RANDOM VARIABLE

INTRODUCTION

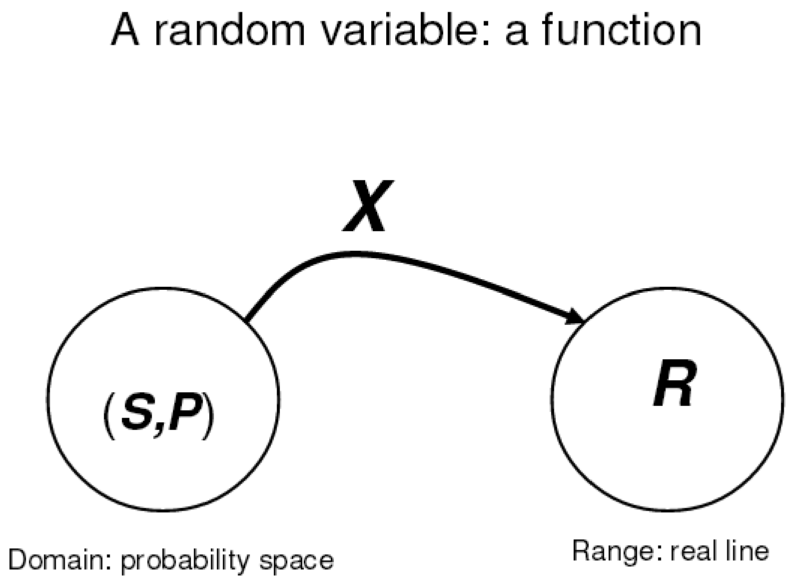
In many situations, we are interested in numbers associated with the outcomes of a random experiment. In application of probabilities, we are often concerned with numerical values which are random in nature. For example, we may consider the number of customers arriving at a service station at a particular interval of time or the transmission time of a message in a communication system. These random quantities may be considered as real-valued function on the sample space. Such a real-valued function is called real random variable and plays an important role in describing random data. We shall introduce the concept of random variables in the following sections.

Random Variable Definition

A random variable is a function that maps outcomes of a random experiment to real numbers. (or)

A random variable associates the points in the sample space with real numbers

A (real-valued) random variable, often denoted by X(or some other capital letter), is a function mapping a probability space (S; P) into the real line R. This is shown in Figure 1.Associated with each point s in the domain S the function X assigns one and only one value X(s) in the range R. (The set of possible values of X(s) is usually a proper subset of the real line; i.e., not all real numbers need occur. If S is a finite set with m elements, then X(s) can assume at most an m different value as s varies in S.)



Example1
A fair coin is tossed 6 times. The number of heads that come up is an example of a random variable.
HHTTHT – 3, THHTTT -- 2.
These random variables can only take values between 0 and 6.
The Set of possible values of random variables is known as its
Range. **Example2**
A box of 6 eggs is rejected once it contains one or more broken eggs. If we examine 10 boxes of eggs and define the randomvariablesX1, X2 as
1 X1- the number of broken eggs in the 10
boxes 2 X2- the number of boxes rejected
Then the range of X1 is {0, 1,2,3,4-------------- 60} and X2 is {0,1,2 --- 10}
Figure 2: A (real-valued) function of a random variable is itself a random variable, i.e., a function mapping a probability space into the real line.

Example 3 Consider the example of tossing a fair coin twice. The sample space is S={ HH,HT,TH,TT} and all four outcomes are equally likely. Then we can define a random variable X as follows

Sample Point	Value of the random Variable
HH	0
HT	1
TH	2
TT	3

Here $R_X = \{0,1,2,3\}$.

Example 4 Consider the sample space associated with the single toss of a fair die. The sample space is given by $S = \{1,2,3,4,5,6\}$.

If we define the random variable X that associates a real number equal to the number on the face of the die, then $X = \{1,2,3,4,5,6\}$.

Types of random variables:

There are two types of random variables, ***discrete*** and ***continuous***.

1. ***Discrete random variable:***

A *discrete random variable* is one which may take on only a countable number of distinct values such as 0, 1,2,3,4,.........Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete

(Or)

A random variable X is called a ***discrete random variable*** if $F_X(x)$ is piece-wise constant. Thus $F_X(x)$ is flat except at the points of jump discontinuity. If the sample space S is discrete the random variable X defined on it is always discrete.

•A discrete random variable has a finite number of possible values or an infinite sequence of countable real numbers.

–X: number of hits when trying 20 free throws.

–X: number of customers who arrive at the bank from 8:30 – 9:30AM Mon--Fri.

–E.g. Binomial, Poisson...

2. ***Continuous random variable:***

A ***continuous random variable*** is one which takes an infinite number of possible values. Continuous rando variables are usually measurements. E

A continuous random variable takes all values in an interval of real numbers.

(or)

X is called a ***continuous random variable*** if $F_X(x)$ ı absolutely continuous function of x. Thus $F_X(x)$ is continuous everywhere on $\mathbb{R}$ and $F_X'(x)$ exists everywhere except at finite or countably infinite points

3. ***Mixed random variable:***

X is called a ***mixed random variable*** if $F_X(x)$ has jump discontinuity at countable number of points and increases continuously at least in one interval of *X*. For a such type RV X.

Conditions for a Function to be a Random Variable:

1. Random variable must not be a multi valued function. i.e two or more values of X cannot assign for Single outcome.
2. P(X=∞)=P(X= -∞)=0
3. The probability of the event (X≤x) must be equal to the sum of probabilities of all the events Corresponding to (X≤x).

Bernoulli's trials:

Bernoulli Experiment with n Trials Here are the rules for a Bernoulli experiment.

1. The experiment is repeated a fixed number of times (n times).
2. Each trial has only two possible outcomes, "success" and "failure". The possible outcomes are exactly the same for each trial.
3. The probability of success remains the same for each trial. We use p for the probability of success (on each trial) and q = 1 − p for the probability of failure.
4. The trials are independent (the outcome of previous trials has no influence on the outcome of the next trial).
5. We are interested in the random variable X where X = the number of successes. Note the possible values of X are 0, 1, 2, 3, . . . , n.

An experiment in which a single action, such as flipping a coin, is repeated identically over and over. The possible results of the action are classified as "success" or "failure". The binomial probability formula is used to find probabilities for Bernoulli trials.

$$P(k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k q^{n-k}$$

n = number of trials
k = number of successes
n – k = number of failures
p = probability of success in one trial
q = 1 – p = probability of failure in one trial

Problem 1:

If the probability of a bulb being defective is 0.8, then what is the probability of the bulb not being defective.

Solution:

Probability of bulb being defective, p = 0.8
Probability of bulb not being defective, q = 1 - p = 1 - 0.8 = 0.2

Problem 2:

10 coins are tossed simultaneously where the probability of getting head for each coin is 0.6. Find the probability of getting 4 heads.

Solution:

Probability of getting head, p = 0.6
Probability of getting head, q = 1 - p = 1 - 0.6 = 0.4
Probability of getting 4 heads out of 10,

$$P(X=4) = C_4^{10}(0.6)^4(0.4)^6 = 0.111476736$$

Probability Distribution Or Probabilty mass function

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

More formally, the probability distribution of a discrete random variable X is a function which gives the probability p(xi) that the random variable equals xi, for each value xi:

p(xi) = P(X=xi)

It satisfies the following conditions:

a. $0 \leq p(xi) \leq 1$
b. $\sum p(xi) = 1$

Cumulative Distribution Function

All random variables (discrete and continuous) have a cumulative distribution function. It is a function giving the probability that the random variable X is less than or equal to x, for every value x.

