SCHOOL OF ADVANCED SCIENCES
DEPARTMENT OF MATHEMATICS
FALL SEMESTER – 2020~2021

MAT2001 – Statistics for Engineers
(Embedded Theory Component)

COURSE MATERIAL

Module 2
Random Variables

Syllabus:
Introduction – Random Variables – Probability Mass Function, Distribution and Density Functions – Joint Probability Distribution and Joint Density Functions – Marginal, Conditional Distributions and Density Functions – Mathematical Expectation and its Properties – Covariance – Moment Generating Function – Characteristic Function.

Prepared By: Prof. M. Mubashir Unnissa (In-charge)
Prof. S. Dinesh Kumar
Prof. Rajesh Bhatt

The course in-charges thankfully acknowledge the course materials preparation committee in-charge and members for their significant contribution in bringing out of this course material.

****
Dr. D. Easwaramoorthy
Dr. A. Manimaran
Course In-charges – MAT2001-SE,
Fall Semester 2020~2021,
Department of Mathematics,
SAS, VIT, Vellore.
****

MAT2001- MODULE 2 – RANDOM VARIABLES

In general, to analyze random experiments, we usually focus on some numerical aspects of the experiment. For example, in a soccer game we may be interested in the number of goals, shots, shots on goal, corners kicks, fouls, etc. If we consider an entire soccer match as a random experiment, then each of these numerical results gives some information about the outcome of the random experiment. These are examples of *random variables*. In a nutshell, a random variable is a real-valued variable whose value is determined by an underlying random experiment.

I toss a coin five times. This is a random experiment and the sample space can be written as

$$S = \{TTTTT, TTTTH, \ldots, HHHHH\}.$$

Note that here the sample space S has $2^5 = 32$ elements. Suppose that in this experiment, we are interested in the number of heads. We can define a random variable X whose value is the number of observed heads. The value of X will be one of $0, 1, 2, 3, 4$ or 5 depending on the outcome of the random experiment.

Random Variable

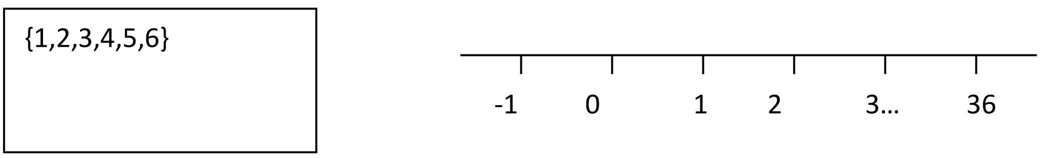
A random variable is a function that associates a real number with each element in the sample space. That is, the function $X(s) = x$ that maps the elements of the sample space S into real numbers is called the random variable associated with the concerned experiment.

Example:

Let the sample space S consisting of the numbers on die i.e. S = {1,2,3,4,5,6}.We define a R.V $X(s)=s^2$ (Square of number of points on a die).

The points in S now map onto the real line as the {1,4,9,16,25,36}.

Solution: A mapping of S = {HH, HT, TH, TT} into the real line



DISCRETE RANDOM VARIABLES

A random variable X is said to be *discrete* if it can assume only a finite or countable infinite number of distinct values. A discrete random variable can be defined on both a countable or uncountable sample space.

Probability function (or) Probability Mass Function (p.m.f):

If X is a discrete R.V. which can take the values $x_1, x_2, x_3, \ldots$ such that $P(X = x_i) = p_i$ then p_i is called the probability mass function and it satisfies the following conditions:

i) $p_i \geq 0$

ii) $\sum_i p_i = 1$

Probability distribution function

The collection of pairs $\{x_i, p_i\}$ is called the probability distribution of the R.V. X.

$X = x_i$	x_1	....	x_i
p_i	p_1	....	p_i

Example:

I have an unfair coin for which $P(H) = p$, where $0 < p < 1$. I toss the coin repeatedly until I observe a heads for the first time. Let Y be the total number of coin tosses. Find the distribution of Y.

First, we note that the random variable Y can potentially take any positive integer, so we have $R_Y = \mathbb{N} = \{1, 2, 3, \ldots\}$. To find the distribution of Y, we need to find $P_Y(k) = P(Y = k)$ for $k = 1, 2, 3, \ldots$. We have

$$P_Y(1) = P(Y = 1) = P(H) = p,$$

$$P_Y(2) = P(Y = 2) = P(TH) = (1-p)p,$$

$$P_Y(3) = P(Y = 3) = P(TTH) = (1-p)^2 p,$$

$$\begin{matrix} . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{matrix}$$

$$P_Y(k) = P(Y = k) = P(TT\ldots TH) = (1-p)^{k-1} p.$$

Thus, we can write the PMF of Y in the following way

$$P_Y(y) = \begin{cases} (1-p)^{y-1} p & \text{for } y = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

For the random variable Y in 1. Check that $\sum_{y \in R_Y} P_Y(y) = 1$.

2. If $p = \frac{1}{2}$, find $P(2 \leq Y < 5)$.

2. if $p = \frac{1}{2}$, to find P$(2 \le Y < 5)$, we can write

$$\begin{aligned} \mathrm{P}(2 \le Y < 5) &= \textstyle\sum_{k=2}^{4} P_Y(k) \\ &= \textstyle\sum_{k=2}^{4} (1-p)^{k-1} p \\ &= \frac{1}{2}\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) \\ &= \frac{7}{16}. \end{aligned}$$

CONTINUOUS RANDOM VARIABLES

A random variable X is said to be a continuous random variable if it takes all possible values between certain limits or in an interval which may be finite or infinite.

