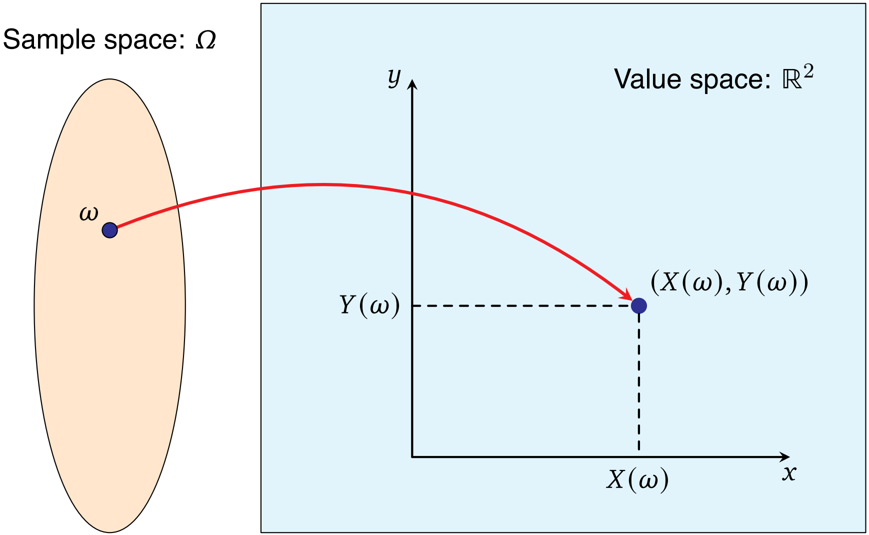
Fall, 2023

Joint distribution functions

Introduction

- We have been considering one-variate random variables.
- How to study the distributions of two random variables X and Y?
- How to study their relationships?

Definition of random vectors



Definition of random vectors

Definition 1

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. We say (X, Y) is a 2-dimensional random vector, or a bivariate random variable if $\omega \mapsto (X(\omega), Y(\omega))$ is a function valued on $\mathbb{R}^2$.

Example 2 (Some examples)

- Let $\Omega = \{$all students at SUSTech$\}$, and $X(\omega) =$ height of ω, $Y(\omega) =$ age of ω.
- Let $\Omega = \{$all products in a supermarket$\}$, and $X(\omega) =$ price of ω, $Y(\omega) =$ date of manufacture of ω.
- Can you give some examples?

p-dimensional random vectors

Definition 3

Generally, we say $(X_1, X_2, \ldots, X_p)$ is a p-variate random variable, or a p-dimensional random vector, if $\omega \mapsto (X_1(\omega), X_2(\omega), \ldots, X_p(\omega))$ is a p-dimensional function on $\mathbb{R}^p$.

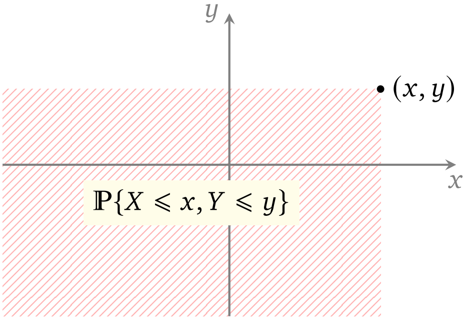
Remark

In the following part of this note, we will focus on the bivariate random variables without further announcement.

Joint distribution functions

- For any two random variables X and Y, the joint distribution function of X and Y is defined by

$$F(x,y) = \mathbb{P}\{X \leqslant x, Y \leqslant y\}, \quad -\infty < x, y < \infty.$$



Properties of $F(x, y)$

Proposition 4

Denote by F_X and F_Y the distribution functions of X and Y, respectively. We have

$$F_X(x) = F(x, \infty), \quad F_Y(y) = F(\infty, y).$$

The distribution functions F_X and F_Y are sometimes referred to as the marginal distribution.

Properties of $F(x, y)$

Proposition 5

For any x and y,

$$\mathbb{P}\{X > x, Y > y\} = 1 - F_X(x) - F_Y(y) + F(x, y).$$

Joint mass function

- If X and Y are both discrete random variables, the joint probability mass function of X and Y is defined as

$$p(x, y) = \mathbb{P}\{X = x, Y = y\}.$$

- The marginal probability mass function of X can be obtained by

$$p_X(x) = \mathbb{P}\{X = x\} = \sum_{y:p(x,y)>0} p(x, y).$$

- Similarly,

$$p_Y(y) = \mathbb{P}\{Y = y\} = \sum_{x:p(x,y)>0} p(x, y).$$

Examples

Example 6

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white, and 5 blue balls. If we let X and Y denote, respectively, the number of red and white balls chosen, find the joint mass function of (X, Y).

Solution.

Let X and Y denote the number of red and white balls chosen, respectively, then the joint probability mass function of X and Y is given by

$$p(0,0) = \mathbb{P}\{X = 0, Y = 0\} = \frac{\binom{5}{3}}{\binom{12}{3}} = \frac{10}{220} \approx 0.0455,$$

$$\vdots$$

Solution

The pmf of (X, Y) can be listed in the following table:

$X = i$ \ $Y = j$	0	1	2	3	Row sum = $\mathbb{P}\{X = i\}$
0	0.0455	0.1818	0.1364	0.0182	0.3819
1	0.1364	0.2727	0.0818	0	0.4909
2	0.0682	0.0545	0	0	0.1227
3	0.0045	0	0	0	0.0045
Column sum = $\mathbb{P}\{Y = j\}$	0.2546	0.5090	0.2182	0.0182	1

Table: Joint pmf of random variables X and Y

Properties of joint pmf

Proposition 7

Joint pmf $p(x, y)$ has the following basic properties:

(a) **Non-negativity:** $p(x, y) \geqslant 0$;

(b) **Normalization:**

$$\sum_x \sum_y p(x, y) = 1.$$

Examples

Example 8

Let X be a number uniformly chosen from $1, 2, 3, 4$, and let Y be a number uniformly chosen from $1, 2, \ldots, X$. Find the joint pmf of (X, Y) and find $\mathbb{P}(X = Y)$.

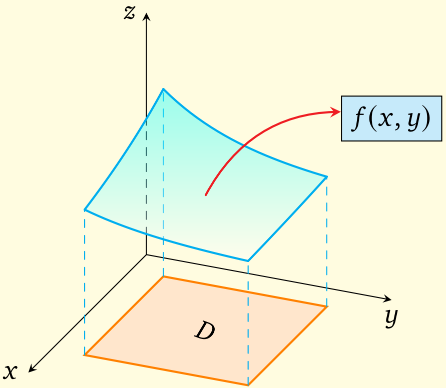
Jointly continuous random variables

Definition 9

X and Y are said to be jointly continuous if there exists a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ such that for every $D \subset \mathbb{R}^2$,

$$\mathbb{P}\{(X, Y) \in D\} = \iint_{(x,y)\in D} f(x, y)dxdy.$$

The function $f(x, y)$ is called the joint probability density function of X and Y.