Formally, the cumulative distribution function F(x) is defined to be:

$$F(x) = P(X \leq x)$$

for

$$-\infty < x < \infty$$

For a discrete random variable, the cumulative distribution function is found by summing up the probabilities as in the example below.

For a continuous random variable, the cumulative distribution function is the integral of its probability density function.

Example

Discrete case : Suppose a random variable X has the following probability distribution p(xi):

xi	0	1	2	3	4	5
p(xi)	1/32	5/32	10/32	10/32	5/32	1/32

This is actually a binomial distribution: Bi(5, 0.5) or B(5, 0.5). The cumulative distribution function F(x) is then:

xi	0	1	2	3	4	5
F(xi)	1/32	6/32	16/32	26/32	31/32	32/32

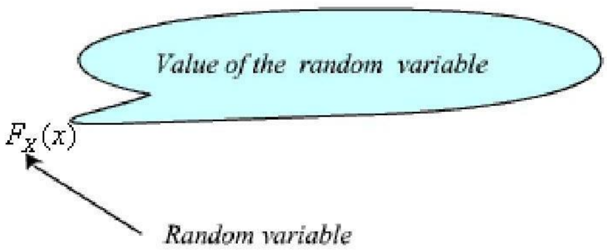
F(x) does not change at intermediate values. For example:

F(1.3) = F(1) = 6/32 and F(2.86) = F(2) = 16/32

Probability Distribution Function

The probability $P(\{X \le x\}) = P(\{s \mid X(s) \le x,\ s \in S\})$ is called the *probability distribution function* (also called the *cumulative distribution function* , abbreviated as *CDF*) of X and denoted by $F_X(x)$. Thus

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



Properties of the Distribution Function

1. $0 \le F_X(x) \le 1$

2. $F_X(x)$ is a *non-decreasing function* of X. Thus, if $x_1 < x_2$, then $F_X(x_1) < F_X(x_2)$

$$\begin{aligned}
& x_1 < x_2 \\
& \Rightarrow \{X(s) \le x_1\} \subseteq \{X(s) \le x_2\} \\
& \Rightarrow P\{X(s) \le x_1\} \le P\{X(s) \le x_2\} \\
& \therefore\ F_X(x_1) < F_X(x_2)
\end{aligned}$$

- $F_X(x)$ Is right continuous.

$$F_X(x^+) = \lim_{\substack{h \to 0 \\ h > 0}} F_X(x+h) = F_X(x)$$

$$\begin{aligned}
\text{Because, } \lim_{\substack{h \to 0 \\ h > 0}} F_X(x+h) &= \lim_{\substack{h \to 0 \\ h > 0}} P\{X(s) \le x+h\} \\
&= P\{X(s) \le x\} \\
&= F_X(x)
\end{aligned}$$

3. $F_X(-\infty) = 0$

Because, $F_X(-\infty) = P\{s \mid X(s) \leq -\infty\} = P(\phi) = 0$.

4. $F_X(\infty) = 1$

5. $P(\{x_1 < X \leq x_2\}) = F_X(x_2) - F_X(x_1) \leq \infty\} = P(S) = 1$

We have,

$$F_X(x^-) = \lim_{\substack{h \to 0 \\ h > 0}} F_X(x-h)$$

$$= \lim_{h \to 0} P\{X(s) \leq x-h\}$$

6. $P(\{X > x\}) = P(\{x < X < \infty\}) = 1 - F_X(x)$

$$= P\{X(s) \leq x\} - P(X(s) = x)$$

$$= F_X(x) - P(X = x)$$

Example:

Value of the random Variable $X = x$	$P(\{X = x\})$
0	1/4
1	1/4
2	1/4
3	1/4

For $x < 0$,

$$F_X(x) = P(\{X \le x\}) = 0$$

For $0 \le x < 1$,

$$F_X(x) = P(\{X \le x\}) = P(\{X = 0\}) = \frac{1}{4}$$

For $1 \le x < 2$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(\{X = 0\} \cup \{X = 1\}) \\ &= P(\{X = 0\}) + P(\{X = 1\}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

For $2 \le x < 3$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\}) \\ &= P(\{X = 0\}) + P(\{X = 1\}) + P(\{X = 2\}) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} \end{aligned}$$

For $x \ge 3$,

$$\begin{aligned} F_X(x) &= P(\{X \le x\}) \\ &= P(S) \\ &= 1 \end{aligned}$$

For $2 \le x < 3$,

$$F_X(x) = P(\{X \le x\})$$
$$= P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\})$$
$$= P(\{X = 0\}) + P(\{X = 1\}) + P(\{X = 2\})$$
$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

Thus we have seen that given $F_X(x),\ -\infty < x < \infty$, we can determine the probability of any event involving values of the random variable X. Thus $F_X(x)\ \forall x \in X$ is a complete description of the random variable X.

Example Consider the random variable X defined by

$$F_X(x) = 0, \qquad x < -2$$
$$= \frac{1}{8}x + \frac{1}{4}, \quad -2 \le x < 0$$
$$= 1, \qquad x \ge 0$$

Find a) $P(X = 0)$.

b) $P\{X \le 0\}$.

c) $P\{X > 2\}$.

d) $P\{-1 < X \le 1\}$.

Solutio

a) $P(X = 0) = F_X(0^+) - F_X(0^-)$

$$= 1 - \frac{1}{4} = \frac{3}{4}$$

b) $P\{X \le 0\} = F_X(0)$

$$= 1$$

c) $P\{X > 2\} = 1 - F_X(2)$

$$= 1 - 1 = 0$$

Probability Density Function

The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

More formally, the probability density function, f(x), of a continuous random variable X is the derivative of the cumulative distribution function F(x):

$$f(x) = \frac{d}{dx} F(x)$$

Since $F(x) = P(X \leq x)$ it follows that:

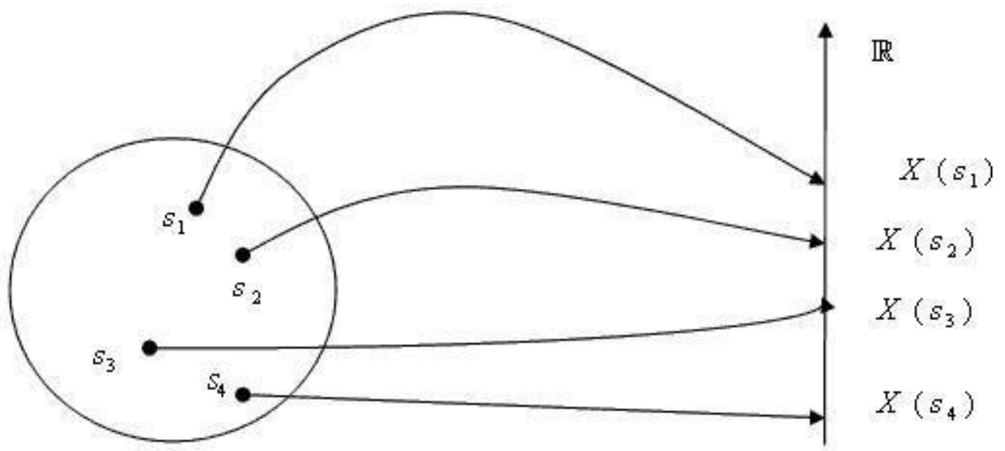
$$\int f(x)\, dx = F(b) - F(a) = P(a < X < b)$$

If f(x) is a probability density function then it must obey two conditions:

a. that the total probability for all possible values of the continuous random variable X is 1:

$$\int f(x)\, dx = 1$$

b. that the probability density function can never be negative: f(x) > 0 for all x.