Example: operating time between two failures of a computer

Probability density function (p.d.f):

If X is a continuous R.V. such that $P(x - \frac{dx}{2} \le X \le x + \frac{dx}{2}) = f(x)dx$ then $f(x)$ is called the probability density function and it satisfies the following conditions:

i) $f(x) \ge 0$, $-\infty < x < \infty$

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Distribution Function (or) Cumulative Distribution Function (cdf)

If X is a R.V, discrete or continuous, then $F(x) = P(X \le x)$ is called the cumulative distribution function of X.

i) If X is discrete then $F(x) = \sum_j p_j$

ii) If X is continuous then $F(x) = P(-\infty \le X \le x) = \int_{-\infty}^{x} f(x)dx$

Relationship between cumulative distribution function and probability distribution function $f(x) = \frac{dF(x)}{dx}$

If $p(x)$ is the p.m.f of a random variable 'X' then $Mean = \sum_{n=0}^{\infty} xp(x)$

Variance $= \sum_{n=0}^{\infty} (x - mean)^2 p(x)$

Moment about the mean = $\sum_{n=0}^{\infty} (x - mean)^r p(x)$

If $f(x)$ is the p.d.f of a random variable 'X' which defined in the interval (a, b) then

(i) Mean = $\int_a^b xf(x)dx$ (ii) Variance = $\int_a^b (x-mean)^2 f(x)dx$

(iii) Moment about mean = $\int_a^b (x-mean)^r f(x)dx$

Problem 1:

If the random variable takes the values 1,2,3 and 4 such that $2P(X=1)=3P(X=2)=P(X=3)=5P(X=4)$, find the probability distribution and the cumulative distribution function of X.

Solution:

Let $2P(X=1)=3P(X=2)=P(X=3)=5P(X=4)=k$

That is, $P(X=1)=\frac{k}{2}$, $P(X=2)=\frac{k}{3}$, $P(X=3)=k$, $P(X=4)=\frac{k}{5}$

We know that, $\sum_i p_i = 1$

That is, $\frac{k}{2}+\frac{k}{3}+k+\frac{k}{5}=1 \Rightarrow k=\frac{30}{61}$

The probability distribution of X is given by

x_i	1	2	3	4
$P(x_i)$	$\frac{15}{61}$	$\frac{10}{61}$	$\frac{30}{61}$	$\frac{6}{61}$

The c.d.f. F(x) is defined as $F(x)=P(X\le x)$.

When $x<1$, $F(x)=0$

When $1\le x<2$, $F(x)=P(X=1)=\frac{15}{61}$

When $2\le x<3$, $F(x)=P(X=1)+P(X=2)=\frac{25}{61}$

When $3\le x<4$, $F(x)=P(X=1)+P(X=2)+P(X=3)=\frac{55}{61}$

When $x\ge 4$, $F(x)=P(X=1)+P(X=2)+P(X=3)+P(X=4)=1$

Problem 2:

A random variable X has the following probability function.

Values of X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$P(X = x)$	**a**	**3a**	**5a**	**7a**	**9a**	**11a**	**13a**	**15a**	**17a**

(i) **Find the value of a** **(ii) Find** $P(X < 3)$, $P(0 < X < 3)$, $P(X \geq 3)$

(iii) Find the distribution function of X.

Solution:

i) We know that , $\sum_i p_i = 1$

a + 3a + 5a + 7a + 9a + 11a + 13a + 15a + 17a = 1

81a = 1 $\Rightarrow a = \frac{1}{81}$

ii) $P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$ = a + 3a + 5a

= 9a = $\frac{1}{9}$

$P(0 < X < 3) = P(X = 1) + P(X = 2)$ = 3a + 5a = 8a = $\frac{8}{81}$

$P(X \geq 3) = 1 - P(X < 3) = 1 - \frac{1}{9} = \frac{8}{9}$

iii) Distribution function F(x) of X

Values of X	0	1	2	3	4	5	6	7	8
$F(x) = P(X \leq x)$	a $\frac{1}{81}$	4a $\frac{4}{81}$	9a $\frac{9}{81}$	16a $\frac{16}{81}$	25a $\frac{25}{81}$	36a $\frac{36}{81}$	49a $\frac{49}{81}$	64a $\frac{64}{81}$	81a 1

Problem 3:

Let X be a random variable with PDF given by

$$f_X(x) = \begin{cases} cx^2 & |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

a. Find the constant c.
b. Find EX and Var(X).
c. Find $P(X \geq \frac{1}{2})$.

a. To find c, we can use $\int_{-\infty}^{\infty} f_X(u)du = 1$:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(u)du \\ &= \int_{-1}^{1} cu^2 du \\ &= \frac{2}{3}c. \end{aligned}$$

Thus, we must have $c = \frac{3}{2}$.

b. To find EX, we can write

$$\begin{aligned} EX &= \int_{-1}^{1} u f_X(u)du \\ &= \frac{3}{2}\int_{-1}^{1} u^3 du \\ &= 0. \end{aligned}$$

In fact, we could have guessed $EX = 0$ because the PDF is symmetric around $x = 0$. To find Var(X), we have

$$\begin{aligned} \mathrm{Var}(X) &= EX^2 - (EX)^2 = EX^2 \\ &= \int_{-1}^{1} u^2 f_X(u)du \\ &= \frac{3}{2}\int_{-1}^{1} u^4 du \\ &= \frac{3}{5}. \end{aligned}$$

c. To find $P(X \geq \frac{1}{2})$, we can write

$$P(X \geq \frac{1}{2}) = \frac{3}{2}\int_{\frac{1}{2}}^{1} x^2 dx = \frac{7}{16}.$$

Problem 4:

Let X be a continuous random variable with PDF given by

$$f_X(x) = \frac{1}{2}e^{-|x|}, \qquad \text{for all } x \in \mathbb{R}.$$

If $Y = X^2$, find the CDF of Y.