Properties

- In particular, if $D = A \times B$, that is, $D = \{(x, y) : x \in A, y \in B\}$, we have

$$\mathbb{P}\{X \in A, Y \in B\} = \int_B \left(\int_A f(x, y) dx \right) dy.$$

- The joint distribution function is given by

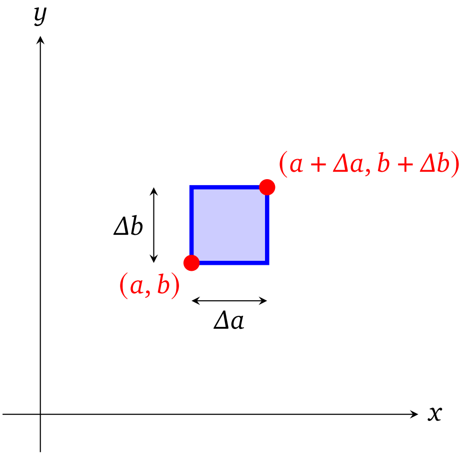
$$F(x, y) = \mathbb{P}\{X \leqslant x, Y \leqslant y\} = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v) du dv.$$

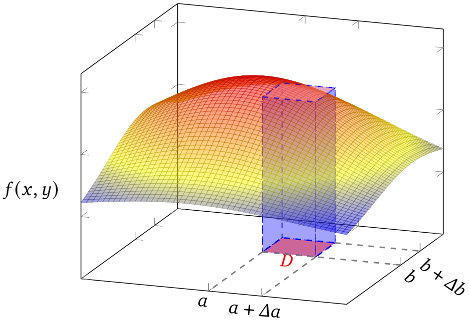
- It follows that

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

- Moreover, if Δa and Δb are small,

$$\mathbb{P}\{a < X < a + \Delta a, b < Y < b + \Delta b\} \approx f(a, b) \Delta a \Delta b.$$


y
x
Δb
Δa
(a, b)
(a + Δa, b + Δb)


f(x, y)
D
a
a + Δa
b
b + Δb

Properties

Joint pdf has the following properties:

(a) **Non-negativity**: $f(x,y) \geqslant 0$ almost everywhere;

(b) **Normalization**:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1.$$

Marginal density of continuous random vectors

Proposition 10

If X and Y are jointly continuous with joint pdf $f(x, y)$, then they are individually continuous, and their pdf are given as follows:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx.$$

Example 11

The joint density function of X and Y is given by

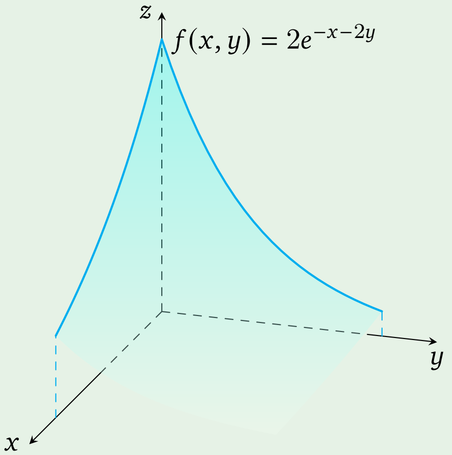
$$f(x,y) = \begin{cases} 2e^{-x-2y} & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute

(a) $\mathbb{P}\{X > 1, Y < 1\}$;

(b) $\mathbb{P}\{X < Y\}$;

(c) $\mathbb{P}\{X < a\}$ for some $a > 0$.



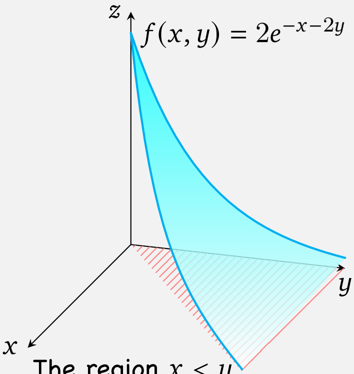
Solution.

(a) We have

$$\begin{aligned}\mathbb{P}\{X > 1, Y < 1\} &= \int_0^1 \left(\int_1^\infty 2e^{-x-2y} dx \right) dy \\ &= \int_0^1 2e^{-2y} \left(\int_1^\infty e^{-x} dx \right) dy \\ &= e^{-1} \int_0^1 2e^{-2y} dy \\ &= e^{-1}(1 - e^{-2}) \approx 0.3181.\end{aligned}$$

(b) We have

$$
\begin{aligned}
\mathbb{P}\{X < Y\} \\
&= \iint\limits_{(x,y):x<y} 2e^{-x-2y}dxdy \\
&= \int_0^\infty \left(\int_0^y 2e^{-x-2y}dx\right)dy \\
&= \int_0^\infty (2e^{-2y})(1-e^{-y})dy \\
&= \int_0^\infty 2e^{-2y}dy - \int_0^\infty 2e^{-3y}dy \\
&= 1 - \frac{2}{3} \\
&= \frac{1}{3}.
\end{aligned}
$$



The region $x < y$

(c) For any $a > 0$,

$$\begin{aligned}\mathbb{P}\{X < a\} &= \int_0^a \left(\int_0^\infty 2e^{-x-2y} dy\right) dx \\ &= \int_0^a e^{-x} dx \\ &= 1 - e^{-a}.\end{aligned}$$

Example 12

Let X and Y follow the joint density function given by

$$f(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \leqslant R^2 \\ 0 & \text{otherwise} \end{cases}$$

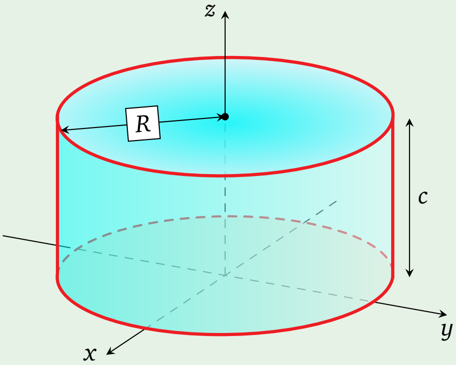
for some value of c.

(a) Determine c.

(b) Find the marginal density functions of X and Y.

(c) Let D be the distance between $(0,0)$ and (X,Y). Compute $\mathbb{P}\{D \leqslant a\}$.

(d) Find $\mathbb{E}[D]$.



Solution.

(a) Because

$$\int_{-\infty}^{\infty} f(x,y)dxdy = 1,$$

it follows that

$$c\iint_{x^2+y^2\leqslant R^2} = 1 \implies c = \frac{1}{\pi R^2}.$$

■

Solution.

(b) Observe that if $|x| \leqslant R$,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\pi R^2}\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy = \frac{2\sqrt{R^2-x^2}}{\pi R^2},$$

and $f_X(x) = 0$ if $|x| > R$.

By symmetry,

$$f_Y(y) = \begin{cases} \dfrac{2\sqrt{R^2-y^2}}{\pi R^2} & \text{if } |x| \leqslant R, \\ 0 & \text{otherwise.} \end{cases}$$

■

Solution.

(c) For any $0 \leqslant z \leqslant R$,

$$\begin{aligned}
\mathbb{P}\{D \leqslant z\} &= \mathbb{P}\{\sqrt{X^2+Y^2} \leqslant z\} \\
&= \mathbb{P}\{X^2+Y^2 \leqslant z^2\} \\
&= \iint\limits_{(x,y):x^2+y^2 \leqslant z^2} f(x,y)dxdy \\
&= \frac{\pi z^2}{\pi R^2} = \frac{z^2}{R^2}.
\end{aligned}$$

For $z > R$, $\mathbb{P}\{D \leqslant z\} = 1$. ■

Solution.

