Joint Distributions

Jointly Distributed Random Variables

For a given experiment, we are often interested not only in probability distribution functions of individual random variables but also in the relationships between two or more random variables.

Examples

- In an experiment into the possible causes of cancer, we might be interested in the relationship between the average number of cigarettes smoked daily and the age at which an individual contracts cancer.

- An engineer might be interested in the relationship between the shear strength and the diameter of a spot weld in a fabricated sheet steel specimen.

Two-dimensional Random Variables

- If the possible values of $[X_1 , X_2]$ are either finite or countably infinite in number, then $[X_1 , X_2]$ will be a <u>two-dimensional discrete random vector</u>.
  The possible values of $[X_1 , X_2]$ are $[x_{1i} , x_{2j}]$, i=1,2,3,... ; j =1,2,3,....

- If the possible values of $[X_1 , X_2]$ are some uncountable set in the Euclidean plane, then $[X_1 , X_2]$ will be a <u>two-dimensional continuous random vector</u>.
  For example, if $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$, then we would have
  $R_{X1 \times X2} = \{ [X_1 , X_2] \ : \ a \leq x_1 \leq b \ , \ c \leq x_2 \leq d \}$

Bivariate Probability Distributions

Handy Notation

If $X = [X_1, X_2]$,

$$p_X(x_1, x_2) = p(x_1, x_2) \quad \text{and} \quad f_X(x_1, x_2) = f(x_1, x_2)$$

where x_1, x_2 are the values which random variables X_1 and X_2 can take respectively.

Discrete Bivariate Probability Distributions

To each outcome $[x_1, x_2]$ of $[X_1, X_2]$, we associate a number,

$$p(x_1, x_2) = P(X_1 = x_1 \text{ and } X_2 = x_2)$$

where

$$p(x_1, x_2) \geq 0 \text{ , for all } x_1, x_2,$$

and

$$\sum_{x1} \sum_{x2} p(x_1, x_2) = 1$$

The values $([x_1, x_2], p(x_1, x_2))$ for all i, j make up the probability distribution for $[X_1, X_2]$.

Discrete Bivariate Probability Distributions

Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries. If we let X_1 and X_2 denote, respectively, the number of new and used but still working batteries that are chosen, then find the joint probability mass function of X_1 and X_2 , $p(i, j) = P\{X_1 = i, X_2 = j\}$

Discrete Bivariate Probability Distributions

$$P\{X_1 = i, X_2 = j\}$$

i \ j	0	1	2	3
0	p(0,0) = 10/220	p(0,1) = 40/220	p(0,2) = 30/220	p(0,3) = 4/220
1	p(1,0) = 30/220	p(1,1) = 60/220	p(1,2) = 18/220	p(1,3) = 0
2	p(2,0) = 15/220	p(2,1) = 12/220	p(2,2) = 0	p(2,3) = 0
3	p(3,0) = 1/220	p(3,1) = 0	p(3,2) = 0	p(3,3) = 0

Continuous Bivariate Probability Distributions

If $[X_1, X_2]$ is a continuous random vector in the Euclidean plane, then f, the joint density function, has the following properties :

$$f(x_1, x_2) \geq 0, \text{ for all } (x_1, x_2) \in R$$

and

$$\iint f(x_1, x_2)\, dx_1\, dx_2 = 1$$

A probability statement is then of the form

$$P\,(a \leq X_1 \leq b\,,\, c \leq X_2 \leq d) = \int_c^d \int_a^b f(x_1, x_2)\, dx_1\, dx_2$$

Continuous Bivariate Probability Distributions

The joint density function of X and Y is given by

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

Compute the following :

- $P\{X > 1, Y < 1\} = e^{-1}(1 - e^{-2})$
- $P\{X < a\} = 1 - e^{-a}$
- $P\{X < Y\} = 1/3$

Marginal Distributions

Having defined the bivariate probability distributions, sometimes called the joint probability distribution (or in continuous case the joint density), a natural question arises :

What is the distribution of X_1 , X_2 alone ?

