Chapter 4 Joint Distributions

Jingrui Sun

Email: sunjr@sustech.edu.cn

Southern University of Science and Technology

September 8, 2024

Outline I

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
   - Part 4.2.1: An Example
   - Part 4.2.2: Joint and Marginal Probability Mass Function
   - Part 4.2.3: Independent Random Variable
   - Part 4.2.4: Joint Cumulative Distribution Function
   - Part 4.2.5: Marginal Cumulative Distribution Function
3. Part 4.3: Continuous Random Variable
   - Part 4.3.1: Joint Cumulative Distribution Function
   - Part 4.3.2: Marginal Cumulative Distribution Function
   - Part 4.3.3: Joint Probability Density Function
   - Part 4.3.4: Marginal Probability Density Function
   - Part 4.3.5: Independence
   - Part 4.3.6: Calculations of Two Continuous r.v.s
   - Part 4.3.7: Bivariate Normal Random Vector

Outline II

4 Part 4.4: Transformation of Random Vector
- Part 4.4.1: Question
- Part 4.4.2: Procedure
- Part 4.4.3: Example
- Part 4.4.4: Further Remarks
- Part 4.4.5: $\max\{X, Y\}$ and $\min\{X, Y\}$

5 Part 4.5: General Case
- Part 4.5.1: Joint c.d.f.
- Part 4.5.2: One-dimensional marginal c.d.f.
- Part 4.5.3: Joint p.d.f. of continuous random vectors
- Part 4.5.4: Notes on General Random Vectors
- Part 4.5.5: One-dimensional marginal p.d.f.

6 Part 4.6: Summary
- Part 4.6.1: Summary of Chapter 2
- Part 4.6.2: Summary of Chapter 3

Outline III

- Part 4.6.3: Summary of Chapter 4

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

Part 4.1: Introduction

Consider two random variables, X and Y, say.
New question: the relation between X and Y?
In general, n random variables $X_1, X_2, \ldots, X_n$?

In many cases, we have to consider several random variables together since there are relations among them.

An example: In ecological studies, several species have to be considered together: prey and predators.

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

1. Example:

Random experiment: A fair coin is tossed three times.

Sample space:

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Let X denote the number of heads on the first toss and Y the total number of heads.

Part 4.2.1: An Example

Then we see that

(1) Both X and Y are random variables: depends upon the outcomes.

For example: let $\omega_1 = \{HHH\}$, then $X(\omega_1) = 1$; $Y(\omega_1) = 3$; $\omega_5 = \{THH\}$, then $X(\omega_5) = 0$; $Y(\omega_5) = 2$; $\cdots$

(2) Both X and Y are discrete random variables.

All possible values of $X : \{0, 1\}$.

All possible values of $Y : \{0, 1, 2, 3\}$.

(3) X and Y are defined on the same sample space and there exists a relation between them.

2. "Joint probability mass function"

Now consider the events A and B:

A: "the number of heads on the 1st toss is zero";

B: "the total number of heads is two".

then

$$A = \{\omega \in \Omega; X(\omega) = 0\} = (X = 0) = \{THH, THT, TTH, TTT\},$$
$$B = \{\omega \in \Omega; Y(\omega) = 2\} = (Y = 2) = \{HHT, HTH, THH\}.$$

Part 4.2.1: An Example

Easy to see, the intersection of A and B is the event

$$A \cap B = \{THH\}.$$

Hence $P(A \cap B) = \frac{1}{8}$. (Equally likely!!) We write it as

$$P(A \cap B) = P\{(X = 0) \cap (Y = 2)\} = \frac{1}{8}.$$

More simply, just write it as

$$P(X = 0, Y = 2) = \frac{1}{8}.$$

Part 4.2.1: An Example

Similarly, let

$$A_1 : \text{“number of heads on } \underline{\text{the first toss}} \text{ is } \underline{\text{one}}\text{”},$$
$$B_1 : \text{“number of } \underline{\text{total}} \text{ heads is } \underline{\text{two}}\text{”}.$$

Then

$$A_1 = (X = 1) = \{HHH, HHT, HTH, HTT\},$$
$$B_2 = (Y = 2) = \{HHT, HTH, THH\},$$
$$A_1 \cap B_2 = \{HHT, HTH\},$$
$$P(A_1 \cap B_2) = \frac{2}{8}, \quad \text{(Equally likely!!)}$$

Part 4.2.1: An Example

Similarly, (check these!)

$$P(X=0,Y=0)=\frac{1}{8}, \qquad P(X=0,Y=1)=\frac{2}{8},$$
$$P(X=0,Y=2)=\frac{1}{8}, \qquad P(X=0,Y=3)=0,$$
$$P(X=1,Y=0)=0, \qquad P(X=1,Y=1)=\frac{1}{8},$$
$$P(X=1,Y=2)=\frac{2}{8}, \qquad P(X=1,Y=3)=\frac{1}{8}.$$

The above are the <u>all possibilities</u>.

Part 4.2.1: An Example

Table:

x \ y	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

Now, we denote $P(X = 0, Y = 0) = p(0, 0)$, then

$$p(0, 0) = \frac{1}{8};$$

Also, denote $P(X = 0, Y = 2) = p(0, 2)$, then $p(0, 2) = \frac{1}{8}$.

Part 4.2.1: An Example

In general, let

$$p(x, y) = P(X = x, Y = y),$$

we get a function $p(x, y)$ depending upon two real variables.

The function $p(\cdot\,,\cdot)$ is called the joint probability mass function of X and Y.

This function gives the full information about the two random variables X and Y.

Part 4.2.2: Joint and marginal p.m.f.'s

1. Joint probability mass function:

(1) **Definition:** Suppose that X and Y are two discrete random variables defined on the same sample space and they take on values $x_1, x_2, \cdots, x_i, \cdots$ (<u>for X</u>) and $y_1, y_2, \cdots, y_j, \cdots$ (<u>for Y</u>), respectively.

Then the following function of two variables

$$p(x_i, y_j) = P(X = x_i, Y = y_j), \quad \text{(for all i and j)} \qquad (4.2.1)$$

is called the joint probability mass function $(\mathrm{p.m.f.})$ of the random variables X and Y (or random vector (X, Y)).

(2) Properties of Joint $\mathrm{p.m.f.}$ p for random vector (X, Y):

(i) $p(x_i, y_j) \geq 0, \quad \forall x_i,\ \forall y_j;$

(ii) $\displaystyle\sum_{y_j}\sum_{x_i} p(x_i, y_j) = 1.$

Part 4.2.2: Joint and marginal p.m.f.'s

2. Marginal probability mass function:

(1) In considering the two random variables X and Y, the random variable X itself has its own distribution, which is called the marginal p.m.f. of X.

Similarly, the marginal p.m.f. of Y.

Hence, for random variables X and Y, we have two marginal p.m.f.'s. So we use different notations:

$$p_X(\cdot) \quad \text{and} \quad p_Y(\cdot).$$

Part 4.2.2: Joint and marginal p.m.f.'s

(2) Suppose X and Y are two discrete random variables defined on the same sample space and taking all possible values: (for X) $x_1, x_2, \cdots, x_i, \cdots$, (and for Y) $y_1, y_2, \cdots, y_i, \cdots$, respectively.

Then the marginal p.m.f. of X is a function defined by

$$p_X(x_i) = P(X = x_i), \quad i = 1, 2, \cdots \tag{4.2.2}$$

Similarly, the marginal p.m.f. of Y:

$$p_Y(y_i) = P(Y = y_i), \quad i = 1, 2, \cdots$$

Part 4.2.2: Joint and marginal p.m.f.'s

(3) Relation: (between joint p.m.f. and marginal p.m.f.)

  (i) joint p.m.f. determines marginal p.m.f.'s;

  (ii) (But usually) all marginal p.m.f.'s can not determine the joint p.m.f.;

  (iii) For some special cases, all marginal p.m.f.'s can decide the joint p.m.f..

No wonder! Since joint p.m.f.: full information of X and Y, including the relation between X and Y.

But marginal p.m.f.'s can only give the information about the X and Y, individually.

However, if the relation is known then ....

(4) How to get Marginal p.m.f.'s from joint one.

Random variables X and Y.

Possible values:

$$X: x_1, x_2, \cdots, x_i, \cdots$$

$$Y: y_1, y_2, \cdots, y_j, \cdots$$

Joint p.m.f. $p(x_i, y_j)$:

$$p_X(x_i) = ?, \quad p_Y(y_j) = ?$$

Part 4.2.2: Joint and marginal p.m.f.'s

Conclusions:

$$p_X(x_i) = \sum_j p(x_i, y_j), \tag{4.2.3}$$

$$p_Y(y_j) = \sum_i p(x_i, y_j). \tag{4.2.4}$$

Can you give a proof by yourself?

Part 4.2.2: Joint and marginal p.m.f.'s

(5) Example: (See the example in 4.2.1)

Joint p.m.f.

$x \backslash y$	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

$$
\begin{aligned}
p_X(0) = P(X = 0) &= \sum_{k=0}^{3} P(X = 0, Y = k) \\
&= p(0,0) + p(0,1) + p(0,2) + p(0,3) \\
&= \frac{1}{2},
\end{aligned}
$$

Part 4.2.2: Joint and marginal p.m.f.'s

$$\begin{aligned} p_X(1) = P(X=1) &= \sum_{k=0}^{3} P(X=1, Y=k) \\ &= p(1,0) + p(1,1) + p(1,2) + p(1,3) \\ &= \frac{1}{2}. \end{aligned}$$

So the marginal p.m.f. of X is

$$p_X(0) = \frac{1}{2}, \quad p_X(1) = \frac{1}{2}.$$

(We see: the row sum !!)