(d) From (c), we have

$$f_D(z) = \frac{2z}{R^2} \quad 0 \leqslant a \leqslant R.$$

Hence,

$$\mathbb{E}[D] = \int_0^R \frac{2z^2}{R^2} dz = \frac{2R}{3}. \qquad \blacksquare$$

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 13

Assume that (X, Y) has the following pdf

$$f(x,y) = \begin{cases} 1 & 0 < x < 1, |y| < x, \\ 0 & \text{otherwise.} \end{cases}$$

Find

(a) the marginal density of X and Y, respectively;

(b) $\mathbb{P}(X < 1/2)$;

(c) $\mathbb{P}(Y > 1/2)$.

Question

If we are given the marginal distributions of X and Y, can we determine the joint distribution of (X, Y)?

Example 14

The following two joint densities have same marginal distributions:

$$f(x,y) = \begin{cases} x+y & 0 \leqslant x, y \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$g(x,y) = \begin{cases} (0.5+x)(0.5+y) & 0 \leqslant x, y \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

Independent random variables

Independence of random variables

- We have already defined independence of events:

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\,\mathbb{P}(F).$$

- For two random variables X and Y, if the behavior of X does not affect the distribution of Y, or verse versa, we can also say that X and Y are “independent”.

Definition 15

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, on which defining two random variables X and Y, respectively. We say X and Y are independent if

$$\{X \in A\} \text{ is independent of } \{Y \in B\} \quad \text{for all } A, B \subset \mathbb{R}.$$

Remark

In particular, for all $x, y \in \mathbb{R}$,

$$\mathbb{P}\{X \leqslant x, Y \leqslant y\} = \mathbb{P}\{X \leqslant x\}\, \mathbb{P}\{Y \leqslant y\}.$$

Proposition 16

If X and Y are jointly discrete random variables with joint pmf $p(x, y)$, then the following two arguments are equivalent:

(i) X and Y are independent.

(ii) $p(x, y) = p_X(x)p_Y(y)$ for all $x, y \in \mathbb{R}$.

If X and Y are jointly continuous random variables with joint pdf $f(x, y)$, then the following are equivalent:

(i) X and Y are independent.

(ii) $f(x, y) = f_X(x)f_Y(y)$ for all $x, y \in \mathbb{R}$.

Example 17 (Poker cards)

Suppose there are two piles of cards: $\{2\spadesuit, 3\spadesuit, 4\spadesuit, 5\spadesuit\}$ and $\{2\heartsuit, 3\heartsuit, 4\heartsuit, 5\heartsuit\}$. We randomly select one card from each pile, and let X be the card value of $\spadesuit$, and let Y be the card value of $\heartsuit$.

Are X and Y independent?

Solution.

The sample space if $\Omega = \{(2\spadesuit, 2\heartsuit), (2\spadesuit, 3\heartsuit), \ldots, (5\spadesuit, 5\heartsuit)\}$ containing 16 elements. Let $\mathscr{F}$ be the power set of Ω, and $\mathbb{P}$ is the classical probability. Then,

$$\mathbb{P}\{X = 2\} = \mathbb{P}\{(2\spadesuit, 2\heartsuit), (2\spadesuit, 3\heartsuit), (2\spadesuit, 4\heartsuit), (2\spadesuit, 5\heartsuit)\} = \frac{4}{16} = \frac{1}{4},$$

and similarly, $\mathbb{P}\{X = i\} = \mathbb{P}\{Y = j\} = 1/4$ for all $2 \leqslant i, j \leqslant 5$. Moreover, for any $2 \leqslant i, j \leqslant 5$, we have $\mathbb{P}\{X = i, Y = j\} = 1/16$. Then, X and Y are independent. ∎

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 18

Suppose that X and Y are jointly discrete random variables with marginal distribution as follows:

X	-1	0	1
p_X	$1/4$	$1/2$	$1/4$

and

Y	0	1
p_Y	$1/2$	$1/2$

If $\mathbb{P}\{XY = 0\} = 1$, find

(a) the joint pmf of X and Y,

(b) whether X and Y are independent?

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 19

Suppose that X and Y are independent continuous random variables with pdf f_X and f_Y, respectively. Find the distribution functions of $Z = \max(X, Y)$ and $W = \min(X, Y)$.

Solution.

The cdf of Z is

$$\begin{aligned} F_Z(z) &= \mathbb{P}(\max(X,Y) \leqslant z) \\ &= \mathbb{P}\{X \leqslant z, Y \leqslant z\} \\ &= \mathbb{P}\{X \leqslant z\}\,\mathbb{P}\{Y \leqslant z\} \\ &= F_X(z)F_Y(z). \end{aligned}$$

■

Independence of many random variables

Definition 20

Discrete random variables $X_1, \ldots, X_n$ are said to be mutually independent (or simply, independent) if

$$\mathbb{P}\{X_1 = x_1, \ldots, X_n = x_n\} = \mathbb{P}\{X_1 = x_1\} \cdots \mathbb{P}\{X_n = x_n\}$$

for all $x_1, \ldots, x_n \in \mathbb{R}$.

Examples

Example 21

Suppose that the number of people who enter a post office on a given day is a Poisson random variable with parameter λ. If each person who enters the post office is a male with probability p and a female with probability $1 - p$, then the number of males and females entering the post office are independent Poisson random variables with respective parameters λp and $\lambda(1 - p)$.

Solution.

Let X and Y denote the number of males and females that enter the post office, respectively. We shall show the independence of X and Y by showing that

$$\mathbb{P}\{X = i, Y = j\} = \mathbb{P}\{X = i\}\,\mathbb{P}\{Y = j\}.$$

Note that

$$\mathbb{P}\{X = i, Y = j\} = \mathbb{P}\{X = i, Y = j | X + Y = i + j\}\,\mathbb{P}\{X + Y = i + j\}.$$

Because $X + Y$ is the total number of people who enter the post office, it follows that

$$\mathbb{P}\{X + Y = i + j\} = e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}.$$

Given that $i + j$ people enter the post office, it follows that $X \sim$ Binomial$(i + j, p)$, and therefore,

$$\mathbb{P}\{X = i, Y = j | X + Y = i + j\} \binom{i+j}{i} p^i (1 - p)^j.$$

Hence,

$$\mathbb{P}\{X = i, Y = j\} = e^{-\lambda} \frac{(\lambda p)^i}{i!j!} [\lambda(1 - p)]^j = \left\{ \frac{e^{-\lambda p} (\lambda p)^i}{i!} \right\} \left\{ \frac{e^{-\lambda(1-p)} [\lambda(1-p)]^j}{j!} \right\}.$$

■

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Example 22 (Buffon's needle problem)

A table is ruled with equidistant parallel lines a distance D apart. A needle of length L, where $L \leqslant D$, is randomly thrown on the table. What is the probability that the needle will intersect one of the lines (the other possibility being that the needle will be completely contained in the strip between two lines)?

Figure: Buffon (1707–1788)

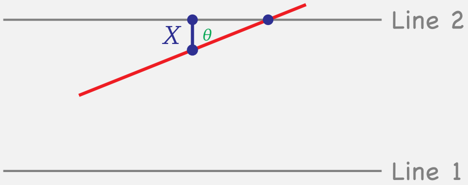
Solution.

