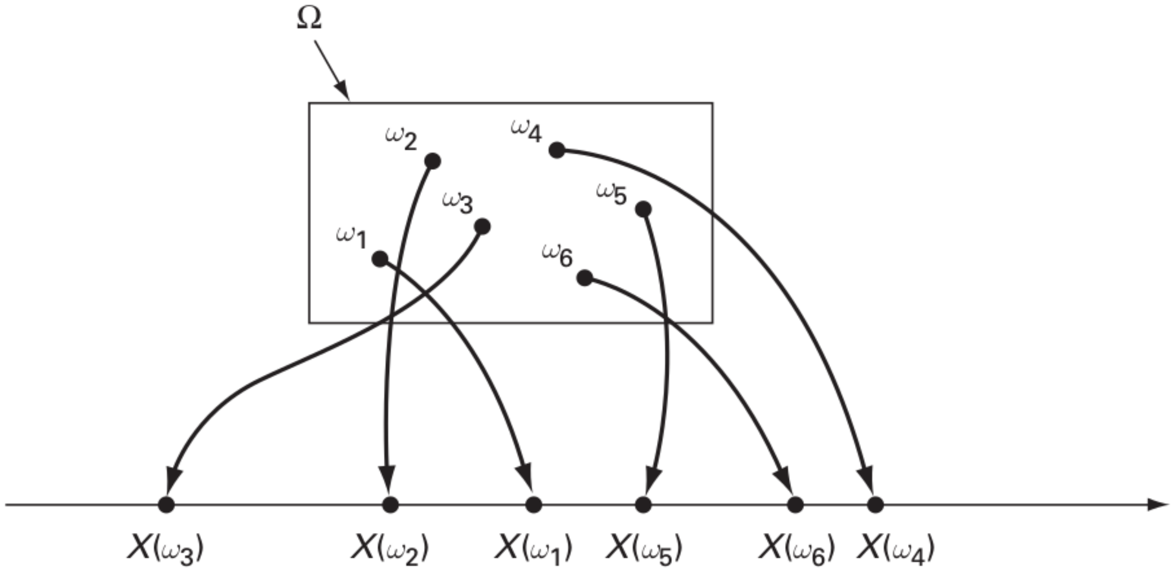
Random Variables

Random Signals & Processes
Lecture 2
Eng. (Mrs.) PN Karunanayake

Random Variables

A real-valued function X (ω) defined on a sample space Ω of points ω is called a random variable

$$X(\omega) = \begin{cases} 1, & \text{if } \omega = \text{head}, \\ 0, & \text{if } \omega = \text{tail}, \end{cases}$$



A random variable $X(\omega)$ as a mapping from Ω to the real line.

Consider a **random variable** as a symbol for a number that is going to be produced by a random experiment!

Random Variables

- Once produced, the number is, no longer random and is called a **realization or instance of the RV**.
- The word random applies to the process that produces the number, rather than to the number itself.
- Example 1:

  In the experiment of tossing a fair coin three times

  (i) Find the sample space S

  (ii) If X is the r.v. giving the number of heads obtained, find

  (a) $P(X = 2)$; (b) $P(X < 2)$

Distribution function

- A random variable X is characterized by its distribution function $F_X(x)$:

$$\boxed{F_X(x) \triangleq P\left[\{\omega : X(\omega) \leq x\}\right],}$$

$$F_X(x) = P\left[X \leq x\right].$$

Property 1. $F_X(x) \geq 0, \quad \text{for } -\infty < x < \infty.$

Property 2. $F_X(-\infty) = 0.$

Property 3. $F_X(\infty) = 1.$

Property 4. If $b > a, \; F_X(b) - F_X(a) = P\left[a < X \leq b\right] \geq 0.$

Example 2

- Consider the r.v. X defined in Example 1. Find and sketch the cdf $F_X(x)$ of X.

Fx(x) = P(X <= x) for x = - 1, 0, 1 , 2, 3, 4

x	$(X \leq x)$	$F_X(x)$
-1	$\varnothing$	0
0	(TTT)	$\frac{1}{8}$
1	$(TTT, TTH, THT, HTT\}$	$\frac{4}{8} = \frac{1}{2}$
2	$\{TTT, TTH, THT, HTT, HHT, HTH, THH\}$	$\frac{7}{8}$
3	S	1
4	S	1

Example 3

- Consider the experiment of throwing a fair die. Let X be the r.v. which assigns 1 if the number that appears is even and 0 if the number that appears is odd.

  (a) What is the range of X ?

  (b) Find P(X = 1) and P(X = 0).