Similarly, the marginal p.m.f. of Y is

$$p_Y(0) = \frac{1}{8}, \quad p_Y(1) = \frac{3}{8}, \quad p_Y(2) = \frac{3}{8}, \quad p_Y(3) = \frac{1}{8}.$$

(Column sum!!)

$x \backslash y$	0	1	2	3	$p_X(\cdot)$
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{2}$
$p_Y(\cdot)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

Part 4.2.3: Independent Random Variable

1. Recall: Two events A and B are called independent if $P(A \cap B) = P(A) \cdot P(B)$.

2. **Definition**: Let X and Y be two discrete random variables. Suppose all the possible values of X and Y are:

$$X : x_1, x_2, x_3, \cdots, x_i, \cdots$$
$$Y : y_1, y_2, y_3, \cdots, y_j, \cdots$$

We call X and Y are independent random variables if for all x_i and y_i, we have

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j), \tag{4.2.5}$$

Part 4.2.3: Independent Random Variable

that is,

$$p(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j),$$

or

$$p(x, y) = p_X(x) \cdot p_Y(y), \quad \forall x, \ \forall y, \tag{4.2.6}$$

where $p(x, y)$ is the joint probability mass function and $p_X(x)$ and $p_Y(y)$ are marginal probability mass functions.

3. Meaning

   (1) For fixed x_i and y_j, (4.2.5) means the two events $(X = x_i)$ and $(Y = y_j)$ are independent events.

   Since (4.2.5) holds true for all x_i and y_j and so there are many pairs of independent events.

   (2) (4.2.5) and (4.2.6) are totally the same.

   (3) (4.2.6) tells us, if X and Y are independent, then all the marginal p.m.f.'s can determine the joint p.m.f.

Part 4.2.3: Independent Random Variable

4. Remark: (Important)

   (1) Joint p.m.f. $\Rightarrow$ all marginal p.m.f.'s;

   (2) Joint p.m.f. $\not\Leftarrow$ all marginal p.m.f.'s;

   (3) For independent r.v.'s X and Y,

$$\text{Joint p.m.f.} \Leftrightarrow \text{all marginal p.m.f.'s}$$

Part 4.2.4: Joint Cumulative Distribution Function

1. **Definition**: Suppose X and Y are two random variables. The function $F(x, y)$ defined by

$$F(x, y) = P(X \leq x, Y \leq y), \tag{4.2.7}$$

where $-\infty < x, y < +\infty$, is called the joint c.d.f. (cumulative distribution function) of the random variables X and Y.

More exactly, the joint c.d.f. of X and Y should be denoted by $F_{(X,Y)}(x, y)$.

Part 4.2.4: Joint Cumulative Distribution Function

2. Properties of joint c.d.f. $F(x, y) \triangleq F_{(X,Y)}(x, y)$:

   (i) For fixed x, $F(x, y)$ is an increasing function of y;
   For fixed y, $F(x, y)$ is an increasing function of x.

   (ii) For fixed x, $F(x, y)$ is a right continuous function of y;
   For fixed y, $F(x, y)$ is a right continuous function of x.

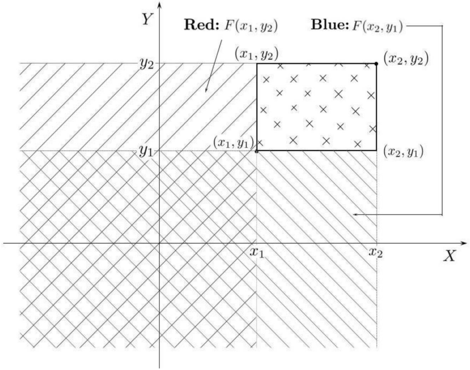
   (iii) $F(+\infty, +\infty) \triangleq \lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y) = 1.$

(iv) $\forall y, \lim_{x \to -\infty} F(x, y) = 0, \quad [F(-\infty, y) = 0];$
$\forall x, \lim_{y \to -\infty} F(x, y) = 0, \quad [F(x, -\infty) = 0].$

(v) For $x_1 < x_2, y_1 < y_2$,

$$\begin{aligned} &P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\ &\quad = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \geq 0. \end{aligned}$$

3. Remark on Properties of joint c.d.f.:
   (1) The properties stated in the above are true for <u>any</u> two r.v.'s (no matter discrete, continuous or even mixed ones).
   (2) Conclusion (v) has clear geometric interpretation:



Part 4.2.4: Joint Cumulative Distribution Function

4. Relationship between joint p.m.f. and joint c.d.f. for discrete random vector.

$$
\begin{aligned}
F_{(X,Y)}(x,y) &= P\{X \le x, Y \le y\} \\
&= \sum_{y_j \le y} \sum_{x_i \le x} P(X = x_i, Y = y_j) \\
&= \sum_{y_j \le y} \sum_{x_i \le x} p(x_i, y_j).
\end{aligned}
$$

Hence, essentially speaking, all the calculations regarding the probabilities of the discrete random vector (X, Y) can be done in terms of joint p.m.f..

For example,

$$P\{X \geq x, Y \leq y\} = \sum_{y_j \leq y} \sum_{x_i \geq x} p(x_i, y_j).$$

Part 4.2.5: Marginal c.d.f.

1. Suppose the random variables X and Y have the joint c.d.f. $F(x, y)$. Then the c.d.f. of X, i.e.

$$F_X(x) = P(X \le x) \tag{4.2.13}$$

is called the marginal c.d.f. of X.

Similarly, the c.d.f. of Y, i.e.

$$F_Y(y) = P(Y \le y) \tag{4.2.14}$$

is called the marginal c.d.f. of Y.

2. Relation with joint c.d.f.:

   Suppose the random variables X and Y have the joint c.d.f. $F(x, y)$, then the marginal c.d.f. $F_X(x)$ of X can be obtained by

$$F_X(x) = \lim_{y \to \infty} F(x, y). \tag{4.2.15}$$

   Similarly, the marginal c.d.f. of Y:

$$F_Y(y) = \lim_{x \to \infty} F(x, y). \tag{4.2.16}$$

Part 4.2.5: Marginal c.d.f.

3. Independence (Can prove that)

   Two r.v.s X and Y are independent iff for any x and y,

   $$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y), \tag{4.2.17}$$

   i.e.,

   $$F(x, y) = F_X(x) \cdot F_Y(y), \tag{4.2.18}$$

   where $F(x, y)$ is the joint c.d.f. of X and Y, and $F_X(x)$ and $F_Y(y)$ are the marginal c.d.f.'s.

   Again, (4.2.18) tells us that for independent r.v.s, the joint c.d.f. can be determined by the marginal c.d.f.'s.

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

Part 4.3.1: Joint Cumulative Distribution Function

For joint c.d.f. totally the same as the discrete case.

1. **Definition:** Suppose X and Y are two continuous random variables. Then the function

$$F_{(X,Y)}(x, y) \triangleq P(X \leq x, Y \leq y) \tag{4.3.1}$$

or more simply,

$$F(x, y) \triangleq P(X \leq x, Y \leq y) \tag{4.3.2}$$

is called the joint c.d.f. of X and Y.

2. **Properties of Joint c.d.f.:** The joint c.d.f. of two continuous random variables X and Y has the following properties:

   (i) For any fixed x, $F(x, y)$ is an increasing function of y;
   For any fixed y, $F(x, y)$ is an increasing function of x;

   (ii) $F(+\infty, +\infty) \triangleq \lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y) = 1$;

   (iii) For any fixed y, $F(-\infty, y) \triangleq \lim_{x \to -\infty} F(x, y) = 0$;
   For any fixed x, $F(x, -\infty) \triangleq \lim_{y \to -\infty} F(x, y) = 0$;

(iv) For $x_1 < x_2$, $y_1 < y_2$,

$$
\begin{aligned}
&P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\
&\quad = F(x_2, y_2) - F(x_1 y_2) - F(x_2, y_1) + F(x_1 y_1) \geq 0.
\end{aligned}
$$

Furthermore, ($\because$ X and Y are both continuous r.v.'s)

$$
\begin{aligned}
&P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\
&\quad = P(x_1 \leq X < x_2, y_1 \leq Y < y_2) \\
&\quad = P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) \\
&\quad = \cdots
\end{aligned}
$$

(iv) $F(x, y)$ is a continuous function of x and y. Furthermore, $\partial^2 F(x, y)/\partial x \partial y$ exists (almost everywhere).

1. **Definition:**

$$F_X(x) = P(X \leq x),$$
$$F_Y(y) = P(Y \leq y),$$

   are called the marginal c.d.f..

2. We have

$$F_X(x) = \lim_{y \to +\infty} F(x, y),$$
$$F_Y(y) = \lim_{x \to +\infty} F(x, y),$$

   where $F(x, y)$ is the joint c.d.f. of X and Y and $F_X(x)$ and $F_Y(y)$ are marginal c.d.f.'s.

Part 4.3.2: Marginal C.D.F.

3. **Independence:** Two continuous random variables X and Y are called independent if for all $x, y \in (-\infty, +\infty)$,

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y),$$

i.e.,

$$F_{(X,Y)}(x, y) = F_X(x) \cdot F_Y(y).$$

Hence, again, for independent continuous random variables, the joint c.d.f. can be determined by the marginal c.d.f.'s.

Part 4.3.3: Joint P.D.F.

Recall the definition of joint c.d.f.:

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

and similar to the single continuous random variable, we may consider the joint p.d.f.

Part 4.3.3: Joint P.D.F.

1. **Definition:** Suppose the joint c.d.f. of the continuous random variables X and Y is

$$F(x, y) = P(X \leq x, Y \leq y).$$

Then the joint p.d.f. $f(x, y)$ is defined by

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

and thus by the basic formula in Calculus (Double integral!!)