Let us determine the position of the needle by specifying (1) the distance X from the middle point of the needle to the nearest parallel line and (2) the angle θ between the needle and the projected line of length X. Then, the needle will intersect a line if the hypotenuse of the right triangle is less than $L/2$:

$$\frac{X}{\cos\theta} < \frac{L}{2}.$$

Now, notice that X varies between 0 and $D/2$ and θ between 0 and $\pi/2$, and assume that they are independent. Hence,

$$\mathbb{P}\{X < \frac{L}{2}\cos\theta\} = \iint_{x<L\cos y/2} f_X(x) f_\theta(y) dx dy = \frac{2L}{\pi D}.$$ ■



An important proposition

Proposition 23

The continuous (discrete) random variables X and Y are independent if and only if their joint probability density (mass) function can be expressed as

$$f(x,y) = h(x)g(y) \qquad -\infty < x, y < \infty.$$

Proof (Continuous case only).

(i) $\Longleftarrow$. First, note that independence implies that the joint density is the product of the marginal densities of X and Y, so the preceding factorization will hold when the random variables are independent.

(ii) $\Longrightarrow$. Suppose that $f(x, y) = h(x)g(y)$, then

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy \\ &= \left(\int_{-\infty}^{\infty} h(x) dx \right) \left(\int_{-\infty}^{\infty} g(y) dy \right) := C_1 C_2. \end{aligned}$$

Also,

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = C_2 h(x), \quad f_Y(y) = C_1 g(y),$$

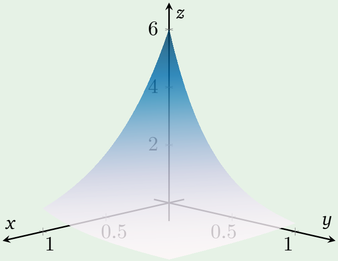
and it follows that $f(x, y) = f_X(x) f_Y(y)$, which proves the result. ∎

Example 24

If the joint density function of X and Y is

$$f(x, y) = \begin{cases} 6e^{-2x-3y} & 0 < x, y < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Are these random variables independent?

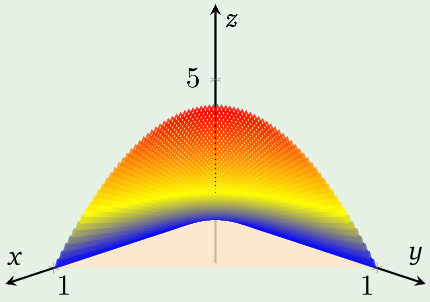


Example 25

What if the joint density function is

$$f(x, y) = 24xy$$
$$0 < x < 1, 0 < y < 1, 0 < x + y < 1$$

and is equal to 0 otherwise?



SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Example 26

Let X, Y, Z be independent and uniformly distributed over $(0, 1)$.
Compute $\mathbb{P}\{X \geqslant YZ\}$.

Solution.

$$
\begin{aligned}
\mathbb{P}\{X \geqslant YZ\} &= \iiint_{x \geqslant yz} f_X(x) f_Y(y) f_Z(z)\, dx dy dz \\
&= \int_0^1 \left(\int_0^1 (1 - yz) dy \right) dz \\
&= \int_0^1 (1 - \frac{1}{2} z) dz \\
&= 1 - \frac{1}{4} = \frac{3}{4}.
\end{aligned}
$$

■

Question

Whether X and Y are independent?

(i) $f(x,y) = \begin{cases} 6xy^2 & 0 < x < 1, 0 < y < 1, \\ 0 & \text{otherwise} \end{cases}$

(ii) $f(x,y) = \begin{cases} 12y^2 & 0 \leqslant y \leqslant x \leqslant 1, \\ 12y^2 & 0 \leqslant y \leqslant x \leqslant 1, \end{cases}$

(iii) $f(x,y) = \begin{cases} 6e^{-2x-3y} & x > 0, y > 0, \\ 0 & \text{otherwise} \end{cases}$

(iv) $f(x,y) = \begin{cases} x^2 + \frac{xy}{3} & 0 < x < 1, 0 < y < 2, \\ 0 & \text{otherwise} \end{cases}$

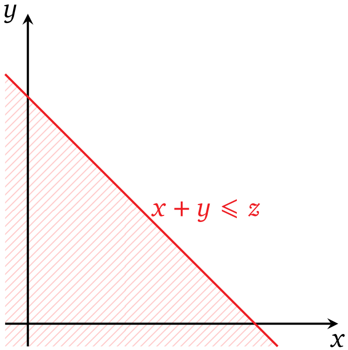
Sums of independent random variables

The distribution of $X + Y$

- Suppose that X and Y are independent, continuous random variables having pdf f_X and f_Y.

- The cdf of $X + Y$ is

$$
\begin{aligned}
F_{X+Y}(z) &= \mathbb{P}\{X + Y \leqslant z\} \\
&= \iint\limits_{x+y \leqslant z} f_X(x) f_Y(y) dx dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} f_X(x) dx \right) f_Y(y) dy \\
&= \int_{-\infty}^{\infty} F_X(z - y) dF_Y(y) \\
&= (F_X \star F_Y)(z).
\end{aligned}
$$



Definition 27 (Convolutions)

The convolution of two distribution functions F and G is defined to be

$$(F \star G)(z) = \int_{-\infty}^{\infty} F(z-y)dG(y).$$

The convolution of two density functions f and g is defined as

$$(f * g)(z) = \int_{-\infty}^{\infty} f(z-y)g(y)dy.$$

Proposition 28

If X and Y are independent, then the pdf of $X + Y$ is

$$f_{X+Y}(z) = (f_X * f_Y)(z).$$

Identically distributed variables

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 29

Random variables X and Y are said to be identically distributed if $F_X = F_Y$.
If X and Y are independent and identically distributed, then we write X and Y are i.i.d. for brevity.

Remark

- X and Y may not be defined on the same probability space. For example, let $\Omega_1 = \{$, $\}$, and $\Omega_2 = (0, 1]$, and define the classical probabilities on them. Let $X($$) = 1$, $X($$) = 0$, and let $Y(\omega) = 1$ if $\omega \in (0, 1/2]$ and 0 otherwise. Then, X and Y are identically distributed.

- Identically random variables are not necessarily independent. To see this, let Y be defined as above, and let $Z = 1 - Y$. Then, Y and Z are identically distributed but not independent.

Sum of two independent uniform random variables

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 30

If X and Y are i.i.d. Uniform random variables over $(0,1)$, calculate the probability density of $X+Y$.

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Proof.

Note that

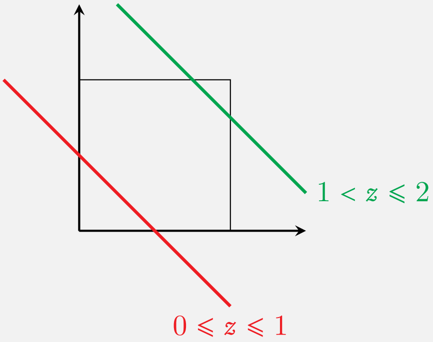
$$f_X(x) = f_Y(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

we obtain

$$\begin{aligned}(f_X * f_Y)(z) &= \int_0^1 f_X(z-y) f_Y(y) dy = \int_0^1 \mathbf{I}(0 < z - y < 1) dy \\ &= \int_{0 \vee (z-1)}^{1 \wedge z} dy = \begin{cases} z & 0 \leqslant z \leqslant 1, \\ 2 - z & 1 < z \leqslant 2, \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

■

Sum of two independent uniform random variables



Sum of n independent uniform variables

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 31

Let $X_1, X_2, \dots, X_n$ be independent $U(0,1)$ random variables, and let F_n be the distribution function of $X_1 + \dots X_n$.