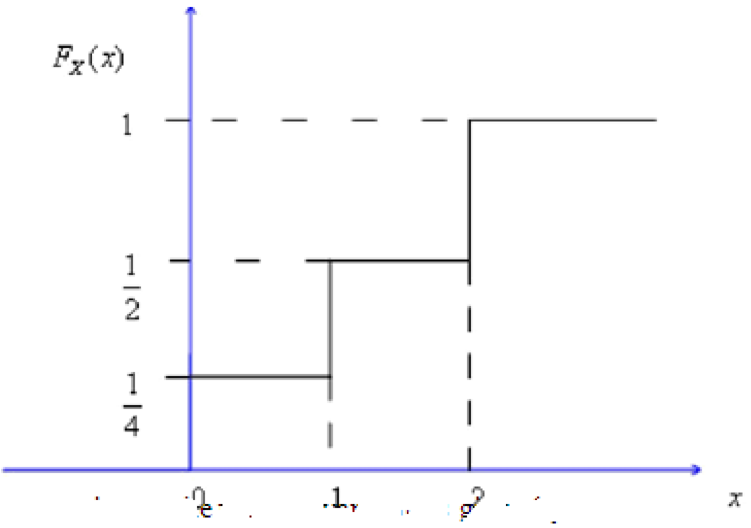


Example 1

Consider the random variable X with the distribution function

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \le x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

The plot of the $F_X(x)$ is shown in Figure



The probability mass function of the random variable is given by

Value of the random variable $X = x$	$p_X(x)$
0	$\frac{1}{4}$
1	$\frac{1}{4}$
2	$\frac{1}{2}$

Properties of the Probability Density Function

1. $f_X(x) \geq 0$.------- This follows from the fact that $F_X(x)$ is a non-decreasing function

2. $F_X(x) = \int_{-\infty}^{x} f_X(u)du$

3. $\int_{-\infty}^{\infty} f_X(x)dx = 1$

4. $P(x_1 < X \leq x_2) = \int_{-x_1}^{x_2} f_X(x)dx$

Conditional Distribution and Density functions:

We discussed conditional probability in an earlier lecture. For two events *A* and *B* with $P(B) \neq 0$, the conditional probability $P(A/B)$ was defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Clearly, the conditional probability can be defined on events involving a Random Variable X

Conditional distribution function:

Consider the event $\{X \leq x\}$ and any event *B* involving the random variable *X* . The conditional distribution function of *X* given *B* is defined as

$$\begin{aligned} F_X(x/B) &= P\left[\{X \leq x\}/B\right] \\ &= \frac{P\left[\{X \leq x\} \cap B\right]}{P(B)} \qquad\qquad P(B) \neq 0 \end{aligned}$$

We can verify that $F_X(x/B)$ satisfies all the properties of the distribution function. Particularly.

- $F_X(-\infty/B) = 0$ And $F_X(\infty/B) = 1$.
- $0 \leq F_X(x/B) \leq 1$.
- $F_X(x/B)$ Is a non-decreasing function of x.
- $$\begin{aligned} P\left(\{x_1 < X \leq x_2\}/B\right) &= P(\{X \leq x_2\}/B) - P(\{X \leq x_1\}/B) \\ &= F_X(x_2/B) - F_X(x_1/B) \end{aligned}$$

Conditional Probability Density Function

In a similar manner, we can define the conditional density function $f_X(x/B)$ of the random variable *X* given the event *B* as

$$f_X(x/B) = \frac{d}{dx} F_X(x/B)$$

All the properties of the pdf applies to the conditional pdf and we can easily show that

- $f_X(x/B) \geq 0$
- $\int_{-\infty}^{\infty} f_X(x/B)dx = F_X(\infty/B) = 1$
- $F_X(x/B) = \int_{-\infty}^{x} f_X(u/B)du$

$$\begin{aligned} P(\{x_1 < X \leq x_2\}/B) &= F_X(x_2/B) - F_X(x_1/B) \\ &= \int_{x_1}^{x_2} f_X(x/B)dx \end{aligned}$$

OPERATIONS ON RANDOM VARIABLE

Expected Value of a Random Variable:

- The *expectation* operation extracts a few parameters of a random variable and provides a summary description of the random variable in terms of these parameters.
- It is far easier to estimate these parameters from data than to estimate the distribution or density function of the random variable.
- Moments are some important parameters obtained through the expectation operation.

Expected value or mean of a random variable

The expected value of a random variable X is defined by

$$EX = \int_{-\infty}^{\infty} x f_X(x)dx$$

provided $\int_{-\infty}^{\infty} x f_X(x)dx$ exists.

EX is also called the mean or statistical average of the random variable X and denoted by μ_X.

Note that, for a discrete RV X defined by the *probability mass function* (pmf) $p_X(x_i), i = 1, 2, ..., N,$ the pdf $f_X(x)$ is given by

$$f_X(x) = \sum_{i=1}^{N} p_X(x_i)\delta(x - x_i)$$

$$\therefore \mu_X = EX = \int_{-\infty}^{\infty} x \sum_{i=1}^{N} p_X(x_i)\delta(x - x_i)dx$$

$$= \sum_{i=1}^{N} p_X(x_i) \int_{-\infty}^{\infty} x\delta(x - x_i)dx$$

$$= \sum_{i=1}^{N} x_i p_X(x_i)$$

Thus for a discrete random variable X with $p_X(x_i), i = 1, 2, ..., N,$

$$\mu_X = \sum_{i=1}^{N} x_i p_X(x_i)$$

Interpretation of the mean

- The mean gives an idea about the average value of the random value. The values of the random variable are spread about this value.
- Observe that

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x)dx$$

$$= \frac{\int_{-\infty}^{\infty} x f_X(x)dx}{\int_{-\infty}^{\infty} f_X(x)dx} \qquad \because \int_{-\infty}^{\infty} f_X(x)dx = 1$$

Therefore, the mean can be also interpreted as the *centre of gravity* of the pdf curve.

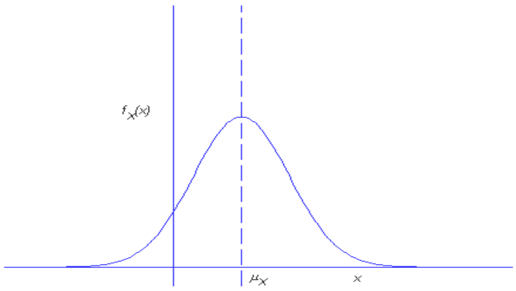


Fig. Mean of a random variable

Example 1 Suppose X is a random variable defined by the pdf

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x f_X(x)dx \\ &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{a+b}{2} \end{aligned}$$

Example 2 Consider the random variable X with pmf as tabulated below

Value of the random variable x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$

$$\begin{aligned} \therefore \quad \mu_X &= \sum_{i=1}^{N} x_i p_X(x_i) \\ &= 0 \times \frac{1}{8} + 1 \times \frac{1}{8} + 2 \times \frac{1}{4} + 3 \times \frac{1}{2} \\ &= \frac{17}{8} \end{aligned}$$

Remark If $f_X(x)$ is an even function of x, then $\int_{-\infty}^{\infty} x f_X(x)dx = 0.$ Thus the mean of a RV with an even symmetric pdf is 0.