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

2. Properties of joint p.d.f. $f(x, y)$:

   Let $f(x, y)$ be the joint p.d.f. of X and Y, then

   (a) $f(x, y) \geq 0$ for all $x, y \in \mathbb{R}$;

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

Part 4.3.3: Joint P.D.F.

Recall for single continuous r.v. X, the p.d.f. $f(x)$ satisfies

$$f(x) \geq 0 \text{ for all } x, \quad \int_{-\infty}^{+\infty} f(x)dx = 1.$$

Note also that for discrete random variables X and Y, the joint p.m.f. has the similar properties:

(i) $p(x, y) \geq 0$;

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

Part 4.3.4: Marginal probability density function

1. **Definition:** Suppose X and Y are two continuous random variables with joint p.d.f. $f(x, y)$. The p.d.f. $f_X(x)$ of the r.v. X is called the marginal p.d.f. of X.

   Similarly, $f_Y(y)$, the marginal p.d.f. of Y.

Part 4.3.4: Marginal probability density function

2. Relation with joint p.d.f.

   Let $f(x, y)$ be the joint p.d.f. of X and Y. Then the marginal probability density functions $f_X(x)$ and $f_Y(y)$ can be obtained by

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy,$$

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

   (Compare with the discrete case).

Part 4.3.5: Independence

1. **Conclusion:** Two random variables X and Y are independent iff

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

The definition tells us that X and Y are independent iff

$$F(x, y) = F_X(x) \cdot F_Y(y),$$

where $F(x, y)$, $F_X(x)$, $F_Y(y)$ are (joint/marginal) c.d.f.'s and $f(x, y)$, $f_X(x)$, $f_Y(y)$ are (joint/marginal) p.d.f.'s.

2. Remark (Important): For two continuous (and discrete) random variables X and Y,

   (1) Joint p.d.f. $\Rightarrow$ Marginal ones;

   (2) In general, Joint p.d.f. $\overset{\text{Not enough}}{\Longleftarrow}$ Marginal ones;

   (3) However, if X and Y are independent, then

   Joint p.d.f. $\Longleftrightarrow$ Marginal ones.

1. **An important conclusion:**

   Suppose the random vector (X, Y) has joint pdf $f(x, y)$. Then for any set G in the plane, we have

$$P\{(X, Y) \in G\} = \iint_G f(x, y)dxdy. \qquad (*)$$

2. Remarks:

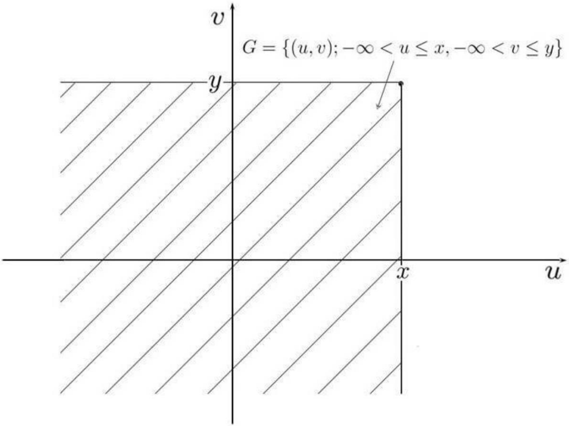
(1) The meaning of the ω-set $\{(X, Y) \in G\}$ is

$$\{\omega \in \Omega; (X(\omega), Y(\omega)) \in G\};$$

(2) The definition of joint c.d.f. could be viewed as a special case of $(*)$ (the above basic formula). Indeed, recall

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

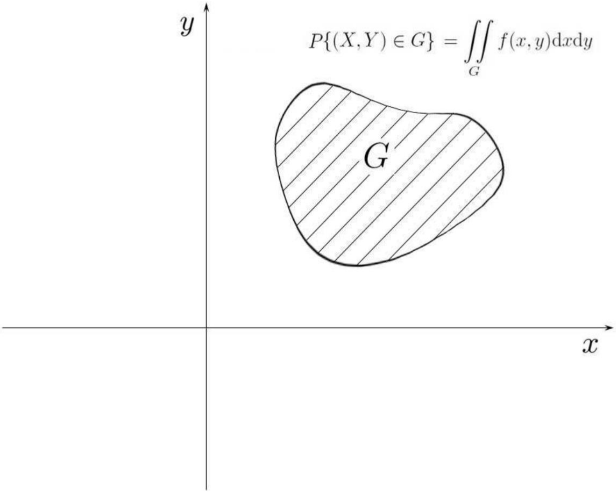
Now let $G = \{(u, v); -\infty < u \leq x, -\infty < v \leq y\}$.



Then

$$\begin{aligned}
F(x,y) &= P\{X \le x, Y \le y\} \\
&= P\{(X,Y) \in G\} \\
&= \iint\limits_{G} f(x,y)dxdy \\
&= \iint\limits_{G} f(u,v)dudv \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv.
\end{aligned}$$

(4) Geometric "meaning" (recall single r.v. case!)



Example 1. Suppose the random vector (X, Y) has joint p.d.f.

$$f(x,y) = \begin{cases} cxy^2, & \text{if } 0 \le x \le 1, 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the constant c.

(b) Find the two marginal p.d.f.'s: $f_X(x)$ and $f_Y(y)$.

(c) Find the probability $P(X \le \frac{1}{2}, Y \le \frac{1}{2})$.

(d) Find $P(X < Y)$.

(e) Are X and Y independent?

Solution: (a) Let

$$G = \{(x, y), 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

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

$$\begin{aligned} 1 &= \int_0^1 \int_0^1 f(x, y)dxdy + \iint_{\mathbb{R}^2 \setminus G} f(x, y)dxdy \\ &= \int_0^1 \int_0^1 cxy^2 dxdy + \iint_{\mathbb{R}^2 \setminus G} 0dxdy \\ &= c \cdot \int_0^1 y^2 \left[\int_0^1 xdx\right] dy \\ &= c\frac{1}{2} \int_0^1 y^2 dy = \frac{c}{6} \quad \Rightarrow \quad \underline{c = 6}. \end{aligned}$$

(b) $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$.

- If $x < 0$ or $x > 1$, then

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{+\infty} 0dy = 0.$$

Part 4.3.6: Calculations of Two Continuous r.v.s

- If $0 \leq x \leq 1$, then

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{0} f(x,y)dy + \int_{0}^{1} f(x,y)dy + \int_{1}^{+\infty} f(x,y)dy \\
&= \int_{-\infty}^{1} 0dy + \int_{0}^{1} 6xy^2dy + \int_{1}^{+\infty} 0dy \\
&= 6x \cdot \int_{0}^{1} y^2dy \\
&= 6x \cdot \left[\frac{y^3}{3}\right]\Bigg|_0^1 \\
&= 2x.
\end{aligned}$$

Therefore

$$f_X(x) = \begin{cases} 2x, & \text{if } 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Similarly,

- If $y < 0$ or $y > 1$, then $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx = 0$.
- If $0 \leq y \leq 1$, then $f_Y(y) = \int_0^1 6xy^2 dx = 6y^2 \cdot \left[\frac{x^2}{2}\right]\Big|_0^1 = 3y^2$.

$$\therefore \quad f_Y(y) = \begin{cases} 3y^2, & \text{if } 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

(e) Now we can see that $\underline{X \text{ and } Y \text{ are independent}}$.

Indeed, let $G = \{(x, y); 0 \leq x \leq 1, 0 \leq y \leq 1\}$. Then

- for any $(x, y) \notin G$,

$$f(x, y) = 0, \quad f_X(x) = f_Y(y) = 0,$$

- for $(x, y) \in G$,

$$f(x, y) = 6xy^2, \quad f_X(x) = 2x, \quad f_Y(y) = 3y^2.$$

Hence $f(x, y) = f_X(x) \cdot f_Y(y)$, $\boxed{\forall (x, y) \in \mathbb{R}^2.}$

(c) $P(X \leq \frac{1}{2}, Y \leq \frac{1}{2})$. Let

$$\hat{G} = \left\{(x, y); -\infty < x \leq \frac{1}{2}, -\infty < y \leq \frac{1}{2}\right\},$$

and

$$G = \{(x, y); 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

$$
\begin{aligned}
\therefore P\left(X \le \frac{1}{2}, Y \le \frac{1}{2}\right) &= P\{(X, Y) \in \hat{G}\} \\
&= P\{(X, Y) \in \hat{G} \cap G\} + P\left\{(X, Y) \in \hat{G} \cap G^c\right\} \\
&= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} 6xy^2 dxdy + 0 \\
&= 6 \int_0^{\frac{1}{2}} y^2 \left[\int_0^{\frac{1}{2}} xdx\right] dy \\
&= \frac{1}{32}.
\end{aligned}
$$

Of course, you could do as follows (totally the same)

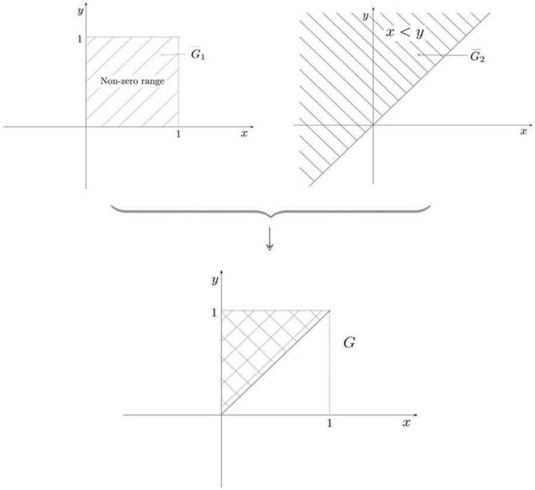
$$P\left(X \leq \frac{1}{2}, Y \leq \frac{1}{2}\right) = F\left(\frac{1}{2}, \frac{1}{2}\right) = \int_{-\infty}^{\frac{1}{2}} \int_{-\infty}^{\frac{1}{2}} f(x, y) dxdy$$
$$= \int_{0}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} 6xy^2 dxdy + 0 = \frac{1}{32}.$$

(d) How about $P\{X < Y\}$? Let

$$G = \{(x, y) \in \mathbb{R}^2; 0 \leq x \leq 1, 0 \leq y \leq 1, x < y\}.$$

Then $P\{X < Y\} = \iint\limits_G f(x, y) dxdy$.