First, we note that $R_Y = [0, \infty)$. For $y \in [0, \infty)$, we have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}e^{-|x|}dx \\ &= \int_{0}^{\sqrt{y}} e^{-x}dx \\ &= 1 - e^{-\sqrt{y}}. \end{aligned}$$

Thus,

$$F_Y(y) = \begin{cases} 1 - e^{-\sqrt{y}} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Problem 5:

Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X \leq \frac{2}{3} | X > \frac{1}{3})$.

We have

$$\begin{aligned} P(X \leq \tfrac{2}{3} | X > \tfrac{1}{3}) &= \frac{P(\frac{1}{3} < X \leq \frac{2}{3})}{P(X > \frac{1}{3})} \\ &= \frac{\int_{\frac{1}{3}}^{\frac{2}{3}} 4x^3 \, dx}{\int_{\frac{1}{3}}^{1} 4x^3 \, dx} \\ &= \frac{3}{16}. \end{aligned}$$

Problem 6

EXAMPLE: Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the value of C?

(b) Find $P\{X > 1\}$.

SOLUTION **(a)** Since f is a probability density function, we must have that $\int_{-\infty}^{\infty} f(x)\,dx = 1$, implying that

$$C \int_0^2 (4x - 2x^2)\,dx = 1$$

or

$$C\left[2x^2 - \frac{2x^3}{3}\right]\Big|_{x=0}^{x=2} = 1$$

or

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

(b) Hence

$$P\{X > 1\} = \int_1^{\infty} f(x)\,dx = \frac{3}{8} \int_1^2 (4x - 2x^2)\,dx = \frac{1}{2} \quad ■$$

Problem 7:

If a random variable 'X' has the p.d.f. $f(x) = \begin{cases} \dfrac{(x+1)}{2} & ,-1 < x < 1 \\ 0 & ,otherwise \end{cases}$ Then find the mean and variance.

Solution:

We know that, Mean = $\int_{-1}^{1} xf(x)dx$ = $\int_{-1}^{1} x\frac{(x+1)}{2}dx$ = $\frac{1}{2}\int_{-1}^{1}(x^2+x)dx$ = $\frac{1}{3}$

Variance = $\int_{-1}^{1}(x-mean)^2 f(x)dx$ = $\int_{-1}^{1}(x-\frac{1}{3})^2\frac{(x+1)}{2}dx$

= $\int_{-1}^{1}(9x^2+1-6x)(x+1)dx$ = $\frac{2}{9}$

MAT2001
Statistics for Engineers

Two Dimension Random Variables

Two Dimensional Random Variables

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, ...n, j = 1, 2, 3, ...m.

Example: 1

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. Then

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.

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) $\sum_j \sum_i p_{ij} = 1$

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

Joint Probability Distribution

The probabilities of the two events $A = \{X \le x\}\, and\, B = \{Y \le y\}$ have defined as functions of x and y respectively called probability distribution functions.

$$F_x(x) = P(X \le x)\, and\, F_y(y) = P(Y \le y)$$

Joint Probability Distribution of two random variables X and Y:

The Joint Probability Distribution of two random variables X and Y is defined as $F_{x,y}(x, y) = P\{X \le x, Y \le y\}$

Properties of the joint distribution:

A joint distribution function for the two random variables X and Y has several properties

$$1.\ F_{xy}(-\infty,-\infty)=0;\ F_{xy}(-\infty,y)=0;\ F_{xy}(x,-\infty)=0$$

$$2.\ F_{xy}(\infty,\infty)=1$$

$$3.\ 0\le F_{xy}(x,y)\le 1$$

$$4.\ F_{xy}(x,y)\ is\ a\ non-decreasing\ function\ of\ x\ and\ y\ and\ so\ on..$$

Joint probability function of Discrete R.V

For a Discrete RV,

The joint probability function of X and Y is defined as:

1. $p(x, y) \geq 0$

$$\sum_x \sum_y p(x, y) = 1.$$

The Marginal probability function is defined as

$$p_X(x) = \sum_y p(x, y) \qquad p_Y(y) = \sum_x p(x, y)$$

- And the conditional probability function is defined as

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)} \qquad p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)}$$

Independence

<u>Definition</u>: Independence

Two random variables X and Y are defined to be *independent* if

$$p(x,y) = p_X(x)\,p_Y(y) \quad \text{if } X \text{ and } Y \text{ are discrete}$$

Note: $$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)} = \frac{p_X(x)\,p_Y(y)}{p_X(x)} = p_Y(y)$$

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)} = \frac{p_X(x)\,p_Y(y)}{p_Y(y)} = p_X(x)$$

Thus, in the case of independence

marginal distributions $\equiv$ conditional distributions

Example: 2

Consider the random variables X and Y with the joint probability mass function as presented in the following table

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	

The marginal probabilities are as shown in the last column and the last row

$$p_{Y/X}(0/1) = \frac{p_{X,Y}(0,1)}{p_X(1)}$$

$$= \frac{0.14}{0.39}$$

Example: 2.A

The joint PMF of X and Y is given by

Px,y(x,y)	y = 0	y = 1	y = 2
x = 0	0.01	0	0
x = 1	0.09	0.09	0
x = 2	0	0	0.81

Find the marginal PMFs for the random variables X and Y

Solution:

Marginal PMF: $P_X(x)$ and $P_Y(y)$ by rewriting the matrix in the above Example and placing the row sums and column sums in the margins

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$	$P_X(x)$
$x = 0$	0.01	0	0	0.01
$x = 1$	0.09	0.09	0	0.18
$x = 2$	0	0	0.81	0.81
$P_Y(y)$	0.10	0.09	0.81	

Example:

For the bivariate probability distribution of (X, Y) given below, find $P(X \leq 1)$, $P(Y \leq 3)$, $P(X \leq 1, Y \leq 3)$, $P(X \leq 1/Y \leq 3)$, $P(Y \leq 3/X \leq 1)$ and $P(X + Y \leq 4)$.

