Probability and Statistics for Engineers

Dr. Sanae KOUISMI

AUI, School of Science and Engineering
Ifrane, Summer 25

Lectures 8 and 9 :
3.4 Joint Probability Distributions

Outline

1 Joint Probability Distributions

1 Joint Probability Distributions

- Joint Probability Distribution and Joint Density Function
- Marginal Distribution
- Conditional Distribution
- Statistical Independence
- Exercise 6

Joint Probability Distributions

Definition (Couple of two random variables)

Let X and Y be two random variables defined on the same sample space S. A ***couple of two random variables*** *(X, Y) is the mapping :*

$$\begin{array}{rcl} (X,Y): \; S & \to & \mathbb{R}^2 \\ \omega & \mapsto & (X(\omega), Y(\omega)) \subset \{(x_i, y_j),\; i \in I,\; j \in J\} \end{array}$$

where $I, J \subset \mathbb{N}$ are two sets on which the values taken by X and Y respectively are indexed.

Joint Probability Distributions

Definition (Joint probability distribution)

The function $f(x,y)$ is a ***joint probability distribution*** *or* ***probability mass function*** *of* ***the discrete random variables*** *X and Y if*

1. $f(x,y) \geq 0$ *for all* (x,y),
2. $\sum_x \sum_y f(x,y) = 1$,
3. $P(X = x, Y = y) = f(x,y)$.

For any region A in the xy plane,
$P[(X,Y) \in A] = \sum_{(x,y) \in A} f(x,y)$.

Joint Probability Distributions

Example 1

Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

(a) the joint probability function $f(x, y)$,

(b) $P[(X, Y) \in A]$, where A is the region $\{(x, y) \,|\, x + y \leq 1\}$.

Joint Probability Distributions

(a)

The possible pairs of values (x, y) are (0, 0), (0, 1), (1, 0), (1, 1), (0, 2), and (2, 0)

$f(0, 1)$, for example, represents the probability that a red and a green pen are selected. The total number of equally likely ways of selecting any 2 pens from the 8 is

$$\binom{8}{2} = 28.$$

The number of ways of selecting 1 red pen from 2 red pens and 1 green pen from 3 green pens is

$$\binom{2}{1} \cdot \binom{3}{1} = 6.$$

Joint Probability Distributions

Hence, $f(0,1) = \frac{6}{28} = \frac{3}{14}$.
Similar calculations yield the probabilities for the other cases, which are presented in the following table

		y			**Totals**
	$f(x,y)$	**0**	**1**	**2**	
x	**0**				
	1				
	2				
	Totals				

Joint Probability Distributions

		x			**Totals**
	$f(x,y)$	**0**	**1**	**2**	
y	**0**	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	Totals	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Joint Probability Distributions

(b) The probability that (X, Y) fall in the region A is $P[(X, Y) \in A] = P(X + Y \leq 1) =$
$f(0,0) + f(0,1) + f(1,0) = \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}$.

Joint Probability Distributions

Definition (Joint density function)

The function $f(x, y)$ is a ***joint density function*** *of* ***the continuous random variables*** *X and Y if*

1. $f(x, y) \geq 0$ *for all* (x, y),
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$,
3. $P[(X, Y) \in A] = \int \int_{A} f(x, y)\, dx\, dy$ *for any region* A *in the xy plane.*

Joint Probability Distributions

Example 2

A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify condition 2 of the previous definition.

Joint Probability Distributions

(b) Find $P[(X, Y) \in A]$, as
$A = \{(x, y) \,|\, 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$

Joint Probability Distributions

Solution :

(a) The integration of $f(x,y)$ over the whole region is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^1\int_0^1 \tfrac{2}{5}(2x+3y)\,dx\,dy$$

$$= \int_0^1 \left[\frac{2x^2}{5} + \frac{6xy}{5}\right]_0^1 dy = \int_0^1 \left(\frac{2}{5} + \frac{6y}{5}\right) dy$$

$$= \left[\frac{2y}{5} + \frac{3y^2}{5}\right]_0^1 = \frac{2}{5} + \frac{3}{5} = 1$$

Joint Probability Distributions

(b) To calculate the probability, we use

$$P[(X,Y) \in A] = P\left\{0 < X < \tfrac{1}{2}, \tfrac{1}{4} < Y < \tfrac{1}{2}\right\}$$

$$= \int_{1/4}^{1/2} \int_{0}^{1/2} \tfrac{2}{5}(2x + 3y)\, dx\, dy = \int_{1/4}^{1/2} \left[\frac{2x^2}{5} + \frac{6xy}{5}\right]_0^{1/2} dy$$

$$= \int_{1/4}^{1/2} \left(\frac{1}{10} + \frac{3y}{5}\right) dy = \left[\frac{y}{10} + \frac{3y^2}{10}\right]_{1/4}^{1/2}$$

$$= \frac{1}{10}\left(\frac{1}{2} + \frac{3}{4}\right) - \frac{1}{10}\left(\frac{1}{4} + \frac{3}{16}\right) = \frac{13}{160}$$

Marginal Distribution

Definition (Marginal Distribution)

The marginal distributions of X alone and of Y alone are

$$g(x) = \sum_{y} f(x, y) \quad \text{and} \quad h(y) = \sum_{x} f(x, y)$$

for the discrete case, and

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

for the continuous case.

Marginal Distribution

Example 3

Considering the example 1, give the marginal distribution of X alone and of Y alone.