Expected value of a function of a random variable

Suppose $Y = g(X)$ is a function of a random variable X as discussed in the last class. Then,

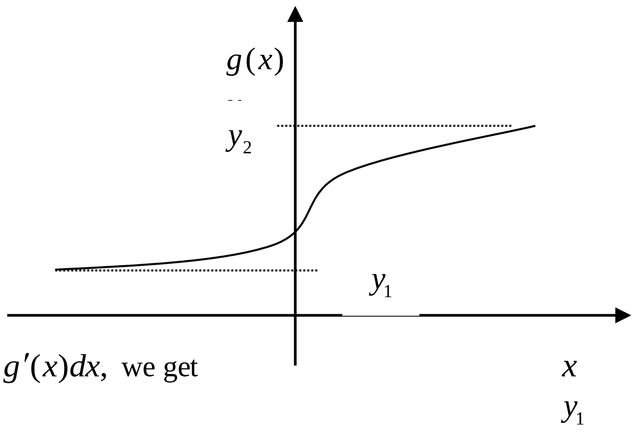
$$EY = Eg(X) = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

We shall illustrate the theorem in the special case $g(X)$ when $y = g(x)$ is one-to-one and monotonically increasing function of x. In this case,

$$f_Y(y) = \left.\frac{f_X(x)}{g'(x)}\right]_{x=g^{-1}(y)}$$

$$EY = \int_{-\infty}^{\infty} y f_Y(y)dy$$

$$= \int_{y_1}^{y2} y \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))} dy$$



where $y_1 = g(-\infty)$ and $y_2 = g(\infty)$.

Substituting $x = g^{-1}(y)$ so that $y = g(x)$ and $dy = g'(x)dx$, we get

$$EY = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

The following important properties of the expectation operation can be immediately derived:

(a) If c is a constant,

$$\boxed{Ec = c}$$

Clearly

$$Ec = \int_{-\infty}^{\infty} c f_X(x)dx = c\int_{-\infty}^{\infty} f_X(x)dx = c$$

(b) If $g_1(X)$ and $g_2(X)$ are two functions of the random variable X and c_1 and c_2 are constants,

$$\boxed{E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E g_1(X) + c_2 E g_2(X)}$$

$$E[c_1 g_1(X) + c_2 g_2(X)] = \int_{-\infty}^{\infty} c_1[g_1(x) + c_2 g_2(x)] f_X(x)dx$$

$$= \int_{-\infty}^{\infty} c_1 g_1(x) f_X(x)dx + \int_{-\infty}^{\infty} c_2 g_2(x) f_X(x)dx$$

$$= c_1 \int_{-\infty}^{\infty} g_1(x) f_X(x)dx + c_2 \int_{-\infty}^{\infty} g_2(x) f_X(x)dx$$

$$= c_1 E g_1(X) + c_2 E g_2(X)$$

The above property means that E is a linear operator.

Mean-square value

$$EX^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx$$

Variance

For a random variable X with the pdf $f_X(x)$ and men μ_X, the variance of X is denoted by σ_X^2 and defined as

$$\sigma_X^2 = E(X - \mu_X)^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx$$

Thus for a discrete random variable X with $p_X(x_i), i = 1, 2, ..., N,$

$$\sigma_X^2 = \sum_{i=1}^{N} (x_i - \mu_X)^2 p_X(x_i)$$

The standard deviation of X is defined as

$$\sigma_X = \sqrt{E(X-\mu_X)^2}$$

Example3: Find the variance of the random variable discussed in Example 1.

$$\begin{aligned}\sigma_X^2 &= E(X-\mu_X)^2\\ &= \int_a^b (x-\frac{a+b}{2})^2 \frac{1}{b-a}dx\\ &= \frac{1}{b-a}[\int_a^b x^2dx - 2\times\frac{a+b}{2}\int_a^b xdx + \left(\frac{a+b}{2}\right)^2\int_a^b dx\\ &= \frac{(b-a)^2}{12}\end{aligned}$$

Example4: Find the variance of the random variable discussed in Example 2.

As already computed

$$\begin{aligned}\mu_X &= \frac{17}{8}\\ \sigma_X^2 &= E(X-\mu_X)^2\\ &= (0-\frac{17}{8})^2\times\frac{1}{8}+(1-\frac{17}{8})^2\times\frac{1}{8}+(2-\frac{17}{8})^2\times\frac{1}{4}+(3-\frac{17}{8})^2\times\frac{1}{2}\\ &= \frac{117}{128}\end{aligned}$$

Remark

- Variance is a central moment and measure of dispersion of the random variable about the mean.
- $E(X-\mu_X)^2$ is the average of the square deviation from the mean. It gives information about the deviation of the values of the RV about the mean. A smaller σ_X^2 implies that the random values are more clustered about the mean, Similarly, a bigger σ_X^2 means that the random values are more scattered.

For example, consider two random variables X_1 and X_2 with pmf as shown below. Note that each of X_1 and X_2 has zero means. $\sigma_{X_1}^2 = \frac{1}{2}$ and $\sigma_{X_2}^2 = \frac{5}{3}$ implying that X_2 has more spread about the mean

Properties of variance:

(1) $\sigma_X^2 = EX^2 - \mu_X^2$

$$\begin{aligned}\sigma_X^2 &= E(X-\mu_X)^2 \\ &= E(X^2 - 2\mu_X X + \mu_X^2) \\ &= EX^2 - 2\mu_X EX + E\mu_X^2 \\ &= EX^2 - 2\mu_X^2 + \mu_X^2 \\ &= EX^2 - \mu_X^2\end{aligned}$$

(2) If $Y = cX + b$, where c and b are constants, then $\sigma_Y^2 = c^2\sigma_X^2$

$$\begin{aligned}\sigma_Y^2 &= E(cX + b - c\mu_X - b)^2 \\ &= Ec^2(X-\mu_X)^2 \\ &= c^2\sigma_X^2\end{aligned}$$

(3) If c is a constant,

$\text{var}(c) = 0.$

MULTIPLE RANDOM VARIABLES

Multiple Random Variables

In many applications we have to deal with more than two random variables. For example, in the navigation problem, the position of a space craft is represented by three random variables denoting the x, y and z coordinates. The noise affecting the R, G, B channels of color video may be represented by three random variables. In such situations, it is convenient to define the vector-valued random variables where each component of the vector is a random variable.

In this lecture, we extend the concepts of joint random variables to the case of multiple random variables. A generalized analysis will be presented for random variables defined on the same sample space.

Example1: Suppose we are interested in studying the height and weight of the students in a class. We can define the joint RV (X,Y) where X represents height and Y represents the weight.

Example 2 Suppose in a communication system X is the transmitted signal and Y is the corresponding noisy received signal. Then (X,Y) is a joint random variable.

Joint Probability Distribution Function:

Recall the definition of the distribution of a single random variable. The event $\{X \leq x\}$ was used to define the probability distribution function $F_X(x)$. Given $F_X(x)$, we can find the probability of any event involving the random variable. Similarly, for two random variables X and Y, the event $\{X \leq x, Y \leq y\} = \{X \leq x\} \cap \{Y \leq y\}$ is considered as the representative event.

The probability $P\{X \leq x, Y \leq y\}\ \forall (x,y) \in \mathbb{R}^2$ is called the *joint distribution function of the random variables* X *and* Y and denoted by $F_{X,Y}(x,y)$.