Part 4.3.6: Calculations of Two Continuous r.v.s



$$
\begin{aligned}
\therefore \quad P\{X < Y\} &= \iint\limits_{G} f(x,y)dxdy \\
&= \int_0^1 \int_0^y 6xy^2 dxdy = 6\int_0^1 y^2 \left[\int_0^y xdx\right] dy \\
&= 6\int_0^1 y^2 \left[\frac{x^2}{2}\right]\Bigg|_0^y dy = 6 \cdot \int_0^1 y^2 \cdot \frac{1}{2} y^2 dy \\
&= 3\int_0^1 y^4 dy = 3\left[\frac{y^5}{5}\right]\Bigg|_0^1 = \frac{3}{5}.
\end{aligned}
$$

Example 2: Suppose the joint p.d.f. of (X, Y) is given by

$$f(x,y) = \begin{cases} xe^{-(x+xy)}, & \text{if } x \geq 0, \text{ and } y \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

(Easy to check that $f(x,y) \geq 0$ and

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_{0}^{+\infty}\int_{0}^{+\infty} xe^{-(x+xy)}dxdy = 1).$$

(a) Find the two marginal p.d.f.'s and see whether X and Y are independent;

(b) Find the two marginal c.d.f.'s $F_X(x)$ and $F_Y(y)$.

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

First consider $f_X(x)$.

If $x \leq 0$, then $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{+\infty} 0dy = 0$.

Part 4.3.6: Calculations of Two Continuous r.v.s

If $x > 0$, then

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{0} f(x,y)dy + \int_{0}^{+\infty} f(x,y)dy \\
&= \int_{-\infty}^{0} 0dy + \int_{0}^{+\infty} xe^{-x(1+y)}dy \\
&= x\int_{0}^{+\infty} e^{-x} \cdot e^{-xy}dy \\
&= xe^{-x}\int_{0}^{+\infty} e^{-xy}dy = xe^{-x} \cdot \left[-\frac{e^{-xy}}{x}\right]\Bigg|_{0}^{+\infty} \\
&= xe^{-x} \cdot \frac{1}{x} = e^{-x}.
\end{aligned}$$

Hence

$$f_X(x) = \begin{cases} e^{-x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(Exponentially distributed)

For $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx$, we see that

if $y \le 0$, we still have $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx = \int_{-\infty}^{+\infty} 0dx = 0$,

while if $y > 0$, then

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{0} f(x,y)dx + \int_{0}^{+\infty} f(x,y)dx \\
&= \int_{-\infty}^{0} 0dx + \int_{0}^{+\infty} xe^{-(1+y)x}dx \\
&= -\frac{1}{1+y}\int_{0}^{+\infty} xde^{-x(1+y)} \text{Integration by parts!!} \\
&= -\frac{1}{1+y}\left\{xe^{-x(1+y)}\Big|_{x=0}^{x=+\infty} - \int_{0}^{+\infty} e^{-x(1+y)}dx\right\} \\
&= \frac{1}{(1+y)^2}\left[-e^{(1+y)x}\right]\Bigg|_{x=0}^{x=+\infty} = \frac{1}{(1+y)^2}.
\end{aligned}
$$

Hence

$$f_Y(y) = \begin{cases} \frac{1}{(1+y)^2}, & \text{if } y \geq 0, \\ 0, & \text{if } y < 0. \end{cases}$$

Therefore X and Y are NOT independent: Because for $x > 0$, $y > 0$,

$$\begin{aligned} f(x, y) &= xe^{-(x+xy)} \\ &\neq \frac{e^{-x}}{(1+y)^2} = f_X(x) f_Y(y). \end{aligned}$$

(b) We have two ways to find the two c.d.f.'s.

Method 1. Use the two p.d.f.'s. For example,

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

Method 2. For $F_X(x)$, use our general formula. For $a \geq 0$,

$$\begin{aligned} F_X(a) &= P\{X \leq a\} = P\{0 \leq X \leq a\} \quad (\because P\{X \leq 0\} = 0) \\ &= P\{0 \leq X \leq a, -\infty < Y < +\infty\} \\ &= P\{0 \leq X \leq a, 0 \leq Y < +\infty\} \quad (\because y \geq 0) \end{aligned}$$

$$
\begin{aligned}
F_X(a) &= \int_0^a \int_0^{+\infty} f(x,y)dydx = \int_0^a \int_0^{+\infty} xe^{-(x+xy)}dydx \\
&= \int_0^a xe^{-x}\left[\int_0^{+\infty} e^{-xy}dy\right]dx = \int_0^a xe^{-x}\cdot\frac{1}{x}dx \\
&= \int_0^a e^{-x}dx = 1-e^{-a};
\end{aligned}
$$

while if $a<0$, then $F_X(a)=0$.

Hence

$$F_X(x) = \begin{cases} 1 - e^{-x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

Similarly, we could get

$$F_Y(y) = \begin{cases} 1 - \frac{1}{1+y}, & \text{if } y \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Example 3. The joint p.d.f. of X and Y is given by

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

Find the p.d.f. of the random variable $\frac{X}{Y}$.

(Easy to see, $f(x,y) \geq 0, \forall x, y$, and $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)dxdy =$

$$\int_0^{+\infty} \int_0^{+\infty} f(x,y)dxdy = \int_0^{+\infty} \int_0^{+\infty} e^{-(x+y)}dxdy = 1.)$$

Solution. Let $Z = \frac{X}{Y}$, we try to find the c.d.f. of Z.

Let $F_Z(z)$ be the c.d.f. of Z, then

$$F_Z(z) = P\{Z \leq z\} = P\left\{\frac{X}{Y} \leq z\right\}$$
$$= \iint\limits_{\frac{x}{y} \leq z} f(x, y)dxdy.$$

Now, if $z \leq 0$, then clearly, either "$x \leq 0$ and $y > 0$ " or " $x > 0$ and $y < 0$", and thus $f(x, y) = 0$

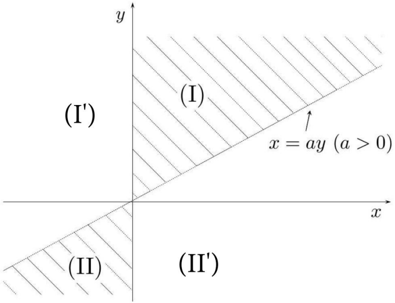
$$\Rightarrow \quad F_Z(z) = 0 \quad \text{for } z \leq 0.$$

Hence only need to consider $z > 0$.

For the notational convenience, let $z = a$ and hence $a > 0$, what is the area G for which $\frac{x}{y} \leq a$? See below:

$$\text{If } y > 0, \text{ then } \frac{x}{y} \leq a \quad \Leftrightarrow \quad x \leq ay$$
$$\text{If } y < 0, \text{ then } \frac{x}{y} \leq a \quad \Leftrightarrow \quad x \geq ay.$$

Hence, $\left\{(x, y); \frac{x}{y} \leq a\right\}$ is as follows:



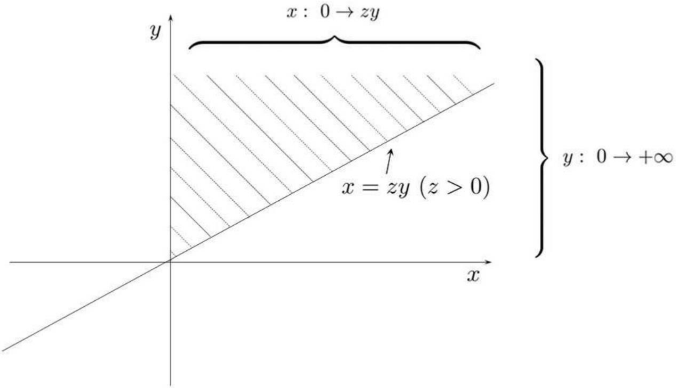
Therefore (on I', II, II', f(x,y)=0)

$$F_Z(z) = \iint\limits_{\frac{x}{y} \leq z} f(x,y)dxdy$$

$$= \iint\limits_{(I)} f(x,y)dxdy = \iint\limits_{\substack{x>0, y>0 \\ \frac{x}{y} \leq z}} e^{-(x+y)}dxdy$$

Hence $F_Z(z) = \iint\limits_{\substack{x>0, y>0 \\ \frac{x}{y} \leq z}} e^{-(x+y)}dxdy$ (view z as a constant).