(i) Prove by induction that

$$F_n(x) = \frac{x^n}{n!} \quad 0 \leqslant x \leqslant 1.$$

(ii) Let

$$N = \min\{n : X_1 + \cdots + X_n > 1\}.$$

Prove that $\mathbb{E}[N] = e$.

Sums of exponential random variables

- Let X and Y are i.i.d. Exp(λ) random variables. What is the density function of $X+Y$?
- We have for $z>0$,

$$\begin{aligned} f_{X+Y}(z) &= \int_0^\infty \lambda^2 e^{-\lambda(z-y)} e^{-\lambda y} \mathbf{I}(z-y>0)dy \\ &= \lambda^2 e^{-\lambda z} \int_0^z 1 dy \\ &= \lambda^2 z e^{-\lambda z}. \end{aligned}$$

- Continue this argument, we have for independent random variables $X_1, \ldots, X_n \sim$ Exp(λ),

$$f_{X_1+\cdots+X_n}(z) = \frac{\lambda e^{-\lambda z}(\lambda z)^{n-1}}{(n-1)!}, \quad 0 < z < \infty.$$

Gamma distribution

Definition 32 (Gamma distribution)

A random variable is said to have a Gamma distribution with parameters k (shape) and λ (scale), written as $X \sim \text{Gamma}(k, \lambda)$, if the density function of X is given by

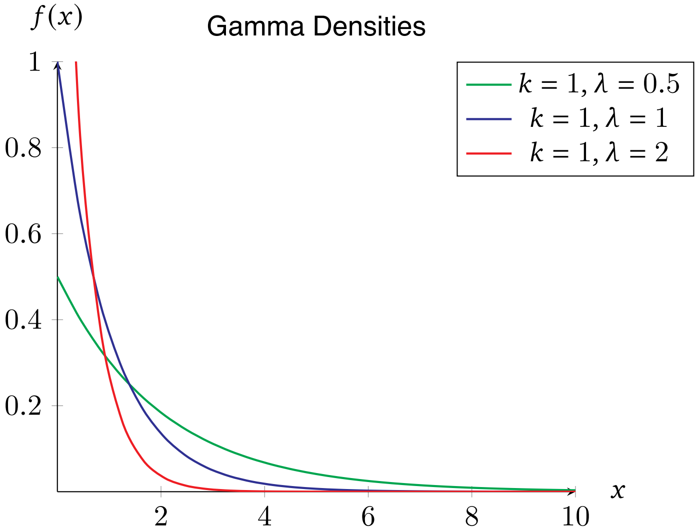
$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, \quad 0 < x < \infty.$$

Remark

Note that $\text{Exp}(\lambda) = \text{Gamma}(1, \lambda)$. Moreover, if $X_1, \ldots, X_n \sim \text{Exp}(\lambda)$, then

$$X_1 + \cdots + X_n \sim \text{Gamma}(n, \lambda).$$

Gamma densities



Proposition 33

If X and Y are independent gamma random variables with respective parameters (k_1, λ) and (k_2, λ), then $X + Y$ is a gamma random variable with parameters $(k_1 + k_2, \lambda)$.

Proof.

We obtain

$$
\begin{aligned}
f_{X+Y}(z) &= \frac{1}{\Gamma(k_1)\Gamma(k_2)} \int_0^z \{\lambda e^{-\lambda(z-y)}[\lambda(z-y)]^{k_1-1}\}\{\lambda e^{-\lambda y}(\lambda y)^{k_2-1}\}dy \\
&= \frac{\lambda^{k_1+k_2} e^{-\lambda z}}{\Gamma(k_1)\Gamma(k_2)} \int_0^z (z-y)^{k_1-1} y^{k_2-1} dy \\
&= \frac{\lambda^{k_1+k_2} e^{-\lambda z} z^{k_1+k_2-1}}{\Gamma(k_1)\Gamma(k_2)} \int_0^1 (1-x)^{k_1-1} x^{k_2-1} dx = \frac{\lambda^{k_1+k_2} z^{k_1+k_2-1} e^{-\lambda z}}{\Gamma(k_1+k_2)}.
\end{aligned}
$$

■

χ^2-distribution

Definition 34 (χ^2-distribution)

If $Z_1, \ldots, Z_n$ are independent $N(0,1)$ random variables, then $Y = \sum_{i=1}^{n} Z_i^2$ is said to have the χ^2 distribution (chi-squared distribution) with n degrees of freedom.

Normal random variables

Proposition 35

If $X_1, \ldots, X_n$ are independent random variables with respective parameters μ_i, σ_i^2 for $i = 1, \ldots, n$, then for any $\alpha_i \in \mathbb{R}$,

$$\sum_{i=1}^{n} \alpha_i X_i \sim N\left(\sum_{i=1}^{n} \alpha_i \mu_i, \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2\right).$$

Example 36

A basketball team will play a 44-game season. Twenty-six of these games are against class A teams and 18 are against class B teams. Suppose that the team will win each game against a class A team with probability 0.4 and will win each game against a class B team with probability 0.7. Suppose also that the results of the different games are independent. Approximate the probability that

(a) the team wins 25 games or more;

(b) the team wins more games against class A teams than it does against class B teams.

Proposition 37

If $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$ are independent random variables, then

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

Proof.

$$\begin{aligned}
\mathbb{P}\{X + Y = n\} &= \sum_{k=0}^{n} \mathbb{P}\{X = k, Y = n - k\} \\
&= \sum_{k=0}^{n} \mathbb{P}\{X = k\}\,\mathbb{P}\{Y = n - k\} && \text{by independence} \\
&= \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n.
\end{aligned}$$

∎

Some commonly used sums

X_1	Ber(p)	Binom(n, p)	Poisson(λ_1)	Exp(λ)	$\Gamma(k_1, \lambda)$	$N(\mu_1, \sigma_1^2)$
X_2	Ber(p)	Binom(m, p)	Poisson(λ_2)	Exp(λ)	$\Gamma(k_2, \lambda)$	$N(\mu_2, \sigma_2^2)$
$X_1 + X_2$	Binom(2, p)	Binom($m + n$, p)	Poisson($\lambda_1 + \lambda_2$)	$\Gamma(2, \lambda)$	$\Gamma(k_1 + k_2, \lambda)$	$N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Table: Sum of Two Independent Variables

Conditional distributions

- Conditional probability
- Conditional distribution: given the information of X, how does Y behave?

Conditional distribution: Discrete case

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 38

If X and Y are discrete random variables with joint pmf $p(x,y)$, the conditional probability mass function of X given $Y=y$ is defined as

$$p_{X|Y}(x|y) = \mathbb{P}\{X = x | Y = y\} = \frac{p(x,y)}{p_Y(y)}$$

for all values of y such that $p_Y(y) > 0$.

The conditional distribution function of X given that $Y = y$ is defined as

$$F_{X|Y}(x|y) = \mathbb{P}\{X \leqslant x | Y = y\} = \sum_{u \leqslant x} p_{X|Y}(u|y).$$

Example 39

Suppose that the joint probability mass function of X and Y is given by

$$p(0,0) = 0.4, \quad p(0,1) = 0.2, \quad p(1,0) = 0.1, \quad p(1,1) = 0.3.$$

Calculate the conditional pmf of X given $Y = 1$.

Solution.