Properties of Joint Probability Distribution Function:

The joint CDF satisfies the following properties:

1. $F_X(x)=F_{XY}(x,\infty)$, for any x (marginal CDF of X);
   Proof:
   $$\{X \leq x\} = \{X \leq x\} \cap \{Y \leq +\infty\}$$
   $$\therefore F_X(x) = P(\{X \leq x\}) = P(\{X \leq x, Y \leq \infty\}) = F_{XY}(x, +\infty)$$

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

2. $F_Y(y)=F_{XY}(\infty,y)$, for any y (marginal CDF of Y);
3. $F_{XY}(\infty,\infty)=1$;
4. $F_{XY}(-\infty,y)=F_{XY}(x,-\infty)=0$;
5. $P(x1<X\leq x2, y1<Y\leq y2)= F_{XY}(x2,y2)-F_{XY}(x1,y2)-F_{XY}(x2,y1)+F_{XY}(x1,y1)$;
6. if X and Y are independent, then $F_{XY}(x,y)=F_X(x)F_Y(y)$
7. $F_{X,Y}(x_1,y_1) \leq F_{X,Y}(x_2,y_2)$ if $x_1 \leq x_2$ and $y_1 \leq y_2$
   Proof:

If $x_1 < x_2$ and $y_1 < y_2$,

$$\{X \le x_1, Y \le y_1\} \subseteq \{X \le x_2, Y \le y_2\}$$

$$\therefore P\{X \le x_1, Y \le y_1\} \le P\{X \le x_2, Y \le y_2\}$$

$$\therefore F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2)$$

Example1:

Consider two jointly distributed random variables X and Y with the joint CDF

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-2x})(1-e^{-y}) & x \ge 0, y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the marginal CDFs

(b) Find the probability $P\{1 < X \le 2, \ \ 1 < Y \le 2\}$

Solution:

(a)

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) = \begin{cases} 1-e^{-2x} & x \ge 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y) = \begin{cases} 1-e^{-y} & y \ge 0 \\ 0 & \text{elsewhere} \end{cases}$$

(b)

$$\begin{aligned} P\{1 < X \le 2, \ \ 1 < Y \le 2\} &= F_{X,Y}(2,2) + F_{X,Y}(1,1) - F_{X,Y}(1,2) - F_{X,Y}(2,1) \\ &= (1-e^{-4})(1-e^{-2}) + (1-e^{-2})(1-e^{-1}) - (1-e^{-2})(1-e^{-2}) - (1-e^{-4})(1-e^{-1}) \\ &= 0.0272 \end{aligned}$$

Jointly distributed discrete random variables

If X and Y are two discrete random variables defined on the same probability space (S, F, P) such that X takes values from the countable subset R_X and Y takes values from the countable subset R_Y. Then the joint random variable (X, Y) can take values from the countable subset in $R_X \times R_Y$. The joint random variable (X, Y) is completely specified by their *joint probability mass function*

$$p_{X,Y}(x,y) = P\{s \mid X(s) = x, Y(s) = y\}, \quad \forall (x,y) \in R_X \times R_Y.$$

Given $p_{X,Y}(x,y)$, we can determine other probabilities involving the random variables X and Y.

Remark

- $p_{X,Y}(x,y) = 0$ for $(x,y) \notin R_X \times R_Y$
- $\sum_{(x,y) \in R_X \times R_Y} \sum p_{X,Y}(x,y) = 1$

  This is because

$$\begin{aligned} \sum_{(x,y) \in R_X \times R_Y} \sum p_{X,Y}(x,y) &= P\Big(\bigcup_{(x,y) \in R_X \times R_Y} \{x,y\}\Big) \\ &= P(R_X \times R_Y) \\ &= P\{s \mid (X(s), Y(s)) \in (R_X \times R_Y)\} \\ &= P(S) = 1 \end{aligned}$$

- ***Marginal Probability Mass Functions***: The probability mass functions $p_X(x)$ and $p_Y(y)$ are obtained from the joint probability mass function as follows

$$p_X(x) = P\{X = x\} \cup R_Y$$
$$= \sum_{y \in R_Y} p_{X,Y}(x,y)$$

and similarly

$$p_Y(y) = \sum_{x \in R_X} p_{X,Y}(x,y)$$

These probability mass functions $p_X(x)$ and $p_Y(y)$ obtained from the joint probability mass functions are called *marginal probability mass functions*.

Example Consider the random variables X and Y with the joint probability mass function as tabulated in Table . The marginal probabilities are as shown in the last column and the last row

X \ Y	0	1	2	$p_Y(y)$
0	0.25	0.1	0.15	0.5
1	0.14	0.35	0.01	0.5
$p_X(x)$	0.39	0.45		

Joint Probability Density Function

If X and Y are two continuous random variables and their joint distribution function is continuous in both x and y, then we can define *joint probability density function* $f_{X,Y}(x,y)$ by

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y), \text{ provided it exists.}$$

Clearly $F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du$

Properties of Joint Probability Density Function:

- $f_{X,Y}(x,y)$ is always a non-negative quantity. That is,

$$f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$$

- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
- Marginal probability density functions can be defined as

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy, \quad \text{for all } x,$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx, \quad \text{for all } y.$$

- The probability of any Borel set B can be obtained by

$$P(B) = \iint_{(x,y) \in B} f_{X,Y}(x,y)\,dx\,dy$$

Marginal Distribution and density Functions:

The probability distribution functions of random variables X and Y obtained from joint distribution function is called ad marginal distribution functions. i.e.

$F_X(x)=F_{XY}(x,\infty)$, for any x (marginal CDF of X);

Proof:

$$\{X \le x\} = \{X \le x\} \cap \{Y \le +\infty\}$$
$$\therefore F_X(x) = P\left(\{X \le x\}\right) = P\left(\{X \le x, Y \le \infty\}\right) = F_{XY}(x, +\infty)$$

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

The marginal density functions $f_X(x)$ and $f_Y(y)$ of two joint RVs X and Y are given by the derivatives of the corresponding marginal distribution functions. Thus

$$\begin{aligned} f_X(x) &= \tfrac{d}{dx} F_X(x) \\ &= \tfrac{d}{dx} F_X(x, \infty) \\ &= \tfrac{d}{dx} \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u, y) dy \right) du \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \end{aligned}$$

and similarly $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$

The marginal CDF and pdf are same as the CDF and pdf of the concerned single random variable. The *marginal* term simply refers that it is derived from the corresponding joint distribution or density function of two or more jointly random variables.

Example2: The joint density function $f_{X,Y}(x, y)$ in the previous example is

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) \\ &= \frac{\partial^2}{\partial x \partial y}[(1-e^{-2x})(1-e^{-y})] \quad x \ge 0, y \ge 0 \\ &= 2e^{-2x}e^{-y} \quad x \ge 0, y \ge 0 \end{aligned}$$

Example3: The joint pdf of two random variables X and Y are given by

$$\begin{aligned} f_{X,Y}(x, y) &= cxy \quad 0 \le x \le 2,\ 0 \le y \le 2 \\ &= 0 \ \text{ otherwise} \end{aligned}$$

(i) Find c.

(ii) Find $F_{X,y}(x, y)$

(iii) Find $f_X(x)$ and $f_Y(y)$.

(iv) What is the probability $P(0 < X \le 1, 0 < Y \le 1)$?