Marginal Distributions - Discrete Case

- Marginal Distribution of X_1 is

$$p_1(x_1) = \sum_{x2} p(x_1, x_2) \text{ for all } x_1$$

- Marginal Distribution of X_2 is

$$p_2(x_2) = \sum_{x1} p(x_1, x_2) \text{ for all } x_2$$

Marginal Distributions - Discrete Case

j / i	0	1	2	3	$P\{X_1 = i\}$
0	p(0,0) = 10/220	p(0,1) = 40/220	p(0,2) = 30/220	p(0,3) = 4/220	$P(X_1 = 0) = 84/220$
1	p(1,0) = 30/220	p(1,1) = 60/220	p(1,2) = 18/220	p(1,3) = 0	$PP(X_1 = 1) = 108/220$
2	p(2,0) = 15/220	p(2,1) = 12/220	p(2,2) = 0	p(2,3) = 0	$P(X_1 = 2) = 27/220$
3	p(3,0) = 1/220	p(3,1) = 0	p(3,2) = 0	p(3,3) = 0	$P(X_1 = 3) = 1/220$
$P\{X_2 = j\}$	$P(X_2 = 0) = 56/220$	$P(X_2 = 1) = 112/220$	$P(X_2 = 2) = 48/220$	$P(X_2 = 3) = 4/220$	Sum = 1

Marginal Distributions - Continuous Case

- Marginal Distribution of X_1 is

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$$

- Marginal Distribution of X_2 is

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$$

Marginal Distributions - Continuous Case

The joint density function of X and Y is given by

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

Compute the following :

- $f_1(x) = 6x(1 - x)$
- $f_2(y) = 3y^2$

Expectation and variance of a marginal distribution

Let us write down the expected value and variance for X_1

- $E(x_1) = \mu_1 = \sum_{x1} x_1 p_1(x_1) = \sum_{x1} \sum_{x2} x_1 p(x_1, x_2)$

- $V(x_1) = (\sigma_1)^2 = \sum_{x1} (x_1 - \mu_1)^2 p_1(x_1)$

  $= \sum_{x1} (x_1)^2 p_1(x_1) - (\mu_1)^2$

  $= \sum_{x1} \sum_{x2} (x_1)^2 p(x_1, x_2) - (\mu_1)^2$

Independent Random Variables

The random variables X and Y are said to be independent if for any two sets of real numbers A and B

$$P\{X \in A, Y \in B\} = P\{X \in A\}\, P\{Y \in B\}$$

In other words, X and Y are independent if, for all A and B, the events $E_A = \{X \in A\}$ and $F_B = \{Y \in B\}$ are independent.

Loosely speaking, X and Y are independent if knowing the value of one does not change the distribution of the other. Random variables that are not independent are said to be dependent.

Independent Random Variables - Discrete Case

When X and Y are discrete random variables, the condition of independence is given by,

$$p(x, y) = p_X(x)\, p_Y(y) \quad \text{for all } x, y$$

Where p_X and p_Y are probability mass functions of X and Y

Independent Random Variables - Discrete Case

Suppose that the successive daily changes of the price of a given stock are assumed to be independent and identically distributed random variables with probability mass function given by

$$P\{\text{daily change is } i\} = \begin{cases} -3 \text{ with probability } .05 \\ -2 \text{ with probability } .10 \\ -1 \text{ with probability } .20 \\ 0 \text{ with probability } .30 \\ 1 \text{ with probability } .20 \\ 2 \text{ with probability } .10 \\ 3 \text{ with probability } .05 \end{cases}$$

What is the probability that the stock's price will increase successively by 1, 2, and 0 points in the next three days ?

Independent Random Variables - Continuous Case

When X and Y are continuous random variables, the condition of independence is given by,

$$f(x, y) = f_X(x)\, f_Y(y) \quad \text{for all } x, y$$

Where f_X and f_Y are probability density functions of X and Y

Independent Random Variables - Continuous Case

Suppose that X and Y are independent random variables having the common density function

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Verify that these two random variables are independent