Example 4

- An information source generates symbols at random from. a four-letter alphabet (a, b, c, d} with probabilities P(a) = 1/2, P(b) = 1/4, and P(c) = P(d) = 1/8. A coding scheme encodes these symbols into binary codes as follows:

  A 0

  B 10

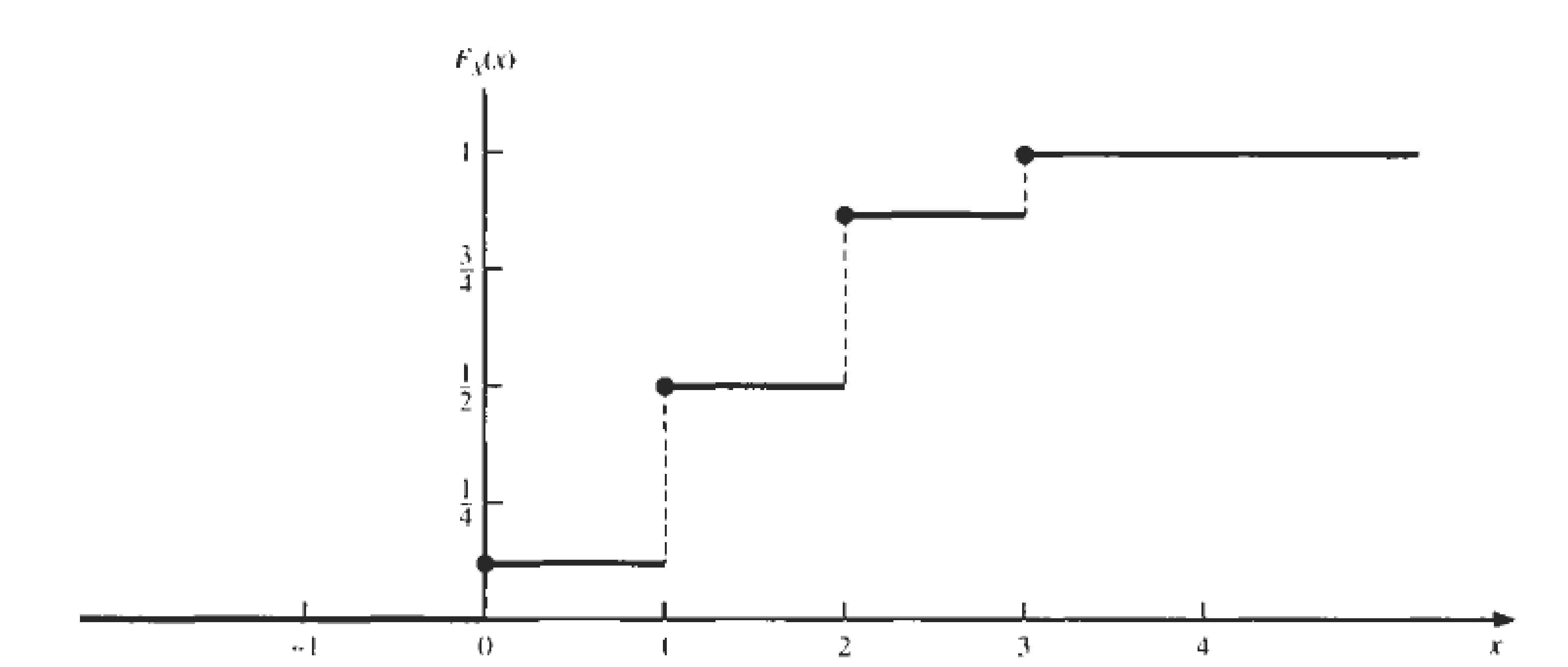
  C 110

  D 111

  Let X be the r.v. denoting the length of the code, that is, the number of binary symbols (bits).

  (a) What is the range of X ?

  (b) Assuming that the generations of symbols are independent, find the probabilities P(X =1), P(X = 2), P(X= 3), and P(X > 3).



$F_X(x)$
1
$\frac{3}{4}$
$\frac{1}{2}$
$\frac{1}{4}$
-1
0
1
2
3
4
x

Two random variables and joint distribution function

- Given functions X (ω) and Y (ω) defined on the sample space , we define the joint distribution function F_{XY} (x, y) of the RVs X and Y by,

$$F_{XY}(x, y) \triangleq P[\{\omega : X(\omega) \leq x, Y(\omega) \leq y\}] = P[X \leq x, Y \leq y].$$

Property 1. $F_{XY}(x, y) \geq 0;$ for $-\infty < x < \infty, \quad -\infty < y < \infty.$