X \ Y	1	2	3	4	5	6
0	0	0	1/32	2/32	2/32	3/32
1	1/16	1/16	1/8	1/8	1/8	1/8
2	1/32	1/32	1/64	1/64	0	2/64

Solution:

$$P(X \le 1) = P(X = 0) + P(X = 1)$$

$$= \sum_{j=1}^{6} P(X = 0, Y = j) + \sum_{j=1}^{6} P(X = 1, Y = j)$$

$$= \left(0 + 0 + \frac{1}{32} + \frac{2}{32} + \frac{2}{32} + \frac{3}{32}\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} +\right.$$

$$= \frac{1}{4} + \frac{5}{8} = \frac{7}{8}$$

$$P(Y \le 3) = P(Y = 1) + P(Y = 2) + P(Y = 3)$$

$$= \sum_{i=0}^{2} P(X = i, Y = 1) + \sum_{i=0}^{2} P(X = i, Y = 2)$$

$$+ \sum_{i=0}^{2} P(X = i, Y = 3)$$

$$= \left(0 + \frac{1}{16} + \frac{1}{32}\right) + \left(0 + \frac{1}{16} + \frac{1}{32}\right) + \left(\frac{1}{32} + \frac{1}{8} + \frac{1}{64}\right)$$

$$= \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$$

P $\sum^{3}$ $\sum^{3}$

$$= \left(0 + 0 + \frac{1}{32}\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8}\right) = \frac{9}{32}$$

$$P(X \le 1 / Y \le 3) = \frac{P(X \le 1, Y \le 3)}{P(Y \le 3)} = \frac{9/32}{23/64} = \frac{18}{23}$$

$$P(Y \le 3 / X \le 1) = \frac{P(X \le 1, Y \le 3)}{P(X \le 1)} = \frac{9/32}{7/8} = \frac{9}{28}$$

$$P(X + Y \le 4) = \sum_{j=1}^{4} P(X = 0, Y = j) + \sum_{j=1}^{3} P(X = 1, Y = j) + \sum_{j=1}^{2} P(X = 2, Y = j)$$

$$= \frac{3}{32} + \frac{1}{4} + \frac{1}{16} = \frac{13}{32}$$

Two – Dimensional continuous random variable

If (X,Y) can assume all values in a specified region R in XY plane (X,Y) is called a two-dimensional continuous random variable.

Joint probability function

- For a Continuous RV, the joint probability function:

$$f(x,y) = \text{Pf}[X = x, Y = y]$$

- Marginal distributions

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

- Conditional distributions

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} \qquad f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

Independence

<u>Definition</u>: Independence

Two random variables X and Y are defined to be *independent* if

$$p(x, y) = p_X(x)\, p_Y(y) \qquad \text{if } X \text{ and } Y \text{ are discrete}$$

$$f(x, y) = f_X(x)\, f_Y(y) \qquad \text{if } X \text{ and } Y \text{ are continuous}$$

Note : $$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)} = \frac{p_X(x)\, p_Y(y)}{p_X(x)} = p_Y(y)$$

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{p_X(x)\, p_Y(y)}{p_Y(y)} = p_X(x)$$

Thus, in the case of independence

marginal distributions $\equiv$ conditional distribution

The Multiplicative Rule for densities

if X and Y are discrete

$$p(x, y) = \begin{cases} p_X(x)\, p_{Y|X}(y|x) \\ p_Y(y)\, p_{X|Y}(x|y) \end{cases}$$

$$= p_X(x)\, p_Y(y) \quad \text{if } X \text{ and } Y \text{ are independent}$$

if X and Y are continuous

$$f(x, y) = \begin{cases} f_X(x)\, f_{Y|X}(y|x) \\ f_Y(y)\, f_{X|Y}(x|y) \end{cases}$$

$$= f_X(x)\, f_Y(y) \quad \text{if } X \text{ and } Y \text{ are independent}$$

Example: 3

For random variables X and Y, the joint probability density function is given by

$$f_{X,Y}(x,y)=\frac{1+xy}{4} \quad |x|\le 1,\ |y|\le 1$$

$$=0 \quad \text{otherwise}$$

Find the marginal density $f_X(x), f_Y(y)$ and $f_{Y/X}(y/x)$. Are X and Y independent?

$$f_X(x)=\int_{-1}^{1}\frac{1+xy}{4}dy$$

$$=\frac{1}{2}$$

Similarly

$$f_Y(y)=\frac{1}{2} \qquad -1\le y\le 1$$

and

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

$$=\frac{1+xy}{2} \qquad \neq f_Y(y)$$

Hence, X and Y are not independent.

Example: 4

Let X and Y have the joint density $f(x,y)=\frac{6}{7}(x+y)^2, 0\le x\le 1, 0\le y\le 1$

By integrating over the appropriate regions, find

(i)P(X>Y),(ii)P(X+Y≤1),(iii)P(X≤½).

Sol:

$$f(x,y)=\frac{6}{7}(x+y)^2, 0\le x\le 1, 0\le y\le 1$$

(i) $P(X>Y)=\int_0^1\int_0^x \frac{6}{7}(x+y)^2\,dydx=\int_0^1 \frac{2}{7}(x+y)^3\Big|_{y=0}^{y=x} dx=\frac{1}{2}$

(ii) $P(X+Y\le 1)=\int_0^1\int_0^{1-x}\frac{6}{7}(x+y)^2\,dydx$

(iii) $P(X\le \frac{1}{2})=\int_0^1\int_0^{\frac{1}{2}}\frac{6}{7}(x+y)^2\,dxdy$

Example: 6

Let X and Y have the joint densities function $f(x,y) = k(x-y)$, $0 \le y \le x \le 1$ and 0 elsewhere.

(a)Find k. (b)Find the marginal densities of X and Y.