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = c\int_0^2\int_0^2 xydydx$$
$$\therefore c = \frac{1}{4}$$
$$F_{X,Y}(x,y) = \frac{1}{4}\int_0^{y}\int_0^{x} uvdudv$$
$$= \frac{x^2y^2}{16}$$
$$f_X(x) = \int_0^2 \frac{xy}{4}dy \quad 0 \le y \le 2$$
$$= \frac{x}{2} \quad 0 \le y \le 2$$
Similarly
$$f_Y(y) = \frac{y}{2} \qquad 0 \le y \le 2$$
$$P(0 < X \le 1, 0 < Y \le 1)$$
$$= F_{X,Y}(1,1) + F_{X,Y}(0,0) - F_{X,Y}(0,1) - F_{X,Y}(1,0)$$
$$= \frac{1}{16} + 0 - 0 - 0$$
$$= \frac{1}{16}$$

Conditional Distribution and Density functions

We discussed conditional probability in an earlier lecture. For two events *A* and *B* with $P(B) \neq 0$, the conditional probability $P(A/B)$ was defined as
$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Clearly, the conditional probability can be defined on events involving a random variable *X*.

Conditional distribution function

Consider the event $\{X \le x\}$ and any event *B* involving the random variable *X*. The conditional distribution function of *X* given *B* is defined as

$$F_X(x/B) = P\left[\{X \le x\}/B\right]$$
$$= \frac{P\left[\{X \le x\} \cap B\right]}{P(B)} \qquad P(B) \neq 0$$

Properties of Conditional distribution function

We can verify that $F_X(x/B)$ satisfies all the properties of the distribution function. Particularly.

- $F_X(-\infty/B) = 0$ and $F_X(\infty/B) = 1$.
- $0 \le F_X(x/B) \le 1$

- $F_X(x/B)$ is a non-decreasing function of x.
- $$P(\{x_1 < X \le x_2\}/B) = P(\{X \le x_2\}/B) - P(\{X \le x_1\}/B)$$ $$= F_X(x_2/B) - F_X(x_1/B)$$

Conditional density function

In a similar manner, we can define the conditional density function $f_X(x/B)$ of the random variable X given the event B as

$$f_X(x/B) = \frac{d}{dx}F_X(x/B)$$

Properties of Conditional density function:

All the properties of the pdf applies to the conditional pdf and we can easily show that

- $f_X(x/B) \ge 0$
- $\int_{-\infty}^{\infty} f_X(x/B)dx = F_X(\infty/B) = 1$
- $F_X(x/B) = \int_{-\infty}^{x} f_X(u/B)du$
- $$P(\{x_1 < X \le x_2\}/B) = F_X(x_2/B) - F_X(x_1/B)$$ $$= \int_{x_1}^{x_2} f_X(x/B)dx$$

Let (X, Y) be a discrete bivariate random vector with joint pmf f(x, y) and marginal pmfs $f_X(x)$ and $f_Y(y)$. For any x such that $P(X = x) = f_X(x) > 0$, the conditional pmf of Y given that X = x is the function of y denoted by f(y|x) and defined by

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

For any y such that $P(Y = y) = f_Y(y) > 0$, the conditional pmf of X given that Y = y is the function of x denoted by f(x|y) and defined by

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

Example 1: Suppose X is a random variable with the distribution function $F_X(x)$. Define $B = \{X \le b\}$.

Then

$$F_X(x/B) = \frac{P(\{X \le x\} \cap B)}{P(B)}$$
$$= \frac{P(\{X \le x\} \cap \{X \le b\})}{P\{X \le b\}}$$
$$= \frac{P(\{X \le x\} \cap \{X \le b\})}{F_X(b)}$$

Case 1: $x<b$

Then

$$F_X(x/B) = \frac{P(\{X \le x\} \cap \{X \le b\})}{F_X(b)}$$

$$= \frac{P(\{X \le x\})}{F_X(b)} = \frac{F_X(x)}{F_X(b)}$$

And $f_X(x/B) = \frac{d}{dx}\frac{F_X(x)}{F_X(b)} = \frac{f_X(x)}{f_X(b)}$

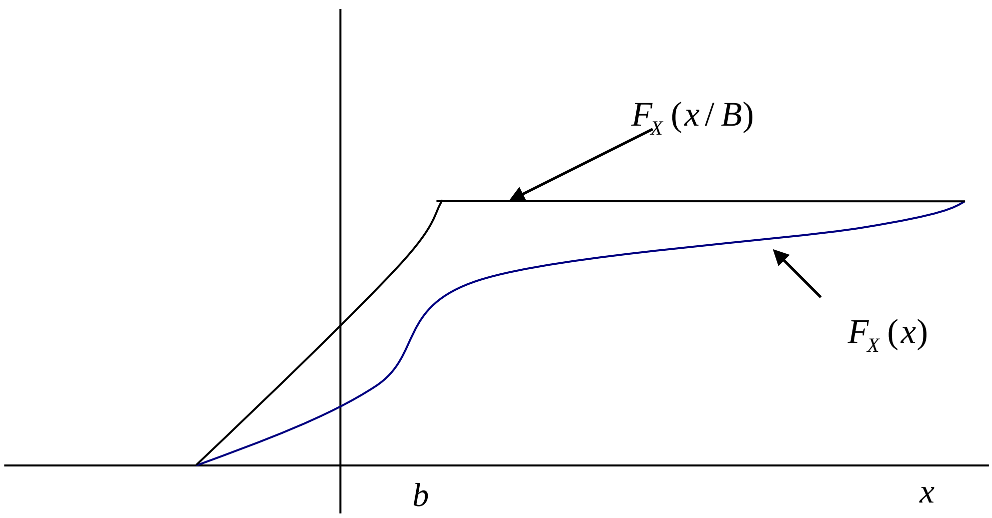
Case 2: $x \ge b$

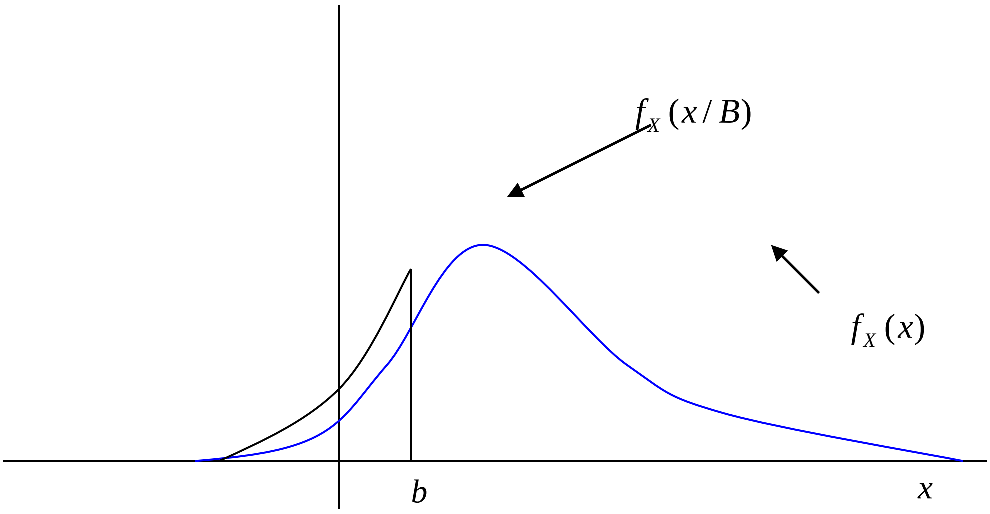
$$F_X(x/B) = \frac{P(\{X \le x\} \cap \{X \le b\})}{F_X(b)}$$

$$= \frac{P(\{X \le x\})}{F_X(b)} = \frac{F_X(b)}{F_X(b)} = 1$$

and $f_X(x/B) = \frac{d}{dx}F_X(x/B) = 0$

$F_X(x/B)$ and $f_X(x/B)$ are plotted in the following figures.





Example 2 Suppose X is a random variable with the distribution function $F_X(x)$ and $B=\{X>b\}$.