Property 2. $F_{XY}(x, -\infty) = 0;$ for $-\infty < x < \infty,$

$F_{XY}(-\infty, y) = 0;$ for $-\infty < y < \infty.$

Property 3. $F_{XY}(\infty, \infty) = 1.$

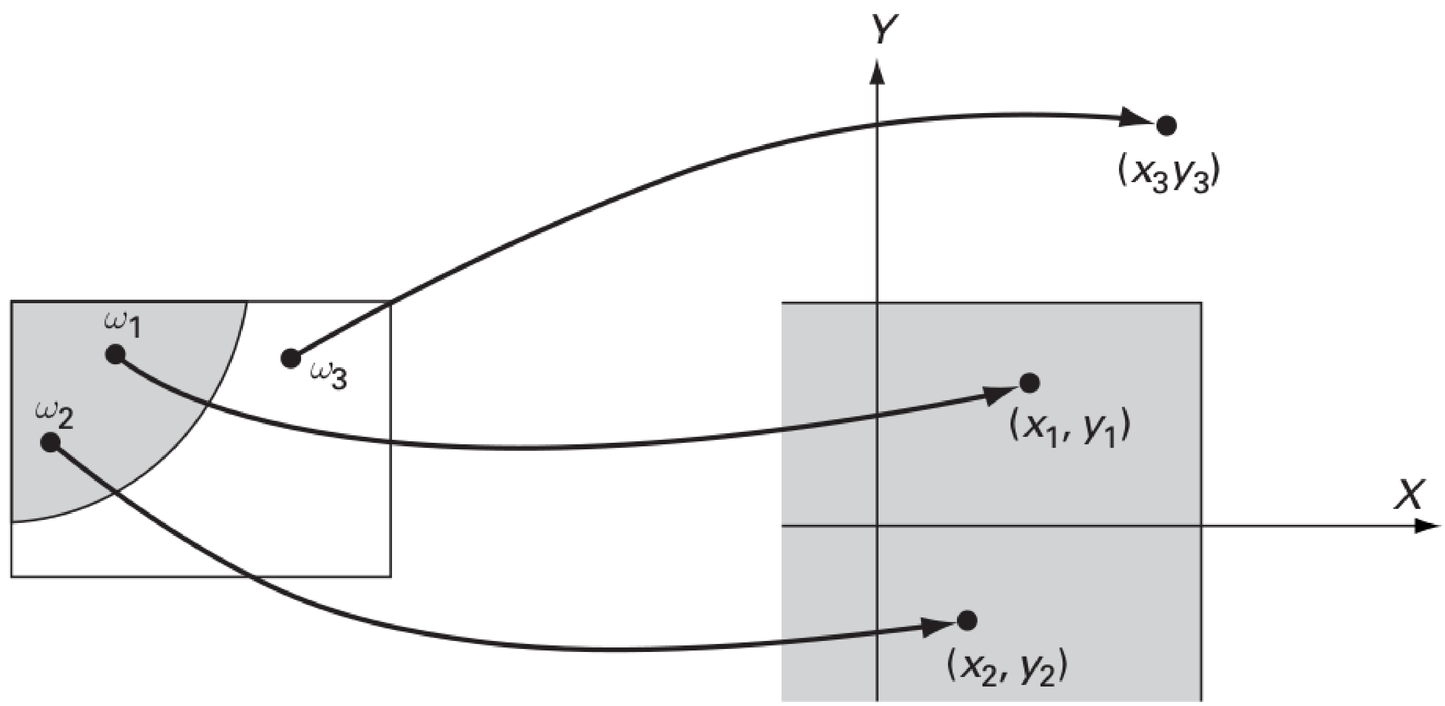
Property 4. If $b > a$ and $d > c$,

$F_{XY}(b, d) \geq F_{XY}(b, c) \geq F_{XY}(a, c).$

Property 5. $F_{XY}(x, \infty) = F_X(x),$

$F_{XY}(\infty, y) = F_Y(y).$

Two random variables



RVs $X(\omega)$ and $Y(\omega)$ as mappings from Ω to the two-dimensional Euclidean space.

m-dimensional joint distribution function $F_X(x)$

$$F_X(\boldsymbol{x}) = P\left[\{\omega : X_1(\omega) \le x_1, X_2(\omega) \le x_2, \ldots, X_m(\omega) \le x_m\}\right]$$
$$= P\left[X_1 \le x_1, X_2 \le x_2, \ldots, X_m \le x_m\right],$$

- where $x = (x_1, x_2, \ldots, x_m)$. We refer to X as an m-dimensional vector of RVs or, simply, as a random vector

Discrete random variables and probability distributions

- Random variable X is called a discrete random variable (discrete RV) if the range of the function X (ω) consists of isolated points on the real line.
- For a discrete RV X ,p_X (x_i) denotes, the probability that X takes the value x_i

$$p_X(x_i) \triangleq P[X = x_i], \quad i = 1, 2, \ldots$$

Cumulative Distribution Function (CDF)

- The complete set of probabilities { p_X (x_i)} associated with the possible values x_i of X is called the probability distribution of the discrete RV X . The probability distribution and the distribution function are related by:

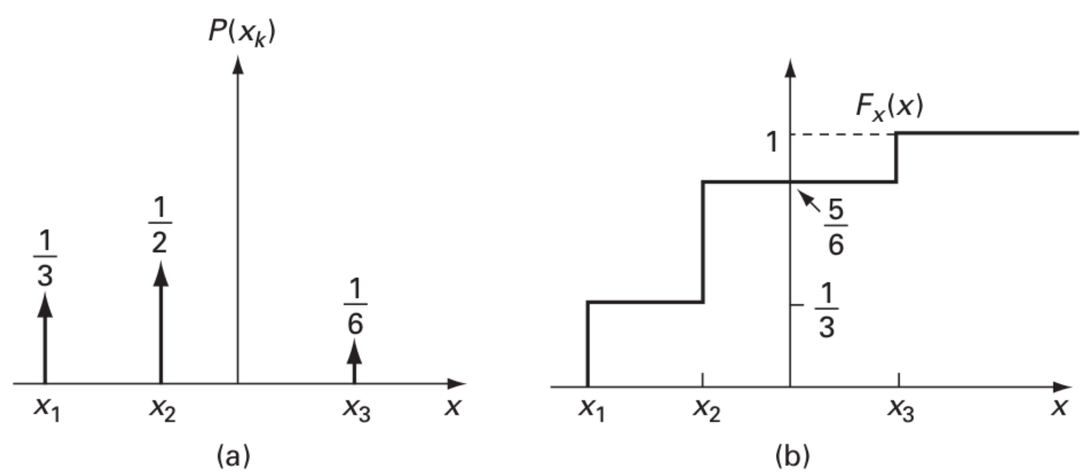
$$F_X(x) = \sum_{x_i \le x} p_X(x_i).$$

Probability Mass Function (PMF)

- PMF provides an equivalent characterization of the discrete RV X as its probability distribution $\{ p_X (x_i)\}$

The function, $p_X : \mathbb{R} \to [0, 1]$, defined by

$$\boxed{p_X(x) = P[X = x], \quad x \in \mathbb{R},}$$



(a) The probability distribution and (b) the distribution function of a discrete RV.

- Suppose that the jumps in $F_X(x)$ of a discrete r.v. X occur at the points x_1, x_2, … ,sequence may be either finite or countably infinite, and we assume $x_i < x_j$ if $i < j$.

- Then $F_X(x_i) - F_X(x_{i-1},) = P(X <= x_i) - P(X <= x_{i-1}) = P(X = x_i)$
- Let $px(x) = P(X = x)$

- The function px(x) is called the probability mass function (pmf) of the discrete r.v. X.

Properties of px(x)

1. $0 \le p_X(x_k) \le 1 \qquad k = 1, 2, \ldots$
2. $p_X(x) = 0 \qquad \text{if } x \ne x_k \ (k = 1, 2, \ldots)$
3. $\sum_k p_X(x_k) = 1$

The cdf $F_X(x)$ of a discrete r.v. X can be obtained by

$$F_X(x) = P(X \le x) = \sum_{x_k \le x} p_X(x_k)$$

Example 5

- Suppose a discrete r.v. X has the following pmfs:

  $p_X(1) = 1/2$
- $p_X(2) = 1/4$
- $p_X(3) = 1/8$
- Px(4) = 1/8

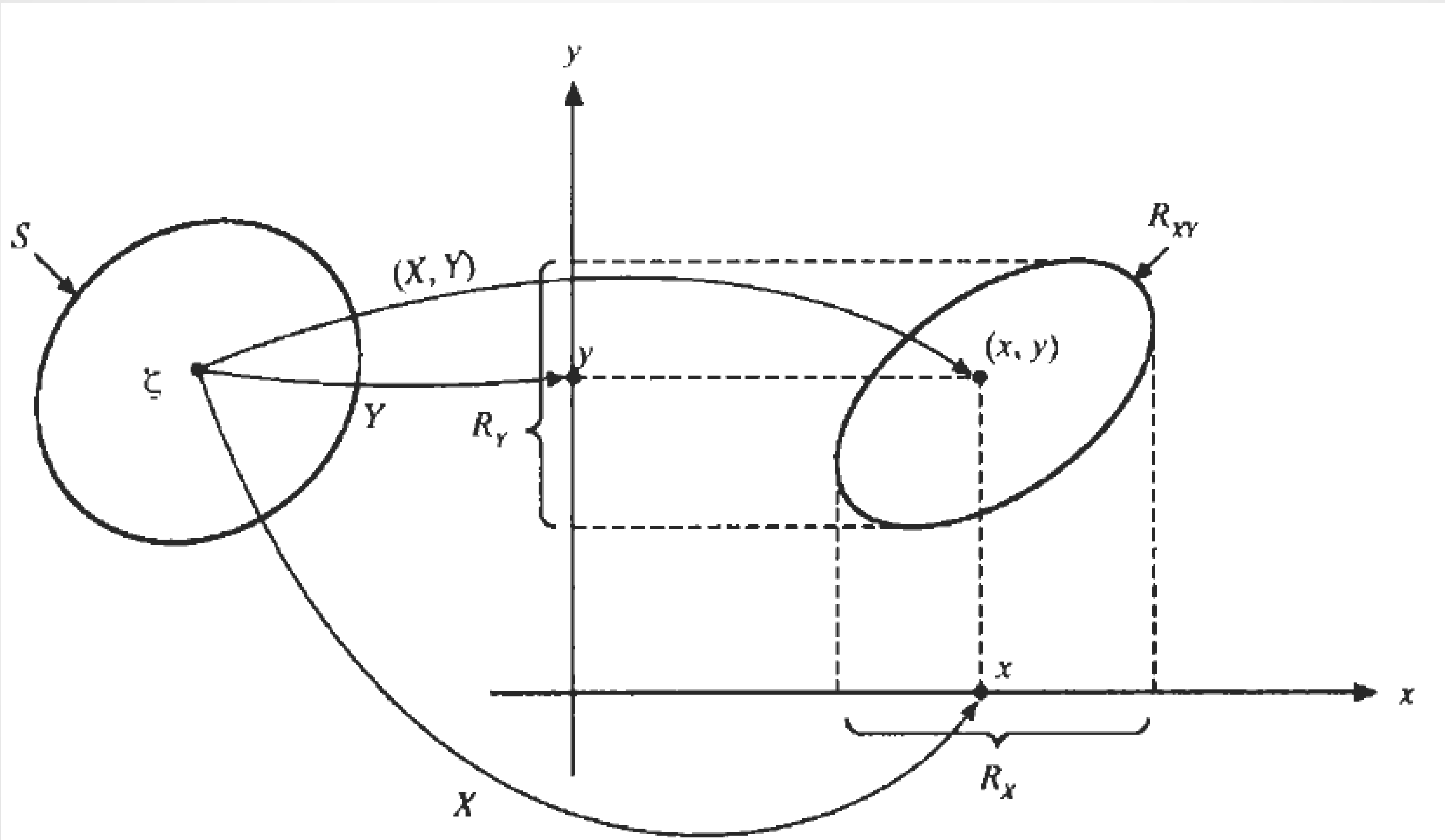
(a) Find and sketch the cdf $F_x(x)$ of the r.v. X.

(b) Find (i) P (X <= I), (ii) P(I < X <= 3), (iii) P (I <= X <=3).

Bivariate Random Variables

- Let **S be the sample space** of a random experiment.
- Let X and Y be two r.v.'s. **Then the pair is called a bivariate** r.v. (or two-dimensional random vector), If each of X and Y associates a real number with every element of S.
- Thus, the bivariate r.v. (X, Y) can be considered as a function that to each point ζ in S assigns a point (x, y) in the plane.
- The range space of the bivariate r.v. (X,Y) is denoted by R_{XY} and defined by

$$R_{xy} = \{(x, y);\ \zeta \in S \text{ and } X(\zeta) = x,\ Y(\zeta) = y\}$$



y
S
R_{XY}
(X, Y)
y
(x, y)
ζ
Y
R_Y
x
x
R_X
X

Joint Distribution Functions

- The joint cumulative distribution function (or joint cdf) of X and Y , denoted by $F_{XY}(x, y)$, is the function defined by,

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

The event (X <= x, Y <=y) is equivalent to the event A n B, where A and B are events of S defined by,

$$A = \{\zeta \in S;\ X(\zeta) \leq x\} \qquad \text{and} \qquad B = \{\zeta \in S;\ Y(\zeta) \leq y\}$$

$$P(A) = F_X(x) \qquad\qquad P(B) = F_Y(y)$$

$$F_{XY}(x, y) = P(A \cap B)$$

Independent Random Variables

- If, for particular values of x and y, A and B were independent events of S , then

$$F_{XY}(x, y) = P(A \cap B) = P(A)P(B) = F_X(x)F_Y(y)$$

Properties of $F_{XY}(x,y)$

1. $0 \le F_{XY}(x, y) \le 1$
2. If $x_1 \le x_2$, and $y_1 \le y_2$, then

$$F_{XY}(x_1, y_1) \le F_{XY}(x_2, y_1) \le F_{XY}(x_2, y_2)$$
$$F_{XY}(x_1, y_1) \le F_{XY}(x_1, y_2) \le F_{XY}(x_2, y_2)$$