Solution:

(a)

$$f(x,y) = k(x-y) \quad , \quad 0 \le y \le x \le 1$$

$$\int_0^1 \int_0^x k(x-y)\,dy\,dx$$

$$= \int_0^1 (kxy - \frac{1}{2}ky^2)\Big|_{y=0}^{y=x} dx$$

$$= \int_0^1 kx^2 - \frac{1}{2}kx^2\,dx$$

$$= \frac{k}{2}(\frac{x^3}{3})\Big|_0^1 = \frac{k}{6} = 1, \therefore k = 6$$

(b) *The* marginal densities of X and Y is

$$f_X(x) = \int_0^x 6(x-y)\,dy$$

$$f_Y(y) = \int_y^1 6(x-y)\,dx$$

Example:

The joint pdf of a two-dimensional RV (X, Y) is given by $f(x, y) = xy^2 + \frac{x^2}{8}$, $0 \le x \le 2, 0 \le y \le 1$.

Compute $P(X > 1)$, $P\left(Y < \frac{1}{2}\right)$, $P(X > 1/Y < 1/2)$

$P\left(Y < \frac{1}{2}/X > 1\right)$, $P(X < Y)$ and $P(X + Y \le 1)$.

Solution:

Here the rectangle defined by $0 \le x \le 2, 0 \le y \le 1$ is the range space R. $R_1, R_2, \ldots$,are event spaces.

(i) $P(X > 1) = \int\int_{\substack{R_1 \\ (x>1)}} f(x, y)\, dx\, dy$

$$= \int_0^1 \int_1^2 \left(xy^2 + \frac{x^2}{8} \right) dxdy = \frac{19}{24}$$

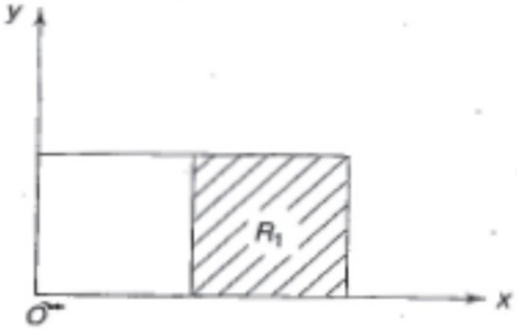


Fig. 2.1

(ii) $P(Y < 1/2) = \int_{\substack{R_2 \\ \left(y<\frac{1}{2}\right)}} \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$

$$= \int_0^{1/2} \int_0^2 \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

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

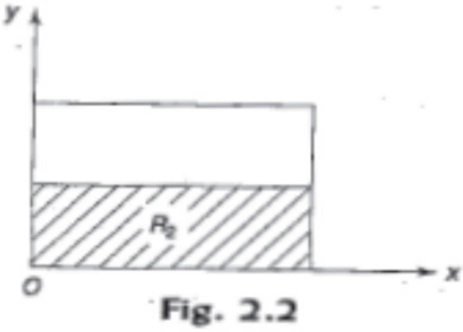


Fig. 2.2

(iii) $P(X > 1, Y < 1/2) = \underset{\substack{R_3 \\ \left(x>1 \,\&\, y<\frac{1}{2}\right)}}{\int\int} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^{1/2}\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$$

$$= \frac{5}{24}$$

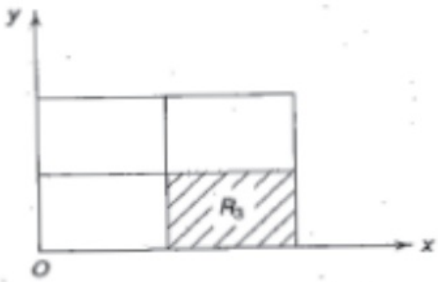


Fig. 2.3

(iv) $P(X > 1/Y < \frac{1}{2}) = \dfrac{P\left(X > 1, Y < \frac{1}{2}\right)}{P\left(Y < \frac{1}{2}\right)} = \dfrac{5/24}{1/4} = \dfrac{5}{6}$

(v) $P(Y < \frac{1}{2}/X > 1) = \dfrac{P\left(X > 1, Y < \frac{1}{2}\right)}{P(X > 1)} = \dfrac{5/24}{19/24} = \dfrac{5}{19}$

(vi) $P(X < Y) = \underset{\substack{R_4 \\ (x<y)}}{\int\int} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^1\int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx\, dy = \frac{53}{480}$$

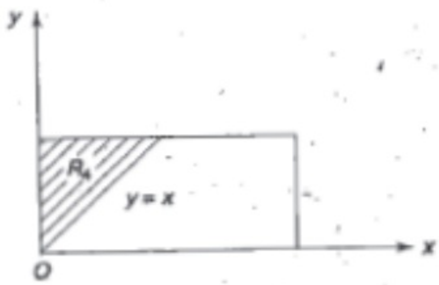


Fig. 2.4

(vii) $P(X + Y \le 1) = \underset{\substack{R_5 \\ (x+y\le 1)}}{\int\int} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy = \frac{13}{480}$$

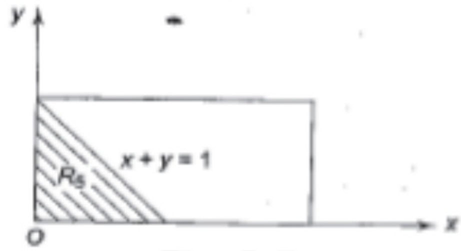


Fig. 2.5

Exercise:

Given $f_{XY}(x, y) = cx\,(x - y)$, $0 < x < 2$, $-x <, y < x$, and 0 elsewhere, (a) evaluate c, (b) find $f_X(x)$, (c) $f_{Y/X}\,(y/x)$ and (d) $f_Y(y)$.