Marginal Distribution

For the random variable X, we see that

$$g(0) = f(0,0) + f(0,1) + f(0,2) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$g(1) = f(1,0) + f(1,1) + f(1,2) = \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28},$$

and

$$g(2) = f(2,0) + f(2,1) + f(2,2) = \frac{3}{28} + 0 + 0 = \frac{3}{28},$$

Marginal Distribution

which are just the column totals of the previous table . In a similar manner, we could show that the values of $h(y)$ are given by the row totals. In tabular form, these marginal distributions may be written as follows :

X	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

Y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Marginal Distribution

Example 4

Find g(x) and h(y) for the joint density function from the example 2

Marginal Distribution

Solution :
By definition, $g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{2}{5}(2x+3y)\,dy$
$= \left[\frac{4xy}{5} + \frac{6y^2}{10}\right]_0^1 = \frac{4x+3}{5}$ for $0 \leq x \leq 1$
and $g(x) = 0$ elsewhere.
Similarly,
$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{2}{5}(2x+3y)\,dx$
$= \frac{2(1+3y)}{5}$ for $0 \leq y \leq 1$,
and $h(y) = 0$ elsewhere.

Conditional Distribution

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad \text{provided } P(A) > 0.$$

Now, considering $P(Y = y|X = x)$, where X and Y are discrete random variables, we have

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x, y)}{g(x)},$$

provided $g(x) > 0$.

Conditional Distribution

Definition (Conditional Probability Distribution)

Let X and Y be two random variables, discrete or continuous. The ***conditional distribution*** *of the random variable Y given that $X = x$ is*

$$f(y|x) = \frac{f(x,y)}{g(x)}, \quad \textit{provided } g(x) > 0.$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad \textit{provided } h(y) > 0.$$

Conditional Distribution

Example 5

Referring to Example 1, find the conditional distribution of X, given that $Y = 1$, and use it to determine $P(X = 0|Y = 1)$.

		x			**Totals**
	$f(x, y)$	**0**	**1**	**2**	
y	**0**	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	Totals	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Conditional Distribution

$$P(X = 0|Y = 1) = f(0|1) = \frac{f(0,1)}{h(1)}.$$

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}$$

$$f(0|1) = \frac{\frac{3}{14}}{\frac{3}{7}} = \frac{1}{2}$$

Conditional Distribution

Remark

- *If we wish to find the probability that the discrete random variable* X *falls between* a *and* b *when it is known that the discrete variable* $Y = y$*, we evaluate*

$$P(a < X < b|Y = y) = \sum_{a<x<b} f(x|y),$$

where the summation extends over all values of X *between* a *and* b*.*

Conditional Distribution

Remark

- *When X and Y are continuous, we evaluate*

$$P(a < X < b|Y = y) = \int_a^b f(x|y)\,dx.$$

Conditional Distribution

Example 6

Given the joint density function :

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find $g(x)$, $h(y)$, $f(x|y)$, and evaluate $P\left(\frac{1}{4} < X < \frac{1}{2} \,\middle|\, Y = \frac{1}{3}\right)$.

Conditional Distribution

By definition of the marginal density, for $0 < x < 2$:

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{x(1+3y^2)}{4}\,dy$$
$$= \int_0^1 \left(\frac{xy}{4} + \frac{xy^3}{4}\right) dy$$
$$= \left[\frac{xy^2}{2} + \frac{xy^4}{16}\right]_0^1 = \frac{x}{2}.$$

Conditional Distribution

And for $0 < y < 1$:

$$
\begin{aligned}
h(y) &= \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^2 \frac{x(1+3y^2)}{4}\, dx \\
&= \int_0^2 \left(\frac{x^2}{8} + \frac{3x^2y^2}{8} \right) dx \\
&= \left[\frac{x^3}{24} + \frac{3x^2y^2}{8} \right]_0^2 = \frac{1+3y^2}{2}.
\end{aligned}
$$

Conditional Distribution

Therefore, using the conditional density definition, for $0 < x < 2$:

$$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{\frac{x(1+3y^2)}{4}}{\frac{1+3y^2}{2}} = \frac{x}{2}$$

And

$$P\left(\frac{1}{4} < X < \frac{1}{2} \,|\, Y = \frac{1}{3}\right) = \int_{\frac{1}{4}}^{\frac{1}{2}} f(x|\frac{1}{3})\,dx = \frac{3}{64}$$

Statistical Independence

Definition

Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables X and Y are said to be ***statistically independent*** *if and only if*

$$f(x, y) = g(x)h(y)$$

Statistical Independence

Example 6

Show that the random variables of Example 1 are not statistically independent

		x			**Totals**
	$f(x,y)$	**0**	**1**	**2**	
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	**1**	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	Totals	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Statistical Independence

Let us consider the point $(0, 1)$. From the previous table, we find the three probabilities $f(0, 1)$, $g(0)$, and $h(1)$ to be

$$f(0,1) = \frac{3}{14},$$

$$g(0) = \sum_{y=0}^{2} f(0,y) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Clearly,

$$f(0,1) \neq g(0)h(1),$$

Statistical Independence

and therefore X and Y are not statistically independent.

Exercise 6

Consider the following joint probability density function of the random variables X and Y :

$$f(x,y) = \begin{cases} \frac{3x-y}{9}, & 1 < x < 3, 1 < y < 2 \\ 0, & \text{elsewhere} \end{cases}$$

(a) Find the marginal density functions of X and Y.
(b) Are X and Y independent ?