3. $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} F_{XY}(x, y) = F_{XY}(\infty, \infty) = 1$
4. $\lim\limits_{x \to -\infty} F_{XY}(x, y) = F_{XY}(-\infty, y) = 0$

$$\lim_{y \to -\infty} F_{XY}(x, y) = F_{XY}(x, -\infty) = 0$$

5. $\lim_{x \to a^+} F_{XY}(x, y) = F_{XY}(a^+, y) = F_{XY}(a, y)$

$\lim_{y \to b^+} F_{XY}(x, y) = F_{XY}(x, b^+) = F_{XY}(x, b)$

6. $P(x_1 < X \le x_2, Y \le y) = F_{XY}(x_2, y) - F_{XY}(x_1, y)$

$P(X \le x, y_1 < Y \le y_2) = F_{XY}(x, y_2) - F_{XY}(x, y_1)$

7. If $x_1 \le x_2$ and $y_1 \le y_2$, then

$$F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \ge 0$$

Marginal Distribution Functions

Now $$\lim_{y \to \infty} (X \le x, Y \le y) = (X \le x, Y \le \infty) = (X \le x)$$

since the condition $y \le \infty$ is always satisfied. Then

$$\lim_{y \to \infty} F_{XY}(x, y) = F_{XY}(x, \infty) = F_X(x)$$

Similarly, $$\lim_{x \to \infty} F_{XY}(x, y) = F_{XY}(\infty, y) = F_Y(y)$$

Discrete Random Variables – Joint Probability Mass Function

- Joint Probability Mass Functions

  Let (X, Y) be a discrete bivariate r.v., and let (X, Y) take on the values (x_i, y_j) for a certain allowable set of integers i and j. Let

$$p_{XY}(x_i, y_j) = P(X = x_i, Y = y_j)$$

The function $p_{xy}(x_i, y_j)$ is called the joint probability mass function (joint pmf) of (X, Y).

Properties of $p_{XY}(x,y)$

1. $0 \le p_{XY}(x_i, y_j) \le 1$
2. $\sum_{x_i} \sum_{y_j} p_{XY}(x_i, y_j) = 1$
3. $P[(X, Y) \in A] = \sum\sum_{(x_i, y_j) \in R_A} p_{XY}(x_i, y_j)$

- where the summation is over the points (x_i, y_j) in the range space RA corresponding to the event A. The joint cdf of a discrete bivariate r.v. (X, Y) is given by

$$F_{XY}(x, y) = \sum_{x_i \le x} \sum_{y_j \le y} p_{XY}(x_i, y_j)$$

Marginal Probability Mass Functions

- Suppose that for a fixed value $X = x_i$, the r.v. Y can take on only the possible values yj (j = 1, 2,. . . , n). Then

$$P(X = x_i) = p_X(x_i) = \sum_{y_j} p_{XY}(x_i, y_j)$$

where the summation is taken over all possible pairs (x_i, y_j) with x_i fixed. Similarly,

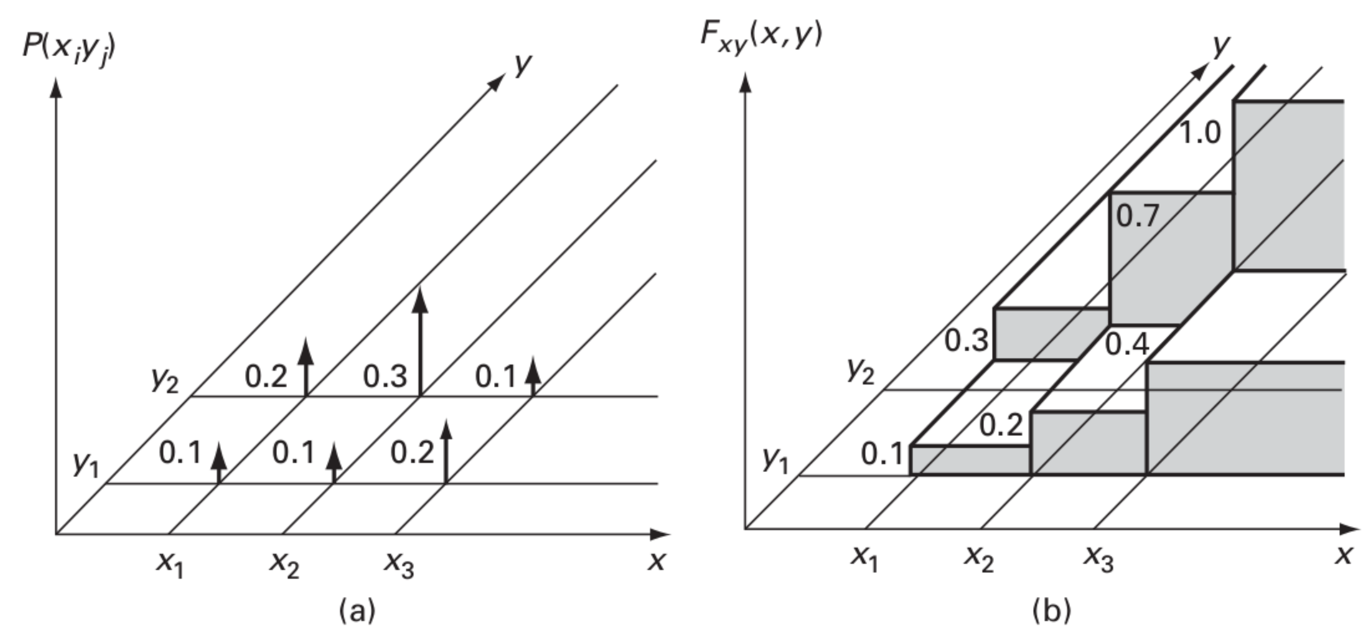
$$P(Y = y_j) = p_Y(y_j) = \sum_{x_i} p_{XY}(x_i, y_j)$$