- ***Theorem:***

Statement:

Given two random variables X and Y, the following equality is true:

$$E[X+Y] = E[X] + E[Y].$$

Proof:

Regarding $X + Y$ as a function of two random variables $g(X, Y)$, we can apply Proposition 7.1 to get

$$\begin{aligned} E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)f(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dydx \;+\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy \\ &= \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx \;+\; \int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f(x,y)dx\right)dy \\ &= \int_{-\infty}^{\infty} xf_X(x)dx \;+\; \int_{-\infty}^{\infty} yf_Y(y)dy \\ &= E[X] + E[Y]. \end{aligned}$$

Covariance between X & Y

- Covariance = 0 for independent X, Y
  - Positive for large X with large Y
  - Negative for large X with small Y (vice versa)
  - Formula is similar to our familiar variance formula

$$Cov[X,Y] = E(XY) - E(X) \bullet E(Y)$$

$$\rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$Var(aX + bY + c) = a^2 Var[X] + b^2 Var[Y] + 2abCov[X,Y]$$

Cumulative distribution function:

If (X,Y) is a two dimensional R.V (discrete or continuous), then

$F(x,y) = P(X \leq x \; and \; Y \leq y)$ is called the cumulative distribution function of (X,Y).

i) In discrete case, $F(x,y) = \sum_{j} \sum_{i} p_{ij}$

ii) In continuous case, $F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x,y)dxdy$

EXAMPLE: 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 and Y denote, respectively, the number of new and used but still working batteries that are chosen, then the joint probability mass function of X and Y, $p(i, j) = P\{X = i, Y = j\}$, is given by

$$p(0,0) = \binom{5}{3} \bigg/ \binom{12}{3} = 10/220$$

$$p(0,1) = \binom{4}{1}\binom{5}{2} \bigg/ \binom{12}{3} = 40/220$$

$$p(0,2) = \binom{4}{2}\binom{5}{1} \bigg/ \binom{12}{3} = 30/220$$

$$p(0,3) = \binom{4}{3} \bigg/ \binom{12}{3} = 4/220$$

$$p(1,0) = \binom{3}{1}\binom{5}{2} \bigg/ \binom{12}{3} = 30/220$$

$$p(1,1) = \binom{3}{1}\binom{4}{1}\binom{5}{1} \bigg/ \binom{12}{3} = 60/220$$

$$p(1,2) = \binom{3}{1}\binom{4}{2} \bigg/ \binom{12}{3} = 18/220$$

$$p(2,0) = \binom{3}{2}\binom{5}{1}\Big/\binom{12}{3} = 15/220$$

$$p(2,1) = \binom{3}{2}\binom{4}{1}\Big/\binom{12}{3} = 12/220$$

$$p(3,0) = \binom{3}{3}\Big/\binom{12}{3} = 1/220$$

These probabilities can most easily be expressed in tabular form as shown in Table below

TABLE $P\{X = i, Y = j\}$

$i \backslash j$	0	1	2	3	Row Sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column Sums = $P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

XAMPLE: The joint density function of X and Y is given by

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

Compute **(a)** $P\{X > 1, Y < 1\}$; **(b)** $P\{X < Y\}$; and **(c)** $P\{X < a\}$.

SOLUTION

(a)

$$\begin{aligned} P\{X > 1, Y < 1\} &= \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dx\,dy \\ &= \int_0^1 2e^{-2y}(-e^{-x}|_1^\infty)\,dy \\ &= e^{-1}\int_0^1 2e^{-2y}\,dy \\ &= e^{-1}(1 - e^{-2}) \end{aligned}$$

(b)

$$\begin{aligned} P\{X < Y\} &= \iint\limits_{(x,y):x<y} 2e^{-x}e^{-2y}\,dx\,dy \\ &= \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dx\,dy \\ &= \int_0^\infty 2e^{-2y}(1 - e^{-y})\,dy \\ &= \int_0^\infty 2e^{-2y}\,dy - \int_0^\infty 2e^{-3y}\,dy \\ &= 1 - \tfrac{2}{3} \\ &= \tfrac{1}{3} \end{aligned}$$

(c)

$$\begin{aligned} P\{X < a\} &= \int_0^a \int_0^\infty 2e^{-2y}e^{-x}\,dy\,dx \\ &= \int_0^a e^{-x}\,dx \\ &= 1 - e^{-a} \quad \blacksquare \end{aligned}$$

Mathematical Expectations:

Let X be a random variable with probability density function $f(x)$ or probability mass function $p(x)$ .Then the mathematical expectation of 'X' is denoted by E(X)

Mean (or) expectation is a significant number representing the behavior of a random variable.

If X is a discrete random variable then

$$E[X] = \sum_{i=0}^{\infty} xp(x)$$

$$E[X^2] = \sum_{i=0}^{\infty} x^2 p(x)$$

$$E[XY] = \sum\sum xyp(x,y)$$

If X is a continuous random variable then

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx$$

$$E[xy] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$$

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

Note : $Mean = E[X]; Mean\ square = E[X^2]$

1 Conditional Distribution and Expected Values of RVs

To begin with, let us consider discrete random vector $[X,\ Y]$ with joint pmf $p_{X,Y}$ and the marginals pmfs p_X and p_Y. Conditional pmf of X given Y is

$$p_{X|Y=y}(x) = \mathrm{P}(X = x|Y = y) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

provided that $p_Y(y) > 0$. $p_{Y|X=x}(y)$ is similarly defined. Once the conditional pmf is available for $p_Y(y) > 0$, the conditional expectation is

$$\mathrm{E}(X|Y = y) = \sum_x x p_{X|Y=y}(x).$$

$$\boldsymbol{Cov(X,Y) = E[XY] - E[X]E[Y]}$$

$$\boldsymbol{r(X,Y) = \frac{Cov(X,Y)}{\sqrt{var(x)}\sqrt{var(Y)}}}$$

Note: If X and Y are independent random variables then $E[XY] = E[X]E[Y]$

Conditional Probability Function of X and Y (Discrete)

i) The conditional probability function of X given $Y = y_j$ is given by

$$P\{X = x_i \big/ Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{\bullet j}}$$

ii) The conditional probability function of Y given $X = x_i$ is given by

$$P\{Y = y_j \big/ X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i\bullet}}$$