Part 4.3.6: Calculations of Two Continuous r.v.s



$$F_Z(z) = \iint\limits_{\substack{x>0,y>0\\x\le yz}} e^{-(x+y)}dxdy = \int_0^{\infty}\int_0^{zy} e^{-(x+y)}dxdy$$

$$= \int_0^{\infty} e^{-y}\left[\int_0^{zy} e^{-x}dx\right]dy = \int_0^{\infty} e^{-y}\left[1-e^{-zy}\right]dy$$

$$= \left[-e^{-y} + \frac{e^{-(z+1)y}}{z+1}\right]\Bigg|_{y=0}^{y=+\infty}$$

$$= \underbrace{-0+0}_{y=+\infty} - \underbrace{\left(-e^{-0} + \frac{e^{-(z+1)\times 0}}{z+1}\right)}_{y=0}$$

$$= 1 - \frac{1}{z+1}.$$

In short

$$F_Z(z) = \begin{cases} 1 - \frac{1}{z+1}, & \text{if } z > 0, \\ 0, & \text{if } z \leq 0. \end{cases}$$

Differentiation yields $\frac{d}{dz} F_Z(z) = f_Z(z) \equiv f_{\frac{X}{Y}}(z)$.

$$f_{\frac{X}{Y}}(z) = \begin{cases} \frac{1}{(z+1)^2}, & \text{if } z > 0, \\ 0, & \text{if } z \leq 0. \end{cases}$$

Part 4.3.6: Calculations of Two Continuous r.v.s

Note: we do have $f_{\frac{X}{Y}}(z) \geq 0, \forall z$ and

$$\int_{-\infty}^{+\infty} f_{\frac{X}{Y}}(z)dz = \int_{0}^{+\infty} \frac{dz}{(z+1)^2} = \left[-\frac{1}{z+1}\right]\Big|_0^{\infty} = 1.$$

Part 4.3.7: Bivariate Normal Random Vector

1.**Expression:** Two random variables X and Y are called bivariate normally distributed, if the joint p.d.f. $f(x, y)$ is given by

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \right] \right\},$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2$, and ρ are constants satisfying

$$\begin{cases} -\infty < \mu_1 < +\infty, \\ -\infty < \mu_2 < +\infty, \\ \sigma_1 > 0, \\ \sigma_2 > 0, \\ -1 < \rho < 1. \end{cases}$$

Part 4.3.7: Bivariate Normal Random Vector

[No need to remember the formula!!]

The above $f(x, y)$ is called the bivariate normal density and the five constants $\mu_1, \mu_2, \sigma_1, \sigma_2$ and ρ are called parameters (and so, five parameters).

2. Marginal p.d.f.'s

(i) First consider a special case: $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1$. Then

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[x^2 + y^2 - 2\rho xy\right]\right\}.$$

$f_Y(y) = ?$

$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}} dx.$

Note that

$$
\begin{aligned}
-\frac{1}{2(1-\rho^2)}\left[x^2+y^2-2\rho xy\right] &= -\frac{x^2-2\rho xy+\rho^2y^2+(1-\rho^2)y^2}{2(1-\rho^2)} \\
&= -\frac{(x-\rho y)^2}{2(1-\rho^2)}-\frac{y^2}{2}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}-\frac{y^2}{2}}\, dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} \cdot e^{-\frac{y^2}{2}}\, dx \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-\frac{y^2}{2}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}\, dx \\
&= \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi}\cdot\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}\, dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.
\end{aligned}
$$

Why $\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}dx = 1$??

Indeed, let $t = \frac{x-\rho y}{\sqrt{1-\rho^2}}$, then $dx = \sqrt{1-\rho^2}dt$ and hence the above is

$$\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} \cdot \sqrt{1-\rho^2}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt = 1.$$

We have got that

$$f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}, \quad -\infty < y < +\infty.$$

Thus, $Y \sim N(0,1)$.

Part 4.3.7: Bivariate Normal Random Vector

By symmetry, we can also obtain $X \sim N(0,1)$.

Note also that X and Y are independent $\Leftrightarrow$

$$
\begin{aligned}
& f(x,y) = f_X(x) \cdot f_Y(y) \\
\Leftrightarrow \quad & \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}}, \\
\Leftrightarrow \quad & \sqrt{1-\rho^2}\, e^{-\frac{x^2+y^2}{2}} = e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}}, \qquad \text{(Easy to see)} \\
\Leftrightarrow \quad & \rho = 0.
\end{aligned}
$$

(ii) The general case.

We recall that if the continuous differentiable functions

$$\begin{cases} x = x(u, v) \\ y = y(u, v) \end{cases}$$

transform the set G' in $O'uv$ to the set G in the plain Oxy, with Jacobian

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \neq 0$$

then

$$\iint_G f(x, y) dxdy = \iint_{G'} f[x(u, v), y(u, v)] \cdot |J| \, dudv$$

Part 4.3.7: Bivariate Normal Random Vector

In the general case,

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \right] \right\}.$$

We do a transformation as

$$U = \frac{X-\mu_1}{\sigma_1} \quad V = \frac{Y-\mu_2}{\sigma_2},$$

Or $u = \frac{x-\mu_1}{\sigma_1}, v = \frac{y-\mu_2}{\sigma_2} \Rightarrow \quad x = \mu_1 + \sigma_1 u, \quad y = \mu_2 + \sigma_2 v.$

$$\Rightarrow \quad \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{vmatrix} = \sigma_1\sigma_2.$$

$$
\begin{aligned}
&P\{(U,V)\in G'\}=P\{(X,Y)\in G\}\\
&\quad=\iint_G f(x,y)dxdy=\iint_{G'} f[x(u,v),y(u,v)]\cdot|J|\,dudv\\
&\quad=\iint_{G'}\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[u^2+v^2-2\rho uv\right]\right\}dudv.
\end{aligned}
$$

Hence the joint p.d.f. of (U,V) is

$$
f_{(U,V)}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{u^2+v^2-2\rho uv}{2(1-\rho^2)}\right\}.
$$

Part 4.3.7: Bivariate Normal Random Vector

Now $U \sim N(0, 1)$ we obtain

$$\boxed{X \sim N(\mu_1, \sigma_1^2)}$$

Similarly

$$\boxed{Y \sim N(\mu_2, \sigma_2^2)}$$

Hence both marginal p.d.f.'s are normal distributions.

Also, easy to see that X and Y are independent iff

$$\rho = 0.$$

3. Independence:

We see that the bivariate normally distributed random variables X and Y are independent if and only if

$$\rho = 0.$$

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

Part 4.4.1: Question

Question:

Let (X, Y) be a continuous random vector with joint p.d.f. $f_{(X,Y)}(x, y)$.

Two known functions $g_1(x, y)$ and $g_2(x, y)$.

Let $U = g_1(X, Y)$ and $V = g_2(X, Y)$, we obtain two new continuous random variables U and V.

Then joint p.d.f. of U and V? i.e., $f_{(U,V)}(u, v) =??$

Part 4.4.2: Procedure

Step 1. From the given functions $g_1(x, y)$ and $g_2(x, y)$ try to find another two functions $h_1(\cdot\,,\cdot)$ and $h_2(\cdot\,,\cdot)$ such that

$$x = h_1(u, v), \quad y = h_2(u, v).$$

Method: $\boxed{\textit{Inverting}!!}$

Step 2. Find the Jacobian of the transformation which is defined as

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u}.$$

Note that the Jacobian might be negative or zero.

Part 4.4.2: Procedure

Since we have found the function forms as

$$x = h_1(u, v), \quad y = h_2(u, v),$$

and thus

$$\frac{\partial x}{\partial u} = \frac{\partial h_1(u, v)}{\partial u}, \qquad \frac{\partial x}{\partial v} = \frac{\partial h_1(u, v)}{\partial v},$$
$$\frac{\partial y}{\partial u} = \frac{\partial h_2(u, v)}{\partial u}, \qquad \frac{\partial y}{\partial v} = \frac{\partial h_2(u, v)}{\partial v}.$$

Part 4.4.2: Procedure

Step 3. Now, the unknown $f_{(U,V)}(u,v)$ is given by the known joint p.d.f. $f_{(X,Y)}(x,y)$ as

$$f_{(U,V)}(u,v) = f_{(X,Y)}(h_1(u,v), h_2(u,v)) \left| \frac{\partial(x,y)}{\partial(u,v)} \right|,$$

i.e., replacing x by the inverted function $h_1(u,v)$ and replacing y by the inverted function $h_2(u,v)$ in the original known joint p.d.f. $f_{(X,Y)}(x,y)$, and then times the absolute value of the Jacobian obtained in Step 2.

Step 4. Pay attention to the region of non-zero area.

Part 4.4.3: Example

Example: $X \sim \exp(\lambda)$, $Y \sim \exp(\lambda)$ (the same λ), X and Y are independent. Define $U = X - Y$ and $V = X + Y$.

Find the joint p.d.f. of the two new r.v.s U and V. Also check the non-zero region and the marginal density of V.

Step 1. What is the joint p.d.f. of X and Y?

$$X \sim \exp(\lambda) \Rightarrow f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

$$Y \sim \exp(\lambda) \Rightarrow f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & \text{if } y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence (since X and Y are independent)

$$f_{(X,Y)}(x,y) = \begin{cases} \lambda^2 e^{-\lambda(x+y)}, & \text{if } x>0, y>0, \\ 0, & \text{otherwise.} \end{cases}$$

Step 2. Find $h_1(u,v)$ and $h_2(u,v)$ by inverting!

Originally $u = x - y, v = x + y$,

$$\Rightarrow \quad x = \frac{u+v}{2}, \quad y = \frac{v-u}{2},$$

i.e., $x = h_1(u,v) = \frac{u+v}{2}, \quad y = h_2(u,v) = \frac{v-u}{2}$.

Part 4.4.3: Example

Step 3. Jacobian $\frac{\partial(x,y)}{\partial(u,v)}$?

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

But $\frac{\partial x}{\partial u} = \frac{1}{2}$, $\frac{\partial x}{\partial v} = \frac{1}{2}$, $\frac{\partial y}{\partial u} = -\frac{1}{2}$, $\frac{\partial y}{\partial v} = \frac{1}{2}$,

and so $\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2} \times \frac{1}{2} - (\frac{1}{2})(-\frac{1}{2}) = \frac{1}{2}$.

