Two-dimensional Random Variable

Our study of random variables and their probability distributions in the preceding sections is restricted to one-dimensional sample spaces.
In that we recorded outcomes of an experiment as values assumed by a single random variable.
There will be situations, however, where we may find it desirable to record the simultaneous outcomes of several random variables

Example:

For instance, blood pressure and cholesterol for each individual are measured simultaneously.

In a study to determine the likelihood of success in college based on high school data, we might use a three dimensional sample space and record for each individual his or her aptitude test score, high school class rank, and grade-point average at the end of freshman year in college.

Definition: Let S be a sample space associated with a random experiment E. Let X and Y be two random variables defined on S. then the pair (X,Y) is called a Two–dimensional random variable.
The value of (X,Y) at a point $s \in S$ is given by the ordered pair of real numbers (X(s),Y(s))=(x,y) where X(s)=x, Y(s)=y.

Two –Dimensional discrete random variable: If the possible values of (X,Y) are finite or countably infinite, then (X,Y) is called a two-dimensional discrete random variable.
When (X,Y) is a two-dimensional discrete random variable the possible values of (X,Y) may be represented as (x_i, y_j), $i = 1, 2, 3, \dots n,\ \ j = 1, 2, 3, \dots m$.

Example: Consider the experiment of tossing a coin twice.

The sample space is S = {HH, HT, TH, TT}.Let X denotes the number of heads obtained in the first toss and Y denote the number of heads in the second toss.

s	HH	HT	TH	TT
X(s)	1	1	0	0
Y(s)	1	0	1	0

(X, Y) is a two-dimensional random variable or bi-variate random variable. The range space of (X, Y) is {(1,1), (1,0), (0,1), (0,0)} which is finite and so (X, Y) is a two-dimensional discrete random variables.

Joint probability distribution of Discrete R.V

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. $P(X = x, Y = y) = f(x, y)$.

$X \backslash Y$	y_1	y_2	y_3	$\cdots$	y_j	$\cdots$	y_m	Total
x_1	p_{11}	p_{12}	p_{13}	$\cdots$	p_{1j}	$\cdots$	p_{1m}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	$\cdots$	p_{2j}	$\cdots$	p_{2m}	$p_{2\cdot}$
x_3	p_{31}	p_{32}	p_{33}	$\cdots$	p_{3j}	$\cdots$	p_{3m}	$p_{3\cdot}$
.	.	.	.	$\cdots$	.	$\cdots$	.	
.	.	.	.	$\cdots$	.	$\cdots$	.	$\vdots$
.	.	.	.	$\cdots$	.	$\cdots$	.	
x_i	p_{i1}	p_{i2}	p_{i3}	$\cdots$	p_{ij}	$\cdots$	p_{im}	$p_{i\cdot}$
.	.	.	.	$\cdots$	.	$\cdots$	.	
.	.	.	.	$\cdots$	.	$\cdots$	.	$\vdots$
.	.	.	.	$\cdots$	.	$\cdots$	.	
x_n	p_{n1}	p_{n2}	p_{n3}	$\cdots$	p_{nj}	$\cdots$	p_{nm}	$p_{n\cdot}$
Total	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$\cdots$	$p_{\cdot j}$	$\cdots$	$p_{\cdot m}$	1

Roll two dice. Let X be the value on the first die and let T be the total on both dice. Here is the joint probability table:

$X \backslash T$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36

Exercise:

Suppose a car showroom has 10 cars of particular brand out of which 5 are good, 2 have defective transmission 3 have defective steering. If 2 cars are selected at random, find the joint probability distribution table.

Answer:

Let X denotes the number of cars with DT

Y denotes the number of cars with DS

$$P(X=0,Y=0)=\frac{\binom{5}{2}}{\binom{10}{2}}=\frac{2}{9}$$

$$P(X=0,Y=1)=\frac{\binom{5}{1}\binom{3}{1}}{\binom{10}{2}}=\frac{1}{3}$$

and so on….

X \ Y	0	1	2
0	10/45	15/45	3/45
1	10/45	6/45	0
2	1/45	0	0

Joint probability distribution of Continuous R.V

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.

Exercise:

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

(a) Verify condition 2

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

Answer:

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

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

$$= \int_0^1 \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Bigg|_{x=0}^{x=1} dy$$

$$= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5}\right) dy = \left(\frac{2y}{5} + \frac{3y^2}{5}\right)\Bigg|_0^1 = \frac{2}{5} + \frac{3}{5} = 1.$$

(b) To calculate the probability, we use

$$
\begin{aligned}
P[(X,Y) \in A] &= P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) \\
&= \int_{1/4}^{1/2} \int_{0}^{1/2} \frac{2}{5}(2x + 3y)\, dx\, dy \\
&= \int_{1/4}^{1/2} \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\bigg|_{x=0}^{x=1/2} dy = \int_{1/4}^{1/2} \left(\frac{1}{10} + \frac{3y}{5}\right) dy \\
&= \left(\frac{y}{10} + \frac{3y^2}{10}\right)\bigg|_{1/4}^{1/2} \\
&= \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}.
\end{aligned}
$$

Marginal Distributions

Given the joint probability distribution $f(x, y)$ of the discrete random variables X and Y, the probability distribution $g(x)$ of X alone is obtained by summing $f(x, y)$ over the values of Y. Similarly, the probability distribution $h(y)$ of Y alone is obtained by summing $f(x, y)$ over the values of X.