Note:

The two R.V's of X and Y are said to be independent if $P\{X = x_i / Y = y_j\} = P(X = x_i)P(Y = y_j)$

That is, $p_{ij} = p_{i\bullet} \times p_{\bullet j}$

Conditional density function of X and Y : (Continuos)

i) The conditional density of X given Y is given by

$$f(x \big/ y) = \frac{f(x, y)}{f_Y(y)}$$

ii) The conditional probability function of Y given X is given by

$$f(y \big/ x) = \frac{f(x, y)}{f_X(x)}$$

Note: If X and Y are independent the $\boldsymbol{f(x, y)} =$ $f_X(y)$. $f_Y(y)$

Properties of Expectation :

1. $\boldsymbol{E[X + Y] = E[X] + E[Y]}$
2. $\boldsymbol{E[XY] = E[X]E[Y]\ X\ and\ Y\ are\ independent}$
3. $\boldsymbol{E[a\} = a\ where\ a\ is\ constant}$
4. $\boldsymbol{E[aX + b] = aE[X] + b}$ **where** $\boldsymbol{a\ and\ b}$ **are constants**
5. $\boldsymbol{E[a + f(x)] = a + E[f(x)]}$
6. $\boldsymbol{E[\sum a_i X_i] = \sum a_i E[X_i]}$ **i=1,2,3…n**

Properties of variance:

$$\mathrm{Var}[aX+b] = a^2 \mathrm{Var}[X]$$

$$Var(a) = 0 \text{ where } a \text{ is constant}$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y).$$

The Variance of a Sum of RVs

- For any random variables $X_1, \ldots, X_n$

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

If X & Y are *independent*, then

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X]+\mathrm{Var}[Y]$$

Properties of covariance

- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- Symmetry: $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- Linearity in each variable:

$$\mathrm{Cov}(aX+b,Y) = a\,\mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$$

$$\mathrm{Cov}(a_1X+b_1, a_2Y+b_2) = a_1a_2\mathrm{Cov}(X,Y)$$

Prob 1:

Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} x^2\left(2x + \frac{3}{2}\right) & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

If $Y = \frac{2}{X} + 3$, find Var(Y).

$$\text{Var}(Y) = \text{Var}\left(\frac{2}{X} + 3\right) = 4\text{Var}\left(\frac{1}{X}\right),$$

$$E\left[\frac{1}{X}\right] = \int_0^1 x\left(2x + \frac{3}{2}\right) dx = \frac{17}{12}$$

$$E\left[\frac{1}{X^2}\right] = \int_0^1 \left(2x + \frac{3}{2}\right) dx = \frac{5}{2}.$$

Thus, Var$\left(\frac{1}{X}\right) = E[\frac{1}{X^2}] - (E[\frac{1}{X}])^2 = \frac{71}{144}$. So, we obtain

$$\text{Var}(Y) = 4\text{Var}\left(\frac{1}{X}\right) = \frac{71}{36}.$$

Problem 2:

I toss a coin twice and define X to be the number of heads I observe. Then, I toss the coin two more times and define Y to be the number of heads that I observe this time. Find $P\Big((X < 2) \text{ and } (Y > 1)\Big)$.

Solution

Since X and Y are the result of different independent coin tosses, the two random variables X and Y are independent. Also, note that both random variables have the distribution we found in Example 3.3. We can write

$$\begin{aligned} P\Big((X < 2) \text{ and } (Y > 1)\Big) &= P(X < 2)P(Y > 1) \qquad \text{(because } X \text{ and } Y \text{ are independent)} \\ &= \big(P_X(0) + P_X(1)\big)P_Y(2) \\ &= \left(\tfrac{1}{4} + \tfrac{1}{2}\right)\tfrac{1}{4} \\ &= \tfrac{3}{16}. \end{aligned}$$

Problem 3:

Suppose that X and Y are independent random variables having density function $f(x, y) = e^{-(x+y)}, x > 0, y > 0$ find the density function of the random variable X/Y

SOLUTION We start by determining the distribution function of X/Y. For $a > 0$

$$F_{X/Y}(a) = P\{X/Y \leq a\}$$
$$= \iint\limits_{x/y \leq a} f(x, y)\, dx\, dy$$
$$= \iint\limits_{x/y \leq a} e^{-x} e^{-y}\, dx\, dy$$
$$= \int_0^{\infty} \int_0^{ay} e^{-x} e^{-y}\, dx\, dy$$

$$= \int_0^{\infty} (1 - e^{-ay}) e^{-y}\, dy$$
$$= \left[-e^{-y} + \frac{e^{-(a+1)y}}{a+1} \right] \Big|_0^{\infty}$$
$$= 1 - \frac{1}{a+1}$$

Differentiation yields that the density function of X/Y is given by

$$f_{X/Y}(a) = 1/(a+1)^2, \qquad 0 < a < \infty \quad \blacksquare$$

Problem 4

EXAMPLE: : 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

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

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

$$P\{X_1 = 1, X_2 = 2, X_3 = 0\} = (.20)(.10)(.30) = .006$$

where we have let X_i denote the change on the ith day. ■

Problem 5

EXAMPLE: Suppose that $p(x, y)$, the joint probability mass function of X and Y, is given by

$$p(0,0) = .4, \quad p(0,1) = .2, \quad p(1,0) = .1, \quad p(1,1) = .3.$$

Calculate the conditional probability mass function of X given that $Y = 1$.