Then

$$F_X(x/B)=\frac{P(\{X\le x\}\cap B)}{P(B)}$$
$$=\frac{P(\{X\le x\}\cap\{X>b\})}{P\{X>b\}}$$
$$=\frac{P(\{X\le x\}\cap\{X>b\})}{1-F_X(b)}$$

For $x\le b$, $\{X\le x\}\cap\{X>b\}=\phi$. Therefore,

$$F_X(x/B)=0 \qquad x\le b$$

For $x>b$, $\{X\le x\}\cap\{X>b\}=\{b<X\le x\}$ Therefore,

$$F_X(x/B)=\frac{P(\{b<X\le x\})}{1-F_X(b)}$$
$$=\frac{F_X(x)-F_X(b)}{1-F_X(b)}$$

Thus,

$$F_X(x/B)=\begin{cases}0 & x\le b\\ \dfrac{F_X(x)-F_X(b)}{1-F_X(b)} & \text{otherwise}\end{cases}$$

The corresponding pdf is given by

$$f_X(x/B)=\begin{cases}0 & x\le b\\ \dfrac{f_X(x)}{1-F_X(b)} & \text{otherwise}\end{cases}$$

Example 3 The joint density of X and Y is given by

$$f(x,y)=\begin{cases}\frac{15}{2}x(2-x-y) & 0<x<1, 0<y<1\\ 0 & \text{otherwise}\end{cases}.$$

Compute the condition density of X, given that $Y=y$, where $0<y<1$.

Solution For $0<x<1, 0<y<1$, we have

$$f_X(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{f(x,y)}{\int_{-\infty}^{\infty}f(x,y)\,dx}$$
$$=\frac{x(2-x-y)}{\int_0^1 x(2-x-y)\,dx}=\frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}}=\frac{6x(2-x-y)}{4-3y}.$$

Example4:

Let the continuous random vector (X, Y) have joint pdf

$$f(x,y) = e^{-y}, \quad 0 < x < y < \infty.$$

The marginal of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{x}^{\infty} e^{-y}dy = e6{-}x.$$

Thus, marginally, X has an exponential distribution. The conditional distribution of Y is

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, & \text{if } y > x, \\ \frac{0}{e^{-x}} = 0, & \text{if } y \leq x \end{cases}$$

Conditional Probability Distribution Function

Consider two continuous jointly random variables X and Y with the joint probability distribution function $F_{X,Y}(x,y).$ We are interested to find the conditional distribution function of one of the random variables on the condition of a particular value of the other random variable.

We *cannot* define the conditional distribution function of the random variable Y on the condition of the event $\{X=x\}$ by the relation

$$\begin{aligned} F_{Y/X}(y/x) &= P(Y \le y / X = x) \\ &= \frac{P(Y \le y, X = x)}{P(X = x)} \end{aligned}$$

as $P(X=x)=0$ in the above expression. The conditional distribution function is defined in the *limiting sense* as follows:

$$\begin{aligned} F_{Y/X}(y/x) &= lim_{\Delta x \to 0} P(Y \le y / x < X \le x + \Delta x) \\ &= lim_{\Delta x \to 0} \frac{P(Y \le y, x < X \le x + \Delta x)}{P(x < X \le x + \Delta x)} \\ &= lim_{\Delta x \to 0} \frac{\int_{\infty}^{y} f_{X,Y}(x,u)\Delta x du}{f_X(x)\Delta x} \\ &= \frac{\int_{\infty}^{y} f_{X,Y}(x,u) du}{f_X(x)} \end{aligned}$$

$$\therefore F_{Y/X}(y/x) = \frac{\int_{\infty}^{y} f_{X,Y}(x,u) du}{f_X(x)}$$

Conditional Probability Density Function

$f_{Y/X}(y/X=x) = f_{Y/X}(y/x)$ is called the *conditional probability density function* of Y given X

Let us define the conditional distribution function .

The conditional density is defined in the limiting sense as follows

$$f_{Y/X}(y/X=x)=\lim_{\Delta y\to 0}(F_{Y/X}(y+\Delta y/X=x)-F_{Y/X}(y/X=x))/\Delta y$$
$$\therefore f_{Y/X}(y/X=x)=\lim_{\Delta y\to 0,\Delta x\to 0}(F_{Y/X}(y+\Delta y/x<X\le x+\Delta x)-F_{Y/X}(y/x<X\le x+\Delta x))/\Delta y$$

Because, $(X=x)=\lim_{\Delta x\to 0}(x<X\le x+\Delta x)$

The right hand side of the highlighted equation is

$$\begin{aligned}
&\lim_{\Delta y\to 0,\Delta x\to 0}(F_{Y/X}(y+\Delta y/x<X<x+\Delta x)-F_{Y/X}(y/x<X<x+\Delta x))/\Delta y\\
&\quad=\lim_{\Delta y\to 0,\Delta x\to 0}(P(y<Y\le y+\Delta y/x<X\le x+\Delta x))/\Delta y\\
&\quad=\lim_{\Delta y\to 0,\Delta x\to 0}(P(y<Y\le y+\Delta y,x<X\le x+\Delta x))/P(x<X\le x+\Delta x)\Delta y\\
&\quad=\lim_{\Delta y\to 0,\Delta x\to 0}f_{X,Y}(x,y)\Delta x\Delta y/f_X(x)\Delta x\Delta y\\
&\quad=f_{X,Y}(x,y)/f_X(x)
\end{aligned}$$

$$\therefore f_{Y/X}(y/x)=f_{X,Y}(x,y)/f_X(x)$$

Similarly we have

$$\therefore f_{X/Y}(x/y)=f_{X,Y}(x,y)/f_Y(y)$$

Two random variables are *statistically independent* if for all $(x,y)\in\mathbb{R}^2$,

$$f_{Y/X}(y/x)=f_Y(y)$$
or equivalently
$$f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

•

Example 2 X and Y are two jointly random variables with the joint pdf given by

$$\begin{aligned}f_{X,Y}(x,y)&=k \text{ for } 0\le x\le 1\\ &=0 \text{ otherwise}\end{aligned}$$

find,

(a) k
(b) $f_X(x)$ and $f_Y(y)$
(a) $f_{X/Y}(x/y)$

Solution:

Since $\int\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = 1$

We get

$$k\, x\frac{1}{2}\, x1x1 = 1$$
$$\Rightarrow k = 2$$

$$\therefore f_{X,Y}(x,y) = 2 \quad for\ 0 \le x \le 1 \text{ as } y \le x$$
$$= 0 \text{ otherwise}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = 2\int_0^x dy = 2x$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = 2\int_y^1 dx = 2(1-y)$$

Independent Random Variables (or) Statistical Independence

Let and be two random variables characterized by the joint distribution function

X Y

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

and the corresponding joint density function $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$

Then X and Y are independent if $\forall (x,y) \in \mathbb{R}^2$, $\{X \le x\}$ and $\{Y \le y\}$ are independent events. Thus,

$$F_{X,Y}(x,y) = P\{X \le x, Y \le y\}$$
$$= P\{X \le x\}P\{Y \le y\}$$
$$= F_X(x)F_Y(y)$$
$$\therefore f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$
$$= \frac{dF_X(x)}{dx}\frac{dF_Y(y)}{dy}$$
$$= f_X(x)f_Y(y)$$
$$\therefore f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

and equivalently $f_{Y/X}(y) = f_Y(y)$

OPERATIONS ON MULTIPLE RANDOM VARIABLES

Expected Values of Functions of Random Variables

Introduction:

In this Part of Unit we will see the concepts of expectation such as mean, variance, moments, characteristic function, Moment generating function on Multiple Random variables. We are already familiar with same operations on Single Random variable. This can be used as basic for our topics we are going to see on multiple random variables.