where the summation is taken over all possible pairs (x_i, y_j) with y_i fixed. Similarly,

Joint and conditional probability distributions

- Joint probability distribution of two discrete Rvs X and Y as the set of probabilities { $p_{XY}(x_i, y_j)$} for all possible values of the pairs (x_i, y_j)

$$p_{XY}(x_i, y_j) \triangleq P[X = x_i, Y = y_j]$$



(a) The joint probability distribution and (b) the joint distribution function.

Joint Distribution Function

$$F_{XY}(x, y) = \sum_{x_i \le x} \sum_{y_j \le y} p_{XY}(x_i, y_j).$$

$$F_{XY}(\infty, \infty) = \sum_{\text{all } i} \sum_{\text{all } j} p_{XY}(x_i, y_j) = 1.$$

Conditional Probability

If (X, Y) is a discrete bivariate r.v. with joint pmf $p_{XY}(x_i, y_j)$, then the conditional pmf of Y , given that X = x_i is defined by

$$p_{Y|X}(y_j \mid x_i) = \frac{p_{XY}(x_i, y_j)}{p_X(x_i)} \qquad p_X(x_i) > 0$$

Similarly, we can define $p_{X|Y}(x_i \mid y_j)$ as

$$p_{X|Y}(x_i \mid y_j) = \frac{p_{XY}(x_i, y_j)}{p_Y(y_j)} \qquad p_Y(y_j) > 0$$

Properties o f $p_{Y|X}(y_j|x_i)$

1. $0 \le p_{Y|X}(y_j \mid x_i) \le 1$
2. $\sum_{y_j} p_{Y|X}(y_j \mid x_i) = 1$

- if X and Y are independent,

$$p_{Y|X}(y_j \mid x_i) = p_Y(y_j) \qquad \text{and} \qquad p_{X|Y}(x_i \mid y_j) = p_X(x_i)$$

Conditional Distribution Function

$$F_{X|Y}(x|y) \triangleq P[X \leq x|Y = y] = \sum_{x_i \leq x} p_{X|Y}(x_i|y),$$

Independent Random Variables

$$\boxed{p_{XY}(x_i, y_j) = p_X(x_i)\,p_Y(y_j), \quad \text{for all values } (x_i, y_j),}$$

or, equivalently, if and only if

$$F_{XY}(x_i, y_j) = F_X(x_i)\,F_Y(y_j), \quad \text{for all values of } x_i \text{ and } y_j.$$

Sample Mean (Emperical Average)

$$\overline{X}_N = \frac{1}{N}\sum_{n=1}^{N} X_n.$$

Ex: Let N (i) denote the number of tosses that result in the integer i, 1 ≤ i ≤ 6, then

$$\overline{X}_N = \frac{1}{N}\sum_{i=1}^{6} iN(i) = \sum_{i=1}^{6} i f_N(i),$$

where f N (i) = N (i)/N is the relative frequency of the outcome i.

when N becomes sufficiently large, f_N (i) will tend to the probability p_X (i). Thus, for large N , X_N stabilize at the value E[X] defined by (E[X] the expectation of X)

$$E[X] \triangleq \sum_{1 \le i \le 6} i p_X(i).$$

Expectation

- The expectation, the expected value, or the mean of a discrete RV X with probability distribution $\{ p_X(x_i) \}$ is defined as:

$$\boxed{\mu_X = E[X] \triangleq \sum_{\text{all } i} x_i p_X(x_i),}$$

expectation extends straightforwardly to a function h(X) of the RV X

$$\boxed{E[h(X)] \triangleq \sum_{\text{all } i} h(x_i) p_X(x_i).}$$

Discrete RVs X and Y are independent if and only if

$$\boxed{E[h(X)g(Y)] = E[h(X)]E[g(Y)]}$$

Moment

- If X is a random variable, so are its k^{th} power X^k and $(X-\mu_X)^k$, define the expectation of these random variables.

$$E[X^k] = \sum_{\text{all } i} x_i^k p_X(x_i),$$

kth moment of X , provided the series converges absolutely.

$$E\left[(X-\mu_X)^k\right] = \sum_{\text{all } i} (x_i - \mu_X)^k p_X(x_i)$$

kth central moment of X .