(In general, the Jacobian may be a function of u and v)

Part 4.4.3: Example

Step 4. The joint p.d.f. of u and v is given by

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

That is, in the form of $f_{(X,Y)}(x,y)$ given above, just replace x by $\frac{u+v}{2}$ and replace y by $\frac{v-u}{2}$ and then times $\left|\frac{1}{2}\right|$, we get the form of $f_{(U,V)}(u,v)$ as a function of u and v only.

Part 4.4.3: Example

Step 5. However, pay attention to the fact that for some x and y, $f_{(X,Y)}(x, y)$ may be zero.

In our example, only for $x > 0$ and $y > 0$, will $f_{(X,Y)}(x, y)$ be non-zero.

Thus we must pay attention to the non-zero region of (u, v) on the (u, v) plane.

Part 4.4.3: Example

Non-zero region: $x > 0$ and $y > 0$.

$$0 < x = \frac{u+v}{2} \quad \Leftrightarrow \quad u + v > 0,$$
$$0 < y = \frac{v-u}{2} \quad \Leftrightarrow \quad v - u > 0, \text{ i.e. } v > u.$$

Hence

$$f_{(U,V)}(u,v) = \begin{cases} \frac{1}{2} f_{(X,Y)}\left(\frac{u+v}{2}, \frac{v-u}{2}\right), & \text{if } u+v > 0 \text{ and } v > u, \\ 0, & \text{otherwise.} \end{cases}$$

Part 4.4.3: Example

Step 6. Finally,

$$f_{(U,V)}(u,v) = \begin{cases} \frac{1}{2}\lambda^2 e^{-\lambda v}, & \text{if } u+v>0 \text{ and } v>u, \\ 0, & \text{otherwise.} \end{cases}$$

Indeed, for $u+v>0$ and $v>u$, we have $x>0$ and $y>0$, and then

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

See the above form of $f_{(X,Y)}(x,y)$.

Step 7. Marginal density of V? $f_V(v)$ should be

$$f_V(v) = \int_{-\infty}^{+\infty} f_{(U,V)}(u, v)du.$$

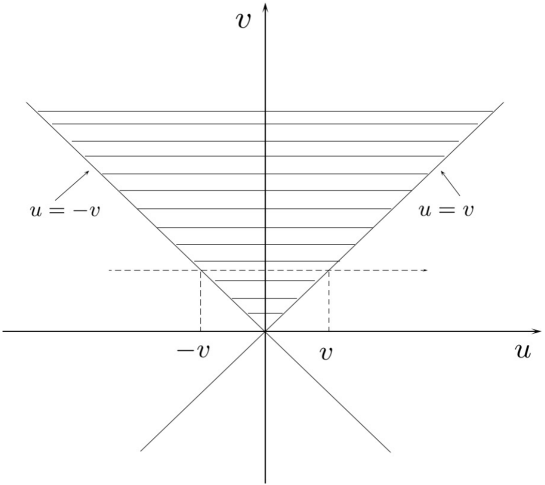
In our example, if $v < 0$, then $f_{(U,V)}(u, v) = 0$ for all u and thus $f_V(v) = 0$.

If $v \geq 0$, then

$$\int_{-\infty}^{+\infty} f_{(U,V)}(u, v)du = \int_{-\infty}^{-v} + \int_{-v}^{+v} + \int_{+v}^{+\infty}.$$

Part 4.4.3: Example

Check the chart



Part 4.4.3: Example

Hence

$$\int_{-\infty}^{+\infty} f_{(U,V)}(uv)du = \int_{-\infty}^{-v} 0du + \int_{-v}^{+v} \frac{\lambda^2}{2} e^{-\lambda v} du + \int_{+v}^{+\infty} 0du$$
$$= \frac{\lambda^2}{2} e^{-\lambda v} \int_{-v}^{+v} du$$
$$= \lambda^2 v e^{-\lambda v}.$$

Thus,

$$f_V(v) = \begin{cases} \lambda^2 v e^{-\lambda v}, & \text{if } v \geq 0, \\ 0, & \text{if } v < 0. \end{cases}$$

i.e., $V \sim \Gamma(\lambda, 2)$.

Part 4.4.4: Further Remarks

1. Conditions: Recall the random vector (X, Y) with known p.d.f. $f_{(X,Y)}(x, y)$, together with two known functions $g_1(x, y)$ and $g_2(x, y)$.

$$U = g_1(X, Y), \quad V = g_2(X, Y), \quad f_{(U,V)}(u, v) = ?$$

To make our conclusion valid, the following conditions need to be satisfied:

Part 4.4.4: Further Remarks

(1) The equations $u = g_1(x, y)$ and $v = g_2(x, y)$ can be uniquely solved for x and y in terms of u and v with solutions given by, say,

$$x = h_1(u, v) \quad \text{and} \quad y = h_2(u, v).$$

In short, $h_1(u, v)$ and $h_2(u, v)$ must be uniquely determined by the known functions $g_1(x, y)$ and $g_2(x, y)$.

[Recall the single random variable case we need $g(x) \downarrow\downarrow$ or $g(x) \uparrow\uparrow$]

(2) The partial derivatives $\frac{\partial x}{\partial u}$, $\frac{\partial x}{\partial v}$, $\frac{\partial y}{\partial u}$ and $\frac{\partial y}{\partial v}$ etc. exist and the Jacobian is not zero.

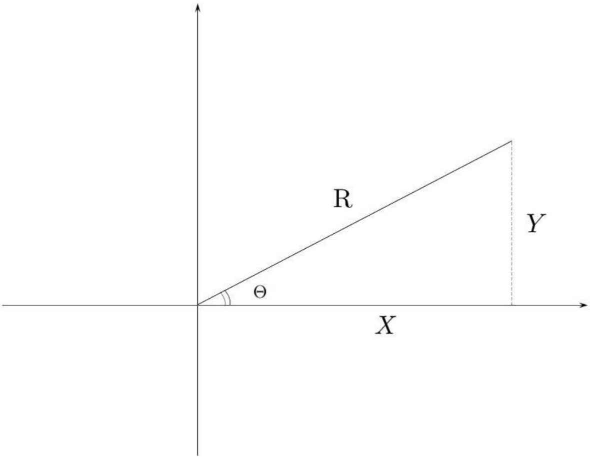
Part 4.4.4: Further Remarks

2. More examples:

Let (X, Y) denote a random point in the plane and assume that the rectangular coordinates X and Y are independent variables and that $X \sim N(0, 1)$ and $Y \sim N(0, 1)$.

Let (R, Θ) be the polar coordinate representation of this point (See below).

2. More examples:



Hence R and Θ are r.v.s as the functions of the r.v.s X and Y.

What is the joint p.d.f. of (R, Θ) ?

$$R = \sqrt{X^2 + Y^2}, \quad \Theta = \tan^{-1}\left(\frac{Y}{X}\right).$$

The function form is

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

Part 4.4.4: Further Remarks

(i) Inverting: $r = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(\frac{y}{x})$

$$\Rightarrow \quad \boxed{x = r\cos\theta \quad y = r\sin\theta}$$

(ii) Jacobian

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= r(\cos^2\theta + \sin^2\theta) = r.$$

(iii) $f_{(R,\Theta)}(r,\theta) = f_{(X,Y)}(r\cos\theta, r\sin\theta)\cdot|J| = r\cdot f_{(X,Y)}(r\cos\theta, r\sin\theta)$.

But X and Y are independent standard Normal r.v.s and thus

$$f_{(X,Y)}(x,y) = f_X(x)\cdot f_Y(y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}.$$

Hence

$$f_{(R,\Theta)}(r,\theta) = r\cdot\frac{1}{2\pi}e^{-\frac{r^2\cos^2\theta+r^2\sin^2\theta}{2}} = r\cdot\frac{1}{2\pi}e^{-\frac{r^2}{2}},$$

where $0 \le r < \infty$ and $0 \le \theta \le 2\pi$.

Part 4.4.4: Further Remarks

That is,

$$f_{(R,\Theta)}(r,\theta) = \frac{1}{2\pi} r e^{-\frac{r^2}{2}}, \quad 0 \le \theta \le 2\pi,\ 0 \le r < \infty.$$

Marginal p.d.f.s of R and Θ?

$$f_R(r) = \int_0^{2\pi} f_{(R,\theta)}(r,\theta) d\theta = \frac{1}{2\pi} \int_0^{2\pi} r e^{-\frac{r^2}{2}} d\theta = r e^{-\frac{r^2}{2}}, \quad 0 < r < \infty.$$

This is usually called <u>the Rayleigh distribution</u>.

$$f_\Theta(\theta) = \int_0^{+\infty} f_{(R,\theta)}(r,\theta)dr = \frac{1}{2\pi}\int_0^{+\infty} re^{-\frac{r^2}{2}}dr$$
$$= \frac{1}{2\pi}\left[-e^{-\frac{r^2}{2}}\right]\Bigg|_0^\infty = \frac{1}{2\pi}, \quad 0 < \theta < 2\pi.$$

That is that, Θ is uniformly distributed over $(0, 2\pi)$.

Also, easy to see

$$f_{(R,\Theta)}(r,\theta) = \frac{1}{2\pi}re^{-\frac{r^2}{2}} = f_\Theta(\theta)\cdot f_R(r).$$

So the two new random variables R and Θ are also independent.