Function of joint random variables:

If g(x,y) is a function of two random variables X and Y with joint density function $f_{x,y}(x,y)$ then the expected value of the function g(x,y) is given as

$$\overline{g} = E\,[g(x,y)] \text{ or}$$

$$\overline{g} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$$

Similarly, for N Random variables $X_1, X_2, \ldots X_N$ With joint density function $f_{x1,x2,\ldots Xn}(x_1,x_2,\ldots x_n)$, the expected value of the function $g(x_1,x_2,\ldots x_n)$ is given as

$$\overline{g} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} g(x_1,x_2,\ldots x_N)\, f_{x_1,x_2,\ldots x_N}(x_1,x_2,\ldots x_N)\, dx_1 dx_2 \ldots dx_N$$

Properties :

The properties of E(X) for continuous random variables are the same as for discrete ones:

1. If X and Y are random variables on a sample space Ω
   then E(X + Y) = E(X) + E(Y). (linearity I)

2. If a and b are constants then E(aX + b) = aE(X) + b.

If $Y = g(X)$ is a function of a discrete random variable X, then $EY = Eg(X) = \sum_{x \in R_X} g(x) p_X(x)$

Suppose $Z = g(X,Y)$ is a function of continuous random variables X and Y then the expected value of Z is given by

$$\begin{aligned} EZ = Eg(X,y) &= \int_{-\infty}^{\infty} z f_Z(z) dz \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy \end{aligned}$$

Thus EZ can be computed without explicitly determining $f_Z(z)$.

We can establish the above result as follows.

Suppose $Z = g(X,Y)$ has n roots $(x_i, y_i),\ i = 1,2,..,n$ at $Z = z$. Then

$$\{z < Z \le z + \Delta z\} = \bigcup_{i=1}^{n} \{(x_i, y_i) \in \Delta D_i\}$$

Where

ΔD_i Is the differential region containing (x_i, y_i). The mapping is illustrated in Figure 1 for $n = 3$.

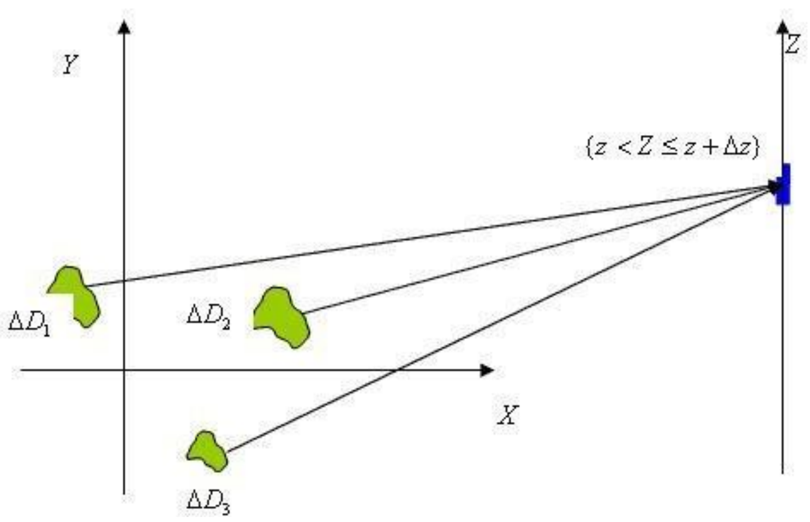


Figure 1

Note that

$$P(\{z < Z \le z + \Delta z\}) = f_Z(z)\Delta z = \sum_{(x_i, y_i) \in D_i} f_{X,Y}(x_i, y_i)\Delta x_i \Delta y_i$$

$$\begin{aligned}\therefore z f_Z(z)\Delta z &= \sum_{(x_i, y_i) \in \Delta D_i} z f_{X,Y}(x_i, y_i)\Delta x_i \Delta y_i \\ &= \sum_{(x_i, y_i) \in \Delta D_i} g(x_i, y_i) f_{X,Y}(x_i, y_i)\Delta x_i \Delta y_i\end{aligned}$$

As Z is varied over the entire Z axis, the corresponding (non-overlapping) differential regions in $X-Y$ plane cover the entire plane.

$$\therefore \int_{-\infty}^{\infty} z f_Z(z)dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dxdy$$

Thus,

$$\boxed{Eg(X,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dxdy}$$

If $Z = g(X,Y)$ is a function of discrete random variables X and Y, we can similarly show that

$$EZ = Eg(X,Y) = \sum_{x,y \in R_X \times R_Y}\sum g(x,y) p_{X,Y}(x,y)$$

Example 1 The joint pdf of two random variables X and Y is given by

$$\begin{aligned}f_{X,Y}(x,y) &= \frac{1}{4}xy \quad 0 \le x \le 2,\ 0 \le y \le 2 \\ &= 0 \quad \text{otherwise}\end{aligned}$$

Find the joint expectation of $g(X,Y) = X^2Y$

$$\begin{aligned}
Eg(X,Y) &= EX^2Y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dxdy \\
&= \int_0^2\int_0^2 x^2 y \frac{1}{4} xydxdy \\
&= \frac{1}{4}\int_0^2 x^3 dx \int_0^2 y^2 dy \\
&= \frac{1}{4} \times \frac{2^4}{4} \times \frac{2^3}{3} \\
&= \frac{8}{3}
\end{aligned}$$

Example 2 If $Z = aX + bY$, where a and b are constants, then

$$EZ = aEX + bEY$$

Proof:

$$\begin{aligned}
EZ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ax+by) f_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} axf_{X,Y}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} byf_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty} ax \int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx + \int_{-\infty}^{\infty} by \int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy \\
&= a\int_{-\infty}^{\infty} xf_X(x)dx + b\int_{-\infty}^{\infty} yf_Y(y)dy \\
&= aEX + bEY
\end{aligned}$$

Thus, expectation is a linear operator.

Example 3

Consider the discrete random variables X and Y discussed in earlier example for joint RV. The joint probability mass function of the random variables are tabulated in Table . Find the joint expectation of $g(X,Y) = XY$.

X \ Y	0	1	2	$p_Y(y)$
0	0.25	0.1	0.15	0.5
1	0.14	0.35	0.01	0.5
$p_X(x)$	0.39	0.45	0.16	

$$\text{Clearly, } EXY = \sum_{x,y \in \mathbb{R}_X \times \mathbb{R}_Y} \sum g(x,y) p_{X,Y}(x,y)$$
$$= 1 \times 1 \times 0.35 + 1 \times 2 \times 0.01$$
$$= 0.37$$

Remark

(1) We have earlier shown that expectation is a linear operator. We can generally write

$$E[a_1 g_1(X,Y) + a_2 g_2(X,Y)] = a_1 E g_1(X,Y) + a_2 E g_2(X,Y)$$

Thus $$E(XY + 5\log_e XY) = EXY + 5E\log_e XY$$

(2) If X and Y are independent random variables and $g(X,Y) = g_1(X)g_2(Y)$,then

$$Eg(X,Y) = Eg_1(X)g_2(Y)$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(X)g_2(Y) f_{X,Y}(x,y)dx$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(X)g_2(Y) f_X(x) f_Y(y)dxdy$$
$$= \int_{-\infty}^{\infty} g_1(X) f_X(x)dx \int_{-\infty}^{\infty} g_2(Y) f_Y(y)dy$$
$$= Eg_1(X)Eg_2(Y)$$