Definition:

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.

Exercise: Find $g(x)$ and $h(y)$ for the following distribution table:

$f(x,y)$		$x=0$	$x=1$	$x=2$	Row Totals
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	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Answer:

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

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

which are just the column totals of 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
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

$\lrcorner$

Expected Value of Two-dimensional RV

Let X and Y be random variables with joint probability distribution $f(x, y)$. The mean, or expected value, of the random variable $g(X, Y)$ is

$$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) \, dx \, dy$$

if X and Y are continuous.

Mean and Variance of Two dimensional random variable:

$$E(X) = \begin{cases} \sum_x \sum_y x f(x,y) = \sum_x x g(x) & \text{(discrete case)}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y)\ dy\ dx = \int_{-\infty}^{\infty} x g(x)\ dx & \text{(continuous case)}, \end{cases}$$

where g(x) is the marginal distribution of X.

Therefore, in calculating E(X) over a two-dimensional space, one may use either the joint probability distribution of X and Y or the marginal distribution of X.

Similarly, we define

$$E(Y) = \begin{cases} \sum_y \sum_x y f(x,y) = \sum_y y h(y) & \text{(discrete case)}, \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y)\ dx dy = \int_{-\infty}^{\infty} y h(y)\ dy & \text{(continuous case)}, \end{cases}$$

where h(y) is the marginal distribution of the random variable Y.

Properties of Expectation:

1. **Addition theorem of Expectation**: $\boldsymbol{E(X+Y) = E(X)+E(Y)}$

Proof:

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dxdy$$

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

$$E(X+Y) = E(X)+E(Y)$$

Generalisation: $E(X_1+X_2+\cdots+X_n) = E(X_1)+E(X_2)+\cdots+E(X_n)$

or $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$

2. Multiplication theorem of Expectation:

If X and Y are independent random variables, then $\boldsymbol{E(X\,Y) = E(X) \cdot E(Y)}$

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy)f(x)\,f(y)dxdy \qquad \text{[Since, X and Y are independent]}$$

$$= \left[\int_{-\infty}^{\infty} x\,f(x)\ dx\right]\ \left[\int_{-\infty}^{\infty} y\,f(y)\ dy\right]$$

$$E(X\,Y) = E(X)\,E(Y)$$

Generalisation: If $X_1, X_2, \dots, X_n$ are n- independent r.v.'s, then

$$E\,(X_1 \cdot X_2 \cdot \cdots \cdot X_n) = E\,(X_1) \cdot E(X_2) \cdot \cdots \cdot E(X_n)$$

or $$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$$

Variance of X and Y:

$$Var(X) = E(X^2) - \left(E(X)\right)^2$$

$$Var(Y) = E(Y^2) - \left(E(Y)\right)^2$$

Example: Two tire-quality experts examine stacks of tires and assign a quality rating to each tire on a 3-point scale. Let X denote the rating given by expert A and Y denote the rating given by B. The following table gives the joint distribution for X and Y . Find mean of X and mean of Y. Find variance of X and variance of Y.

		y		
$f(x,y)$		1	2	3
x	1	0.10	0.05	0.02
	2	0.10	0.35	0.05
	3	0.03	0.10	0.20

Solution:

$$\mu_X = \sum xg(x) = (1)(0.17) + (2)(0.5) + (3)(0.33) = 2.16,$$
$$\mu_Y = \sum yh(y) = (1)(0.23) + (2)(0.5) + (3)(0.27) = 2.04.$$

$$E(X^2) = \sum x_i^2 g(x) = 1^2(0.17) + 2^2(0.5) + 3^2(0.33) = 5.14$$

$$E(Y^2) = \sum y_j^2 h(y) = 1^2(0.23) + 2^2(0.5) + 3^2(0.27) = 4.66$$

$$Var(X) = E(X^2) - \left(E(X)\right)^2 = 5.14 - (2.16)^2 = .47$$

$$Var(Y) = E(Y^2) - \left(E(Y)\right)^2 = 4.66 - (2.04)^2 = .50$$

Exercise: Find mean and variance of X and Y.

The Joint pmf of X and Y is

Y \ X	-1	1
0	1/8	3/8
1	2/8	2/8

Example:

Two R.V's X and Y have joint pdf $f(x,y)=\begin{cases}\dfrac{xy}{96} & ,0<x<4,1<y<5\\ 0 & ,elsewhere\end{cases}$

Find E(X), E(Y), Var(X).

Solution:

i) $E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy$

$$=\int_{1}^{5}\int_{0}^{4}x\left(\frac{xy}{96}\right)dxdy$$

$$=\frac{8}{3}$$

ii) $E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy$

$$=\int_{1}^{5}\int_{0}^{4}y\left(\frac{xy}{96}\right)dxdy$$

$$=\frac{31}{9}$$

We know that, $Var(X) = E(X^2) - [E(X)]^2$

Now, $E(X^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)\,dx\,dy$

$$= \int_1^5\int_0^4 x^2\left(\frac{xy}{96}\right)dx\,dy$$

$$= 8$$

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2$$

$$= 8 - \left(\frac{8}{3}\right)^2 = \frac{8}{9}$$