Note that

$$p_Y(1) = \sum_x p(x,1) = p(0,1) + p(1,1) = 0.5.$$

Hence,

$$p_{X|Y}(0|1) = \frac{p(0,1)}{p_Y(1)} = \frac{2}{5}, \qquad p_{X|Y}(1|1) = \frac{p(1,1)}{p_Y(1)} = \frac{3}{5}.$$ ■

Poisson distribution

If $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$ are independent random variables, then

$$p_{X|X+Y}(k|n) = ?$$

Conditional distribution: continuous case

Definition 40

If X and Y have a joint pdf $f(x, y)$, then the conditional pdf of X given $Y = y$ is defined as

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} \quad \text{for all } y \text{ such that } f_Y(y) > 0.$$

Example 41

The joint density of X and Y is given by

$$f(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

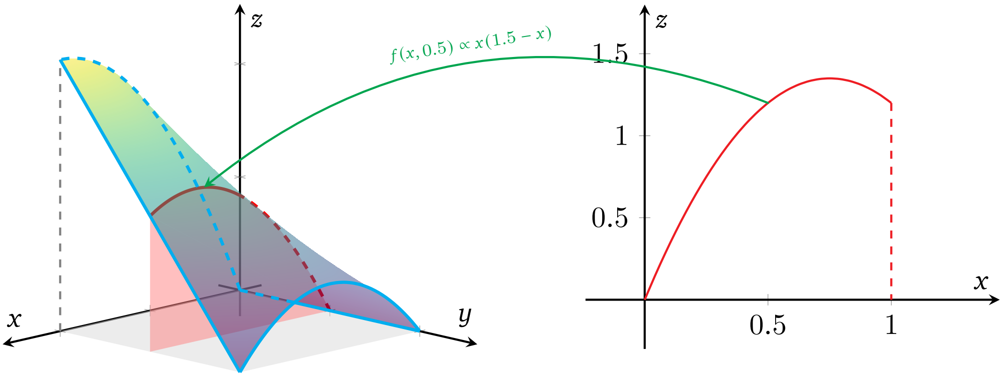
Compute the conditional pdf of X given that $Y = y$, where $0 < y < 1$.

Solution.

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{x(2-x-y)}{\int_0^1 x(2-x-y)dx}$$
$$= \frac{6x(2-x-y)}{4-3y}, \quad 0 < x, y < 1.$$

■

Figure



Joint probability distribution of $g(X, Y)$

- Let X and Y be jointly continuous random variables with joint pdf $f_{X,Y}$.
- In may questions, we want to obtain the joint distribution $f_{U,V}$ of X and Y, where

$$U = g(X, Y), \quad V = h(X, Y).$$

- Assume that g and h both have continuous partial derivatives at all points (x, y) and

$$J(x, y) := \begin{vmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \\ \dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial y} \end{vmatrix} \neq 0 \quad \text{at all points } (x, y).$$

Theorem 42

Under these two conditions, it can be shown that the random variables U and V are jointly continuous with joint pdf

$$f_{U,V}(u,v) = f_{X,Y}(x,y)|J(x,y)|^{-1},$$

where $x = \widetilde{g}(u,v)$ and $y = \widetilde{h}(u,v)$ is the inverse of g and h.

Example 43

If X and Y be jointly continuous random variables with probability density function $f_{X,Y}$. Let $U = X + Y$ and $V = X - Y$. Find the joint pdf of U and V.

Solution.

Let $g(x, y) = x + y$ and $h(x, y) = x - y$. Then,

$$J(x, y) = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2.$$

Also, solve equation system

$$\begin{cases} u = x + y \\ v = x - y \end{cases} \implies \begin{cases} x = \dfrac{u+v}{2} \\ y = \dfrac{u-v}{2} \end{cases}$$

Therefore,

$$f_{U,V}(u, v) = \frac{1}{2} f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right).$$

■

Example 44

Let (X, Y) denote a random point in the plane, and assume that the rectangular coordinates X and Y are independent standard normal random variables. Let (R, Θ) be the polar coordinate representation of (X, Y). Find the joint pdf of (R, θ).

Solution.

For any positive x and y, let

$$r = g_1(x,y) = \sqrt{x^2+y^2}, \quad \theta = g_2(x,y) = \tan^{-1}\frac{y}{x}.$$

Then,

$$\frac{\partial g_1}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}, \qquad \frac{\partial g_1}{\partial y} = \frac{y}{\sqrt{x^2+y^2}},$$

$$\frac{\partial g_2}{\partial x} = -\frac{y}{x^2+y^2}, \qquad \frac{\partial g_2}{\partial y} = \frac{x}{\sqrt{x^2+y^2}},$$

and

$$J(x,y) = \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix} = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{r}.$$

Note that

$$f_{X,Y}(x, y|X > 0, Y > 0) = \frac{2}{\pi}e^{-(x^2+y^2)/2}, \quad x > 0, y > 0.$$

Then,

$$f_{R,\Theta}(r, \theta|X > 0, Y > 0) = \frac{2}{\pi}re^{-r^2/2}, \quad 0 < \theta < \pi/2, 0 < r < \infty.$$

Similarly, we can show that

$$f_{R,\Theta}(r, \theta|X < 0, Y > 0) = \frac{2}{\pi}re^{-r^2/2}, \quad \frac{\pi}{2} < \theta < \pi, 0 < r < \infty,$$
$$f_{R,\Theta}(r, \theta|X < 0, Y < 0) = \frac{2}{\pi}re^{-r^2/2}, \quad \pi < \theta < 3\pi/2, 0 < r < \infty,$$
$$f_{R,\Theta}(r, \theta|X > 0, Y < 0) = \frac{2}{\pi}re^{-r^2/2}, \quad 3\pi/2 < \theta < 2\pi, 0 < r < \infty.$$

Combining, we obtain $f_{R,\Theta}(r, \theta) = \frac{1}{2\pi}re^{r^2/2}, \quad 0 < \theta < 2\pi, 0 < r < \infty.$ ■

Example 45

If $X \sim \text{Gamma}(\alpha, \lambda)$ and $Y \sim \text{Gamma}(\beta, \lambda)$ are independent, compute the joint density of $U = X + Y$ and $V = \frac{X}{X+Y}$.

Solution.

The joint pdf of X and Y is given by

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \frac{\lambda e^{-\lambda y}(\lambda y)^{\beta-1}}{\Gamma(\beta)} \\ &= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda(x+y)} x^{\alpha-1} y^{\beta-1}. \end{aligned}$$

Examples

Now, let $g_1(x,y) = x + y$ and $g_2(x,y) = x/(x+y)$, then

$$\frac{\partial g_1}{\partial x} = \frac{\partial g_1}{\partial y} = 1, \quad \frac{\partial g_2}{\partial x} = \frac{y}{(x+y)^2}, \quad \frac{\partial g_2}{\partial y} = -\frac{x}{(x+y)^2},$$

so

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & \frac{-x}{(x+y)^2} \end{vmatrix} = -\frac{1}{x+y}.$$

Examples

Moreover, as the equations $u = x+y$ and $v = x/(x+y)$ have solutions $x = uv$ and $y = u(1-v)$, we see that

$$f_{U,V}(u,v) = f_{X,Y}(uv, u(1-v))|u|$$

$$= \underbrace{\frac{\lambda e^{-\lambda u}(\lambda u)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)}}_{\Gamma(\alpha+\beta,\lambda)} \underbrace{\frac{v^{\alpha-1}(1-v)^{\beta-1}\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}}_{\text{Beta}(\alpha,\beta)}.$$

■

Further reading

[1] Sheldon M. Ross (谢尔登・M. 罗斯).