Part 4.4.5: $\max\{X, Y\}$ and $\min\{X, Y\}$

Theorem. Let X and Y be independent r.v.s. with c.d.f.s $F_X(x)$ and $F_Y(y)$, respectively. Let

$$U = \max\{X, Y\}, \quad V = \min\{X, Y\}.$$

Then

$$F_{\max}(u) \triangleq F_U(u) = F_X(u) \cdot F_Y(u),$$

$$F_{\min}(v) \triangleq F_V(v) = 1 - [1 - F_X(v)] \cdot [1 - F_Y(v)].$$

Part 4.4.5: $\max\{X, Y\}$ and $\min\{X, Y\}$

Moreover, if $X_i \sim F_{X_i}(x)$, $i = 1, 2, \cdots, n$, and $X_1, X_2, \cdots, X_n$ are independent, then

$$F_{\max}(z) = P\{\max(X_1, X_2, \cdots, X_n) \leq z\}$$
$$= F_{X_1}(z) F_{X_2}(z) \cdots F_{X_n}(z).$$

$$F_{\min}(z) = P\{\min(X_1, X_2, \cdots, X_n) \leq z\}$$
$$= 1 - \prod_{i=1}^{n} \left[1 - F_{X_i}(z)\right].$$

In particular, if $X_1, X_2, \cdots, X_n$ are independent and have the same c.d.f. $F(x)$, then

$$F_{\max}(z) = F^n(z), \quad F_{\min}(z) = 1 - [1 - F(z)]^n.$$

Part 4.4.5: $\max\{X, Y\}$ and $\min\{X, Y\}$

Proof. Since X and Y are independent, we have

$$\begin{aligned} F_{\max}(z) &= P\{\max(X, Y) \leq z\} = P\{X \leq z, Y \leq z\} \\ &= P\{X \leq z\} \cdot P\{Y \leq z\} = F_X(z) \cdot F_Y(z). \end{aligned}$$

$$\begin{aligned} F_{\min}(z) &= P\{\min(X, Y) \leq z\} = 1 - P\{\min(X, Y) > z\} \\ &= 1 - P\{X > z, Y > z\} \\ &= 1 - P\{X > z\} \cdot P\{Y > z\} \\ &= 1 - [1 - P\{X \leq z\}] \cdot [1 - P\{Y \leq z\}] \\ &= 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]. \end{aligned}$$

Part 4.4.5: $\max\{X, Y\}$ and $\min\{X, Y\}$

Example. Suppose that X and Y are independent r.v.s and each is of exponential distribution with parameter $\frac{1}{3}$, i.e.,

$$f(x) = \begin{cases} 3e^{-3x}, & x > 0, \\ 0, & x \leq 0. \end{cases} \qquad f(y) = \begin{cases} 3e^{-3y}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Let $V = X + Y$, $W = \min(X, Y)$ and $Z = \max(X, Y)$.

(i) Find the distribution function of V.

(ii) Find the distribution function of W.

(iii) Find the distribution function of Z.

Solution. (i)

$$F_V(v) = \begin{cases} 1 - (3v+1)e^{-3v}, & v > 0, \\ 0, & v \le 0. \end{cases}$$

(ii)

$$F_W(w) = \begin{cases} 1 - e^{-6w}, & w > 0, \\ 0, & w \le 0. \end{cases}$$

(iii)

$$F_Z(z) = \begin{cases} 1 - 2e^{-3z} + e^{-6z}, & z > 0, \\ 0, & z \le 0. \end{cases}$$

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

The joint c.d.f. of $X_1, X_2, \cdots, X_n$, is defined by

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n).$$

For example, four random variables X_1, X_2, X_3, X_4, then

$$F(x_1, x_2, x_3, x_4) = P(X_1 \le x_1, X_2 \le x_2, X_3 \le x_3, X_4 \le x_4).$$

Part 4.5.2: One-dimensional marginal c.d.f.

$$F_{X_i}(x_i) = P(X_i \le x_i) \quad (i = 1, 2, \cdots, n).$$

For example,

$$F_{X_1}(x_1) = P(X_1 \le x_1).$$

Part 4.5.3: Joint p.d.f. of continuous r.v.s

$$f(x_1, x_2, \cdots, x_n) = \frac{\partial^n F(x_1, x_2, \cdots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n},$$

$$F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f(\mu_1, \cdots, \mu_n) d\mu_1 \cdots d\mu_n.$$

For example, four r.v.s X_1, X_2, X_3 and X_4,

$$f(x_1, x_2, x_3, x_4) = \frac{\partial^4 F(x_1, x_2 x_3, x_4)}{\partial x_1 \partial x_2 \partial x_3 \partial x_4},$$

$$F(x_1, x_2, x_3, x_4) = \int_{-\infty}^{x_4} \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(\mu_1, \mu_2, \mu_3, \mu_4) d\mu_1 d\mu_2 d\mu_3 d\mu_4.$$

The p.d.f. of the random variable X_i $(i = 1, 2, \ldots, n)$

$$f_{X_i}(x) = \frac{d}{dx} F_{X_i}(x) \quad (i = 1, 2, \ldots, n)$$

is called a marginal p.d.f. of X_i.

Part 4.5.5: One-dimensional marginal p.d.f.

1. Suppose $X = (X_1, X_2, \cdots, X_n)$ is a n-dimensional continuous vector with joint p.d.f. $f(x_1, x_2, \cdots, x_n)$.

Then the joint c.d.f. is

$$F(x_1, x_2, \cdots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n\}$$
$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(u_1, u_2, \cdots, u_n) du_1 du_2 \cdots du_n.$$

In more general, for any n-dimensional Borel set G, we have

Part 4.5.5: One-dimensional marginal p.d.f.

$$P\{X \in G\} = P\{(X_1, X_2, \cdots, X_n) \in G\} = \int \cdots \int_G f(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n.$$

In particular, let $G = \mathbb{R}^n$, then

$$\begin{aligned} &P\{(X_1, X_2, \cdots, X_n) \in \mathbb{R}^n\} \\ &= P\{-\infty < X_1 < +\infty, \cdots, -\infty < X_n < +\infty\} \\ &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \\ &= P(\Omega) = 1. \end{aligned}$$

Part 4.5.5: One-dimensional marginal p.d.f.

Hence we get the following well-known “conclusion”:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n)dx_1 dx_2 \cdots dx_n = 1.$$

Furthermore, let

$$G = \prod_{i=1}^{n}(-\infty, x_i],$$

then

$$P\{-\infty < X_1 \leq x_1, -\infty < X_2 \leq x_2, \cdots, -\infty < X_n \leq x_n\}$$
$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(u_1, u_2, \cdots, u_n)du_1 du_2 \cdots du_n.$$

Part 4.5.5: One-dimensional marginal p.d.f.

2. How to get marginal p.d.f. from joint p.d.f.

For example, the 1-D c.d.f. of X_1: $F(x_1) = P\{X_1 \leq x_1\}$.

$$
\begin{aligned}
&F_{X_1}(x_1) = P\{X_1 \leq x_1\} \\
&= P\{-\infty < X_1 \leq x_1, -\infty < X_2 < +\infty, \cdots, -\infty < X_n < +\infty\} \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, u_2, \cdots, u_n) du_1 du_2 \cdots du_n.
\end{aligned}
$$

Part 4.5.5: One-dimensional marginal p.d.f.

$\Rightarrow$ the p.d.f. of X_1 is given by

$$\begin{aligned}
\frac{d}{dx_1}F(x_1) &= \frac{d}{dx_1}\left(\int_{-\infty}^{x_1}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(u_1 u_2,\cdots,u_n)du_1 du_2\cdots du_n\right) \\
&= \frac{d}{dx_1}\int_{-\infty}^{x_1}\left(\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(u_1,u_2,\cdots,u_n)du_2\cdots du_n\right)du_1 \\
&= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,u_2,\cdots,u_n)du_2 du_3\cdots du_n.
\end{aligned}$$

Part 4.5.5: One-dimensional marginal p.d.f.

Similarly, for a 2-D marginal distribution, the joint c.d.f. of (X_1, X_2), say, is

$$
\begin{aligned}
& F_{(X_1,X_2)}(x_1, x_2) = P\{X_1 \le x_1, X_2 \le x_2\} \\
& = P\{-\infty < X_1 \le x_1, -\infty < X_2 \le x_2, -\infty < X_i < +\infty \ (i \ge 3)\} \\
& = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, u_2, \cdots, u_n) du_1 du_2 \cdots du_n.
\end{aligned}
$$

Part 4.5.5: One-dimensional marginal p.d.f.

$\Rightarrow$ the joint p.d.f. of (X_1, X_2) can be obtained as follows:

$$
\begin{aligned}
&\frac{\partial}{\partial x_1} F_{(X_1,X_2)}(x_1, x_2) \\
&= \frac{\partial}{\partial x_1} \int_{-\infty}^{x_1} \left[\int_{-\infty}^{x_2} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, u_2, \cdots u_n) du_2 \cdots du_n \right] du_1 \\
&= \int_{-\infty}^{x_2} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, u_2, \cdots, u_n) du_2 du_3 \cdots du_n.
\end{aligned}
$$

Part 4.5.5: One-dimensional marginal p.d.f.

$$
\begin{aligned}
&\frac{\partial^2}{\partial x_1 \partial x_2} F_{(X_1,X_2)}(x_1, x_2) \\
&= \frac{\partial}{\partial x_2} \int_{-\infty}^{x_2} \left[\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, u_2, \cdots, u_n) du_3 du_4 \cdots du_n \right] du_2 \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, u_3, \cdots, u_n) du_3 du_4 \cdots du_n.
\end{aligned}
$$

Part 4.5.5: One-dimensional marginal p.d.f.

Hence, the 2-D marginal p.d.f. of (X_1, X_2) is given by

$$f_{(X_1,X_2)}(x_1, x_2) \\ = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, u_3 \cdots, u_n) du_3 du_4 \cdots du_n.$$

Similarly, we could get any other marginal p.d.f.s.

Part 4.5.5: One-dimensional marginal p.d.f.