Variance

- Let X be a RV with finite second moment $E[X^2]$ and mean μ_X

$$\boxed{\sigma_X^2 = \mathrm{Var}[X] \triangleq E[(X-\mu_X)^2] = E[X^2] - \mu_X^2.}$$

The square root of the variance, σ_X, is called the standard deviation.

Conditional Variance

- Let X and Y be discrete RVs. The conditional variance of X given Y is defined as

$$\boxed{\operatorname{Var}[X|Y] \triangleq E[(X - E[X|Y])^2|Y].}$$

Bernoulli Distribution

- A r.v. X is called a Bernoulli RV with parameter p if its pmf is given by:

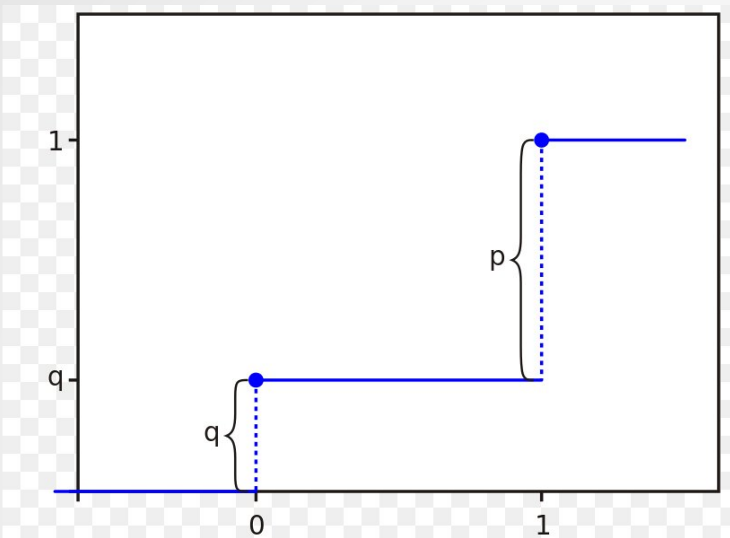
  $p_x(k) = P(X = k) = p^k(l - P)^{1-k}$ $\quad$ $k = 0,1$

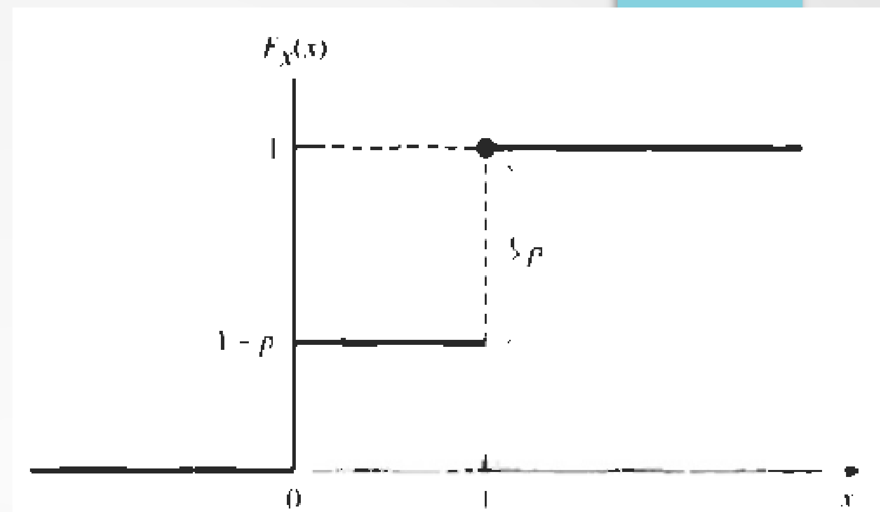
- Where 0 <=k<=1

- The cdf $F_X(x)$ of the Bernoulli RV X is given by,

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

Bernoulli Distribution





PDF

CDF

Mean μ_x= E(X)=p

Variance σ_x^2= Var(X)= p(1-p)

A Bernoulli RV X is associated with some experiment where an outcome can be classified as either a "success" or a "failure," and the probability of a success is p and the probability of a failure is 1 - p. Such experiments are often called Bernoulli trials.

Binomial Distribution

- A RV X is called a binomial RV with parameters (n, p) if its pmf is given by

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n$$

where $0 \le p \le 1$ and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

which is known as the binomial coefficient. The corresponding cdf of X is

$$F_X(x) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \qquad n \le x < n+1$$

Mean $\mu_x = E(X) = np$

Variance $\sigma_x^2 = Var(X) = np(1-p)$

Example

- Binomial distribution for n = 6 and p = 0.6.