A first course in probability (概率论基础教程): Chapter 6.

10th edition (原书第十版), 机械工业出版社

[2] Sheldon M. Ross (谢尔登・M. 罗斯).

Introduction to Probability Models (概率模型导论): Chapter 3.

12th edition (原书第十二版), 人民邮电出版社

Random variables

- We have introduced the concept of random variables.
- We also defined cdf and pmf:

$$\mathbb{P}\{X \leqslant x\}, \quad \mathbb{P}\{X = x\}, \quad \mathbb{P}\{a < X \leqslant b\}, \ldots$$

- Remember that $X : \Omega \longrightarrow \mathbb{R}$, for any $a < b$,

$$\{a < X \leqslant b\} = \{\omega : a < X(\omega) \leqslant b\} \quad \text{need to be measurable}$$

Borel sets and Borel field

Definition 46

The Borel field, denoted by $\mathscr{B}(\mathbb{R})$, or simply $\mathscr{B}$, is the smallest σ-field containing all intervals $\{(a, b] : -\infty < a < b < \infty\}$.

Remark

Observe that

$$\{a\} = \prod_{n=1}^{\infty}(a - \frac{1}{n}, a],$$
$$(a, b) = (a, b] \setminus \{b\},$$
$$[a, b] = (a, b] \cup \{a\},$$
$$[a, b) = ((a, b] \setminus \{b\}) \cup \{a\}$$

all belongs to $\mathscr{B}$.

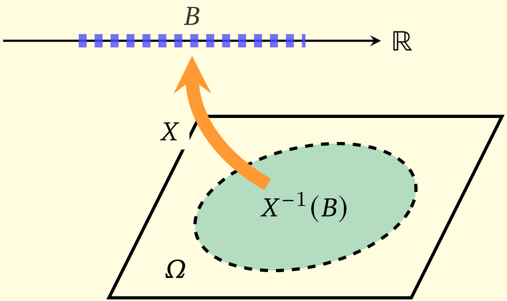
A formal definition of Random variable

Definition 47 (Random variable)

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let $X : \Omega \longrightarrow \mathbb{R}$ be a function. If

$$\{X \in B\} = \{\omega : X(\omega) \in B\} = X^{-1}(B) \in \mathscr{F}$$

for any Borel set $B \in \mathscr{B}$, then X is called a random variable.



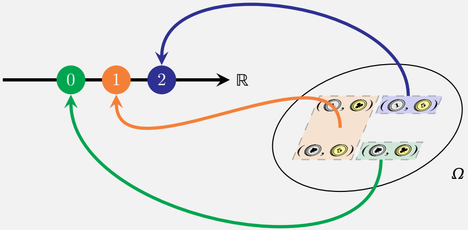
Examples

Example 48

Flip two different coins (coin 1), (coin 5). Let X be numbers of heads. Discuss that why X is a random variable.

Solution.

The sample space $\Omega = \{(\text{1}, \text{5}), (\text{1}, \text{♣}), (\text{♠}, \text{5}), (\text{♠}, \text{♣})\}$. The simga-field $\mathscr{F}$ is the power set of Ω, which contains all subsets of Ω. ■



Random vectors

Definition 49 (Random vectors)

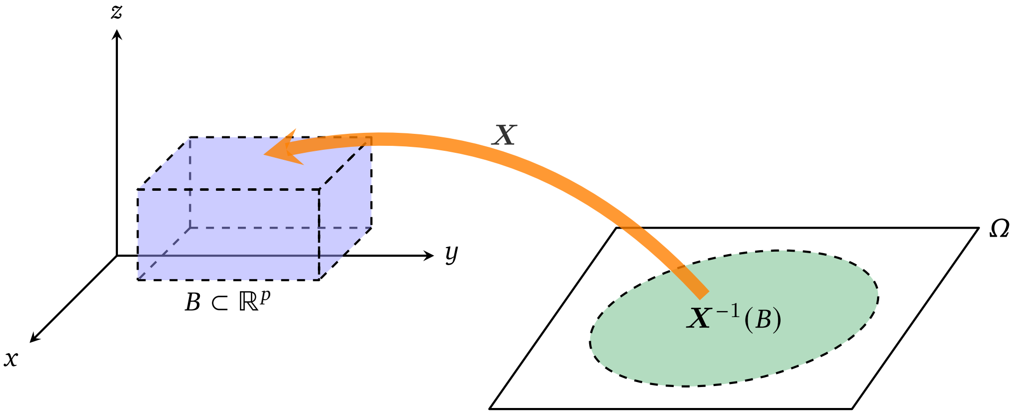
If random variables $X_1, \ldots, X_p$ are defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$, then $\boldsymbol{X} = (X_1, \ldots, X_p)$ is called a p-dimensional random vector.

Remark

Note that for any $x_1, \ldots, x_p$,

$$\{\omega : X_1(\omega) \leqslant x_1, \ldots, X_p(\omega) \leqslant x_p\} = \bigcap_{j=1}^{p} \{\omega : X_j(\omega) \leqslant x_j\} \leqslant \mathscr{F}.$$

n-dimensional random vecters



Joint (cumulative) distribution function

Definition 50

The joint p-dimensional cdf of $\boldsymbol{X}$ is defined as

$$F(x_1, \ldots, x_p) = \mathbb{P}\{X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_p \leqslant x_p\}.$$

Proposition 51

The joint cdf F satisfies the following properties:

(i) **Monotonicity**: For every $x_j \leqslant y_j$,

$$F(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_p) \leqslant F(x_1, \ldots, x_{j-1}, y_j, x_{j+1}, \ldots, x_p).$$

(ii) For any $x_1, \ldots, x_p$,

$$\lim_{x_j \to -\infty} F(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_p) = 0, \quad F(\infty, \infty, \ldots, \infty) = 1.$$

Joint pmf

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 52

For jointly discrete random vector $\boldsymbol{X}$, the joint pmf of $\boldsymbol{X}$ is defined as

$$p(x_1, \ldots, x_p) = \mathbb{P}\{X_1 = x_1, \ldots, X_p = x_p\}.$$

Multinomial distribution

Definition 53

Assume that there are r possibilities in a trial: $A_1, \ldots, A_r$, and $\mathbb{P}(A_j) = p_j$ for $j = 1, \ldots, r$, $p_1 + \cdots + p_r = 1$. If we repeat this trial n times independently, and let X_j denote the number of occurrences of A_j, then

$$\mathbb{P}\{X_1 = k_1, \ldots, X_r = k_r\} = \frac{n!}{k_1! \cdots k_r!} p_1^{k_1} \ldots p_r^{k} \quad \text{for } k_1, \ldots, k_r \geqslant 0 \text{ and } k_1 + \cdots + k_r = n.$$

Joint pdf

Definition 54

If there exists a nonnegative function $f(x_1, \ldots, x_p)$ such that

$$F(x_1, \ldots, x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \ldots du_p,$$

then f is called the joint probability density function (pdf) of $\boldsymbol{X}$. Here, f satisfies the following two conditions:

(i) $f(x_1, \ldots, x_p) \geqslant 0$,

(ii) $\displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_p) dx_1 \ldots dx_p = 1.$

Uniform distribution

Definition 55

If $A \subset \mathbb{R}^p$ such that $\lambda(A) < \infty$, where $\lambda(A)$ is the Lebesgue measure (volume) of A. If $\boldsymbol{X}$ has the joint pdf

$$f(x_1, \ldots, x_p) = \begin{cases} \dfrac{1}{\lambda(A)} & \boldsymbol{x} \in A, \\ 0 & \text{otherwise,} \end{cases}$$

then $\boldsymbol{X}$ is said to have the uniform distribution on A.