Random variables $X_1, X_2, \cdots, X_n$ are called <u>(mutually) independent</u> if for any $x_1, x_2, \cdots, x_n \in (-\infty, +\infty)$,

$$\begin{aligned} &P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n) \\ &= P(X_1 \le x_1) \cdot P(X_2 \le x_2) \cdot \cdots \cdot P(X_n \le x_n), \end{aligned}$$

i.e.,

$$F(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i).$$

Part 4.5.5: One-dimensional marginal p.d.f.

If all the r.v.s are continuous, then they are independent iff

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i),$$

where $f(x_1, x_2, \cdots, x_n)$ is the joint p.d.f. and $f_{X_1}(x_1), f_{X_2}(x_2), \cdots, f_{X_n}(x_n)$ are marginal p.d.f.s.

Part 4.5.5: One-dimensional marginal p.d.f.

Example: Let X, Y, Z be the independent r.v.s with a common uniform distribution over $(0, 1)$. Compute $P\{X \geq YZ\}$.

Solution: By independence property, the joint c.d.f. $F(x, y, z)$ and the joint p.d.f. $f(x, y, z)$ are given by

$$F_{(X,Y,Z)}(x, y, z) = \int_{-\infty}^{x} \int_{-\infty}^{y} \int_{-\infty}^{z} f(u, v, w) dudvdw$$

and

$$f_{(X,Y,Z)}(x, y, z) = f_X(x) \cdot f_Y(y) \cdot f_Z(z), \quad (\because X, Y, Z \text{ independent}).$$

But X, Y, Z are independent, identically distributed with a common distribution $U(0,1)$ and hence it is easily get that

$$f_{(X,Y,Z)}(x,y,z) = \begin{cases} 1, & \text{if } 0 \le x \le 1,\ 0 \le y \le 1,\ 0 \le z \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

$P\{X \ge YZ\} = P\{(X,Y,Z) \in G\}$, where

$$G = \left\{(x,y,z) \in \mathbb{R}^3; x \ge yz\right\}.$$

Let $D = \{(x,y,z); 0 \le x \le 1, 0 \le y \le 1, 0 \le z \le 1\}$, $D^c = \mathbb{R}^3 \backslash D$.

Then $G = G \cap \mathbb{R}^3 = (G \cap D) \cup (G \cap D^c)$.

But $G \cap D$ and $G \cap D^c$ are disjoint, and hence

$$P\{X \geq YZ\} = P\{(X, Y, Z) \in G\}$$
$$= P\{(X, Y, Z) \in G \cap D\} + P\left\{(X, Y, Z) \in G \cap D^c\right\}.$$

That is,

$$P\{X \geq YZ\} = \iiint\limits_{x \geq yz} f_{(X,Y,Z)}(x, y, z)dxdydz$$

$$= \iiint\limits_{\substack{x \geq yz \\ 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \\ 0 \leq z \leq 1}} \overbrace{f_{(X,Y,Z)}(x, y, z)}^{1} dxdydz + \iiint\limits_{\substack{x \geq yz \\ \{0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}^c}} \overbrace{f_{(X,Y,Z)}(x, y, z)}^{0} dxdydz$$

Part 4.5.5: One-dimensional marginal p.d.f.

$$P\{X \geq YZ\} = \iiint\limits_{\substack{x \geq yz \\ 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \\ 0 \leq z \leq 1}} 1 dxdydz + 0 = \iiint\limits_{\substack{x \geq yz \\ 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \\ 0 \leq z \leq 1}} dxdydz$$

$$= \int_0^1 \int_0^1 \int_{yz}^1 dxdydz = \int_0^1 \int_0^1 (1 - yz)dydz$$

$$= \int_0^1 \left[\int_0^1 (1 - yz)dy\right] dz = \int_0^1 \left[y - \frac{y^2}{2}z\right]\Bigg|_0^1 dz$$

$$= \int_0^1 \left[1 - \frac{z}{2}\right] dz = \left[z - \frac{z^2}{4}\right]\Bigg|_0^1 = 1 - \frac{1}{4} = \frac{3}{4}.$$

Outline

1. Part 4.1: Introduction
2. Part 4.2: Discrete Random Variables
3. Part 4.3: Continuous Random Variable
4. Part 4.4: Transformation of Random Vector
5. Part 4.5: General Case
6. Part 4.6: Summary

Part 4.6.1: Summary of Chapter 2

Basic Concepts

1. Sample Space: Ω
2. Events: (Impossible event $\emptyset$, Certain event Ω, Elementary event; General event.)
3. Probability Measure: (Set function: events $\rightarrow R$)
4. Independence:
5. Conditional Probability:
6. Disjoint Events:
7. Partition of Ω:

Part 4.6.1: Summary of Chapter 2

Operations of Events

1. Union: $A \cup B = \{$ Either A or B occurs $\}$
2. Intersection: $A \cap B = \{$ Both A and B occur $\}$
3. Complement: $A^c = \{A$ does not occur $\}$
4. $\cup_{k=1}^{n} A_k$ and $\cup_{k=1}^{\infty} A_k, \cap_{k=1}^{n} A_k$ and $\cap_{k=1}^{\infty} A_k$.

Part 4.6.1: Summary of Chapter 2

Properties of Probability

1. $0 \leq P(A) \leq 1, \forall A.$
2. $P(\emptyset) = 0, \quad P(\Omega) = 1.$
3. $A \subset B \Rightarrow P(A) \leq P(B)$
4. $\{B_k\}$ disjoint $\Rightarrow P(\cup_k B_k) = \sum_k P(B_k)$
5. $\{B_k\}$ independent $\Rightarrow P(\cap_{k=1}^n B_k) = \prod_{k=1}^n P(B_k)$

Part 4.6.1: Summary of Chapter 2

Important Formulas

1. $P(A^c) = 1 - P(A)$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. $P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$
4. $P(A \cup B) = P(A) + P(B)$ if A, B disjoint.
5. $P(A \cap B) = P(A) \cdot P(B)$ if A, B independent
6. If $\{B_k\}$ is a partition of Ω, then for any A,

$P(A) = \sum_k P(B_k) \cdot P(A \mid B_k)$ and

$P(B_n \mid A) = \frac{P(B_n) \cdot P(A|B_n)}{\sum_k P(B_k) \cdot P(A|B_k)}$.

Part 4.6.2: Summary of Chapter 3

Basic Concepts

1. Random Variables: (Definition; meaning; two types)
2. Cumulative Distribution Function: c.d.f.

$$F(x) = P(X \leq x)(\text{ for both types }).$$

3. Probability mass function: p.m.f. (discrete r.v.)
   Probability density function: p.d.f. (continuous r.v.)
4. Poisson Distribution; Normal Distribution; Other distributions

Calculation

1. Basic formula: For $a < b$,

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

where the random variable X has c.d.f. $F(x)$.

2. For continuous r.v. with p.d.f. $f(x)$, we further have

$$\begin{aligned}\mathbb{P}(a < X \leq b) &= \mathbb{P}(a \leq X < b) = \mathbb{P}(a \leq X \leq b) \\ &= \mathbb{P}(a < X < b) = \int_a^b f(x)dx.\end{aligned}$$

3. $X \sim \mathcal{N}(0,1) : \mathbb{P}(a < X < b) = \Phi(b) - \Phi(a)$, then check the table (if $a < 0$, then $\Phi(a) = 1 - \Phi(-a)$).
4. $X \sim \mathcal{N}(\mu, \sigma^2) :$ Let $Y = \frac{X-\mu}{\sigma}$, then $Y \sim \mathcal{N}(0,1)$. (c.d.f.)

Part 4.6.3: Summary of Chapter 4

I. Basic Concepts:

1. Joint c.d.f. and Marginal c.d.f.
2. Joint p.m.f. and Marginal p.m.f. (discrete case)
3. Joint p.d.f. and Marginal p.d.f. (continuous case)

*4. Independence

II. Basic Conclusions:

*1. Two random variables X and Y are independent if and only if

$$F(x, y) = F_X(x) \cdot F_Y(y),$$

where $F(x, y)$: joint c. d.f.

$F_X(x)$ and $F_Y(y)$: marginal c.d.f.

(True for both discrete and continuous case)

For continuous random variables X and Y, they are independent if and only if

$$f(x, y) = f_X(x) \cdot f_Y(y),$$

where $f(x, y)$: joint p.d.f.
$f_X(x)$ and $f_Y(y)$: marginal p.d.f.

For discrete random variables X and Y, they are independent if and only if

$$p(x, y) = p_X(x) \cdot p_Y(y),$$

where $p(x, y)$: joint p.m.f.
$p_X(x)$ and $p_Y(y)$: marginal p.m.f.

2. For general random variable, X and Y,

joint $\mathrm{c.d.f.}$ $\overset{\text{determines}}{\Longrightarrow}$ marginal $\mathrm{c.d.f.}$'s

joint p.m.f. $\overset{\text{determines}}{\Longrightarrow}$ marginal p.m.f.'s (discrete case)

joint p.d.f. $\overset{\text{determines}}{\Longrightarrow}$ marginal p.d.f.'s (continuous case)

vice versa for independent (discrete or continuous) case

Part 4.6.3: Summary of Chapter 4

3. For n random variables $(^*)n$ random variables:

$X_1, X_2, \cdots, X_n$ They are (mutually) independent if and only if "the joint $\mathrm{c.d.f.}$ is the product of n marginal $\mathrm{c.d.f.}$'s".

If all are continuous random variables, then they are "independent" if and only if "the joint p.d.f. is the product of n marginal p.d.f.'s".

If all are discrete random variables, then they are "independent" if and only if "the joint p.m.f. is the product of n marginal p.m.f.'s".