If the possible values of (X, Y) are finite or countably infinite, (X, Y) is called *a two-dimensional discrete* RV. When (X, Y) is a two-dimensional discrete RV the possible values of (X, Y) may be represented as (x_i, y_j), $i = 1, 2, \ldots, m, \ldots; j = 1, 2, \ldots, n, \ldots$.

If (X, Y) can assume all values in a specified region R in the xy-plane, (X, Y) is called a *two-dimensional continuous* RV.

Probability Function of (X, Y)

If (X, Y) is a two-dimensional discrete RV such that $P(x = x_i, y = y_j) = p_{ij}$, then p_{ij} is called the *probability mass function* or simply the *probability function* of (X, Y) provided the following conditions are satisfied.

(i) $p_{ij} \geq 0$, for all i and j

(ii) $\sum\limits_j \sum\limits_i p_{ij} = 1$

The set of triples $\{x_i, y_i, p_{ij}\}$, $i = 1, 2, \ldots, m, \ldots, j = 1, 2, \ldots, n, \ldots$, is called *the joint probability distribution of (X, Y).*

Joint Probability Density Function

If (X, Y) is a two-dimensional continuous RV such that.

$P\left\{x - \dfrac{dx}{2} \leq X \leq x + \dfrac{dx}{2} \text{ and } y - \dfrac{dy}{2} \leq Y \leq y + \dfrac{dy}{2}\right\} = f(x, y)\, dx\, dy$, then $f(x, y)$ is called *the joint pdf* of (X, Y), provided $f(x, y)$ satisfies the following conditions.

(i) $f(x, y) \geq 0$, for all $(x, y) \in R$, where R is the range space.

(ii) $\iint\limits_R f(x, y)\, dx\, dy = 1.$

Moreover if D is a subspace of the range space R, $P\{(X, Y) \in D\}$ is defined as

$$P\{(X, Y) \in D\} = \iint\limits_D f(x, y) dx\, dy. \text{ In particular}$$

$$P\{a \leq X \leq b, c \leq Y \leq d\} = \int_c^d \int_a^b f(x, y) dx\, dy$$

Cumulative Distribution Function

If (X, Y) is a two-dimensional RV (discrete or continuous), then $F(x, y) = P\{X \leq x \text{ and } Y \leq y\}$ is called *the cdf of (X, Y).*

In the discrete case,

$$F(x, y) = \sum_{\substack{j \\ y_j \leq y}} \sum_{\substack{i \\ x_i \leq x}} p_{ij}$$

In the continuous case,

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

Properties of $F(x, y)$

(i) $F(-\infty, y) = 0 = F(x, -\infty)$ and $F(\infty, \infty) = 1$
(ii) $P\{a < X < b, Y \leq y\} = F(b, y) - F(a, y)$
(iii) $P\{X \leq x, c < Y < d\} = F(x, d) - F(x, c)$
(iv) $P\{a < X < b, c < Y < d\} = F(b, d) - F(a, d) - F(b, c) + F(a, c)$
(v) At points of continuity of $f(x, y)$

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

Marginal Probability Distribution

$$P(X = x_i) = P\{(X = x_i \text{ and } Y = y_1) \text{ or } (X = x_i \text{ and } Y = y_2) \text{ or etc.}\}$$

$$= p_{i1} + p_{i2} + \ldots = \sum_j p_{ij}$$

$P(X = x_i) = \sum_j p_{ij}$ is called *the marginal probability function of X*. It is defined for $X = x_1, x_2, \ldots$ and denoted as P_{i*}. The collection of pairs $\{x_i, p_{i*}\}$, $i = 1, 2, 3, \ldots$ is called the *marginal probability distribution of X*.
Similarly the collection of pairs $\{y_j, p_{*j}\}$, $j = 1, 2, 3, \ldots$ is called *the marginal probability distribution of Y*, where $p_{*j} = \sum_i P_{ij} = P(Y = y_j)$.

In the continuous case,

$$P\{x - \frac{1}{2}dx \leq X \leq x + \frac{1}{2}dx, -\infty < Y < \infty\}$$

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

$$= \left[\int_{-\infty}^{\infty} f(x, y)\, dy\right] dx$$ [since $f(x, y)$ may be treated a constant in $(x - 1/2dx, x + 1/2\,dx)$]

$= f_X(x)dx$, say

$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$ is called *the marginal density of X*.

Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$ is called *the marginal density of Y*.

Note $P(a \le X \le b) = P(a \le X \le b, -\infty < Y < \infty)$

$$= \int_{-\infty}^{\infty}\int_{a}^{b} f(x, y)\,dx\,dy$$

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

Similarly, $P(c \le Y \le d) = \int_{c}^{d} f_Y(y)\,dy$

Conditional Probability Distribution

$P\{X = x_i/Y = y_j\} = \dfrac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \dfrac{p_{ij}}{p_{*j}}$ is called *the conditional probability function of* X, given that $Y = y_j$.

The collection of pairs, $\left\{x_i, \dfrac{p_{ij}}{p_{*j}}\right\}$ $i = 1, 2, 3, \ldots,$

is called *the conditional probability distribution of* X, *given* $Y = y_j$.

Similarly, the collection of pairs, $\left\{Y_j, \dfrac{p_{ij}}{p_{i*}}\right\}$, $j = 1, 2, 3, \ldots$, is called the *conditional probability distribution of* Y *given* $X = x_i$. In the continuous case,

$$P\left\{x - \frac{1}{2}dx \le X < x + \frac{1}{2}dx \,/\, Y = y\right\}$$

$$= P\left\{x - \frac{1}{2}dx \le X \le x + \frac{1}{2}dx \,/\, y - \frac{1}{2}dy \le Y \le y + \frac{1}{2}dy\right\}$$

$$= \frac{f(x, y)\,dx\,dy}{f_Y(y)\,dy} = \left\{\frac{f(x, y)}{f_Y(y)}\right\}dx.$$

$\dfrac{f(x, y)}{f_Y(y)}$ is called *the conditional density of* X, *given* Y, and is denoted by $f(x/y)$.

Similarly, $\dfrac{f(x, y)}{f_X(y)}$ is called *the conditional density of* Y, given X, and is denoted by $f(y/x)$.

Independent RVs

If (X, Y) is a two-dimensional discrete RV such that $P\{X = x_i/y = y_j\} = P(X = x_i)$ i.e., $\dfrac{p_{ij}}{p_{*j}} = p_{i*}$, i.e., $p_{ij} = p_{i*} \times p_{*j}$ for all i, j then X and Y are said to be independent RVs.

Similarly if (X, Y) is a two-dimensional continuous RV such that $f(x, y) = f_X(x) \times f_Y(y)$, then X and Y are said to be independent RVs.

Random Vectors

Sometimes we may have to be concerned with Random experiments whose outcomes will have 3 or more simultaneous numerical characteristics. To study the outcomes of such random experiments we require knowledge of *n-dimensional random variables* or *random vectors*. For example, the location of a space vehicle in a cartesian co-ordinate system is a three-dimensional random vector.

Most of the concepts introduced above for the two-dimensional case can be extended to the n-dimensional one.

Definitions: A vector X: $[X_1, X_2, \ldots, X_n]$ whose components X_i are RVs is called a *random vector*. $(X_1, X_2, \ldots, X_n)$ can assume all values in some region R_n of the n-dimensional space. R_n is called the *range space*.

The joint distribution function of $(X_1, X_2, \ldots \ldots, X_n)$ is defined as $F(x_1, x_2, \ldots, x_n) = P[X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n]$

The joint pdf of $(X_1, X_2, \ldots, X_n)$ is defined as $f(x_1, x_2, \ldots, x_n)$

$$= \frac{\partial^n F(x_1, x_2 \cdots x_n)}{\partial x_1 \cdot \partial x_2 \cdots \partial x_n}$$ and satisfies the following conditions.

(i) $f(x_1, x_2, \ldots, x_n) \ge 0$, for all $(x_1, x_2, \ldots, x_n)$

(ii) $$\iint_{R_n}\int \cdots \int f(x_1, x_2, \ldots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n = 1$$

(iii) $$P[(X_1, X_2, \ldots, X_n) \in D] = \iint_D \ldots \int f(x_1, x_2, \ldots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n$$ where D is a subset of the range space R_n.

The marginal pdf of any ...