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Definition 56

Let $X_1, \ldots, X_p$ be p random variables. If for any $B_1, \ldots, B_p \in \mathscr{B}(\mathbb{R})$,

$$\mathbb{P}\{X_1 \in B_1, \ldots, X_p \in B_p\} = \prod_{j=1}^{p} \mathbb{P}\{X_j \in B_j\}.$$

Remark

Specially, if $B_j = (-\infty, x_j]$ for each $j = 1, \ldots, p$, then

$$F(x_1, \ldots, x_p) = F_{X_1}(x_1) \ldots F_{X_p}(x_p).$$

Properties

Proposition 57

If $X_1, \ldots, X_p$ are independent, then for any $1 \leqslant r \leqslant p$, any subset $\{X_{i_1}, \ldots, X_{i_r}\}$ of $\{X_1, \ldots, X_p\}$ are also independent.

Proof.

We only prove $X_1, \ldots, X_{p-1}$ are independent: Note that

$$\begin{aligned} \mathbb{P}\{X_1 \in B_1, \ldots, X_{p-1} \in B_{p-1}\} &= \mathbb{P}\{X_1 \in B_1, \ldots, X_{p-1} \in B_{p-1}, X_p \in \mathbb{R}\} \\ &= \left(\prod_{j=1}^{p-1} \mathbb{P}\{X_j \in B_j\}\right) \mathbb{P}\{X_p \in \mathbb{R}\} \\ &= \prod_{j=1}^{p-1} \mathbb{P}\{X_j \in B_j\}. \end{aligned}$$

■

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proposition 58

If $X_1, \ldots, X_p$ are independent, then for any Borel functions $g_1, \ldots, g_p : \mathbb{R} \to \mathbb{R}$, random variables $g_1(X_1), \ldots, g_p(X_p)$ are also independent.

Proof.

For any Borel sets $B_1, \ldots, B_p$,

$$\begin{aligned}
\mathbb{P}\{g_1(X_1) \in B_1, \ldots, g_p(X_p) \in B_p\} &= \mathbb{P}\{X_1 \in g_1^{-1}(B_1), \ldots, X_p \in g_p^{-1}(B_p)\} \\
&= \prod_{j=1}^{p} \mathbb{P}\{X_j \in g_j^{-1}(B_j)\} \\
&= \prod_{j=1}^{p} \mathbb{P}\{g_j(X_j) \in B_j\}.
\end{aligned}$$

∎

Example 59

Let $X_1, \ldots, X_n$ be independent continuous random variables, with the common distribution function $F(x)$ and probability density function $f(x)$. Let $Y = \max_{1 \leqslant i \leqslant n} X_i$. Find the distribution function and probability density function of Y.

Solution.

Note that

$$\begin{aligned}\mathbb{P}\{Y \leqslant x\} &= \mathbb{P}\{\max(X_1, \ldots, X_n) \leqslant x\} \\ &= \mathbb{P}\{X_1 \leqslant x, X_2 \leqslant x, \ldots, X_n \leqslant x\} \\ &= \prod_{i=1}^{n} \mathbb{P}\{X_i \leqslant x\} = \prod_{i=1}^{n} F(x) = [F(x)]^n.\end{aligned}$$

Then,

$$F_Y(x) = [F(x)]^n.$$

Moreover,

$$f_Y(x) = \frac{d}{dx} F_Y(x) = n[F(x)]^{n-1} f(x).$$

■

Minimum of independent random variables

Example 60

Let $W = \min_{1 \leqslant i \leqslant n} X_i$. Find the distribution function of W.

Solution.

Note that

$$\begin{aligned}\mathbb{P}\{W > x\} &= \mathbb{P}\{\min(X_1, \ldots, X_n) > x\} \\ &= \mathbb{P}\{X_1 > x, X_2 > x, \ldots, X_n > x\} \\ &= \prod_{i=1}^{n} \mathbb{P}\{X_i > x\} \\ &= \prod_{i=1}^{n} (1 - F(x)),\end{aligned}$$

then

$$F_W(x) = 1 - \mathbb{P}\{W > x\} = 1 - \prod_{i=1}^{n} (1 - F(x)).$$

■

Distribution of (W, Y)

Example 61

Find the joint distribution of (W, Y).

Solution.

If $w \geqslant y$, then $\{Y \leqslant y\} \subset \{W \leqslant w\}$, and thus

$$\begin{aligned} F_{W,Y}(w, y) &= \mathbb{P}\{W \leqslant w, Y \leqslant y\} \\ &= \mathbb{P}\{Y \leqslant y\} = [F(y)]^n. \end{aligned}$$

If $w > y$,

$$\begin{aligned} F_{W,Y}(w, y) &= \mathbb{P}\{W \leqslant w, Y \leqslant y\} \\ &= \mathbb{P}\{Y \leqslant y\} - \mathbb{P}\{W > w, Y \leqslant y\}. \end{aligned}$$

Solution

For the second term, the event means that all of the variables $X_1, \ldots, X_n$ are between w and y, which means that

$$\begin{aligned}\mathbb{P}\{W > w, Y \leqslant y\} &= \mathbb{P}\{w < X_1 \leqslant y, \ldots, w < X_n \leqslant y\} \\ &= \prod_{i=1}^{n} \mathbb{P}\{w < X_i \leqslant y\} = \prod_{i=1}^{n} (F(y) - F(w)) = [F(y) - F(w)]^n.\end{aligned}$$

Therefore, if $w > y$, then

$$F_{W,Y}(w, y) = [F(y)]^n - [F(y) - F(w)]^n.$$

The joint pdf of (W, Y) is

$$\begin{aligned}f_{W,Y}(w, y) &= \frac{d^2}{dwdy} F_{W,Y}(w, y) \\ &= \begin{cases} 0 & w \geqslant y \\ n(n-1)[F(y) - F(w)]^{n-2} f(w) f(y) & w < y. \end{cases}\end{aligned}$$

■

Example 62

Let $R = \max(X_1, \ldots, X_n) - \min(X_1, \ldots, X_n)$. The variable R is called the range of the data.

In descriptive statistics, range is the size of the smallest interval which contains all the data and provides an indication of statistical dispersion. Since it only depends on two of the observations, it is most useful in representing the dispersion of small data sets.

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Solution.

Note that for $r \geqslant 0$,

$$
\begin{aligned}
F_R(r) &= \mathbb{P}\{R \leqslant r\} \\
&= \iint_{y-w\leqslant r} f_{W,Y}(w,y)\,dw\,dy \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{w+r} f_{W,Y}(w,y)\,dy\right)dw \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{r} f_{W,Y}(w,w+u)\,du\right)dw \\
&= \int_{-\infty}^{r}\left(\int_{-\infty}^{\infty} f_{W,Y}(w,w+u)\,dw\right)du.
\end{aligned}
$$

Solution (Cont'd).

Therefore,

$$f_R(r) = \frac{d}{dr}F_R(r) = \int_{-\infty}^{\infty} f_{W,Y}(w, w+r)dw$$
$$= n(n-1)\int_{-\infty}^{\infty} [F(x+r) - F(x)]^{n-2} f(x)f(x+r)dx. \qquad \blacksquare$$