SOLUTION We first note that

$$P\{Y = 1\} = \sum_x p(x, 1) = p(0, 1) + p(1, 1) = .5$$

Hence,

$$P\{X = 0|Y = 1\} = \frac{p(0,1)}{P\{Y = 1\}} = 2/5$$

$$P\{X = 1|Y = 1\} = \frac{p(1,1)}{P\{Y = 1\}} = 3/5 \quad ■$$

Problem 6:

EXAMPLE: Find $E[X]$ where X is the outcome when we roll a fair die.

SOLUTION Since $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$, we obtain that

$$E[X] = 1\left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) + 3\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{6}\right) + 5\left(\tfrac{1}{6}\right) + 6\left(\tfrac{1}{6}\right) = \tfrac{7}{2}$$

EXAMPLE: If I is an indicator random variable for the event A, that is, if

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

then

$$E[I] = 1P(A) + 0P(A^c) = P(A)$$

Hence, the expectation of the indicator random variable for the event A is just the probability that A occurs. ■

Problem 7

EXAMPLE: The joint density of X and Y is given by

$$f(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

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

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

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} \\ &= \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dx} \\ &= \frac{x(2-x-y)}{\int_0^1 x(2-x-y)\,dx} \\ &= \frac{x(2-x-y)}{\frac{2}{3} - y/2} \\ &= \frac{6x(2-x-y)}{4-3y} \end{aligned}$$ ■

Problem 8

EXAMPLE: Suppose that you are expecting a message at some time past 5 P.M. From experience you know that X, the number of hours after 5 P.M. until the message arrives, is a random variable with the following probability density function:

$$f(x) = \begin{cases} \dfrac{1}{1.5} & \text{if } 0 < x < 1.5 \\ 0 & \text{otherwise} \end{cases}$$

The expected amount of time past 5 P.M. until the message arrives is given by

$$E[X] = \int_0^{1.5} \frac{x}{1.5}\,dx = .75$$

Hence, on average, you would have to wait three-fourths of an hour. ■

Problem 9

EXAMPLE: Suppose X has the following probability mass function

$$p(0) = .2, \quad p(1) = .5, \quad p(2) = .3$$

Calculate $E[X^2]$.

SOLUTION Letting $Y = X^2$, we have that Y is a random variable that can take on one of the values $0^2, 1^2, 2^2$ with respective probabilities

$$\begin{aligned} p_Y(0) &= P\{Y = 0^2\} = .2 \\ p_Y(1) &= P\{Y = 1^2\} = .5 \\ p_Y(4) &= P\{Y = 2^2\} = .3 \end{aligned}$$

Hence,

$$E[X^2] = E[Y] = 0(.2) + 1(.5) + 4(.3) = 1.7 \quad ■$$

Problem 10

EXAMPLE: The time, in hours, it takes to locate and repair an e
a certain factory is a random variable — call it X — whose density

$$f_X(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

If the cost involved in a breakdown of duration x is x^3, what is the
a breakdown?

SOLUTION Letting $Y = X^3$ denote the cost, we first calculate its di
follows. For $0 \leq a \leq 1$,

$$\begin{aligned} F_Y(a) &= P\{Y \leq a\} \\ &= P\{X^3 \leq a\} \\ &= P\{X \leq a^{1/3}\} \\ &= \int_0^{a^{1/3}} dx \\ &= a^{1/3} \end{aligned}$$

By differentiating $F_Y(a)$, we obtain the density of Y,

$$f_Y(a) = \frac{1}{3}a^{-2/3}, \quad 0 \leq a < 1$$

Hence,

$$\begin{aligned} E[X^3] = E[Y] &= \int_{-\infty}^{\infty} a f_Y(a)\, da \\ &= \int_0^1 a \frac{1}{3} a^{-2/3}\, da \\ &= \frac{1}{3}\int_0^1 a^{1/3}\, da \\ &= \frac{1}{3}\frac{3}{4} a^{4/3}|_0^1 \\ &= \frac{1}{4} \quad \blacksquare \end{aligned}$$

Problem 11

EXAMPLE: A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. If its probabilities of winning the jobs are respectively .2, .8, and .3, what is the firm's expected total profit?

SOLUTION Letting $X_i, i = 1, 2, 3$ denote the firm's profit from job i, then

$$\text{total profit} = X_1 + X_2 + X_3$$

and so

$$E[\text{total profit}] = E[X_1] + E[X_2] + E[X_3]$$

Now

$$E[X_1] = 10(.2) + 0(.8) = 2$$
$$E[X_2] = 20(.8) + 0(.2) = 16$$
$$E[X_3] = 40(.3) + 0(.7) = 12$$

and thus the firm's expected total profit is 30 thousand dollars. ■

Problem 12

EXAMPLE: A secretary has typed N letters along with their respective envelopes. The envelopes get mixed up when they fall on the floor. If the letters are placed in the mixed-up envelopes in a completely random manner (that is, each letter is equally likely to end up in any of the envelopes), what is the expected number of letters that are placed in the correct envelopes?

SOLUTION Letting X denote the number of letters that are placed in the correct envelope, we can most easily compute $E[X]$ by noting that

$$X = X_1 + X_2 + \cdots + X_N$$

where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th letter is placed in its proper envelope} \\ 0 & \text{otherwise} \end{cases}$$

Now, since the ith letter is equally likely to be put in any of the N envelopes, it follows that

$$P\{X_i = 1\} = P\{i\text{th letter is in its proper envelope}\} = 1/N$$

and so

$$E[X_i] = 1P\{X_i = 1\} + 0P\{X_i = 0\} = 1/N$$

Hence, from Equation 4 $E[X_1 + X_2 \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$

$$E[X] = E[X_1] + \cdots + E[X_N] = \left(\frac{1}{N}\right) N = 1$$

Hence, no matter how many letters there are, on the average, exactly one of the letters will be in its own envelope. ■

Problem 13

EXAMPLE: Suppose there are 20 different types of coupons and suppose that each time one obtains a coupon it is equally likely to be any one of the types. Compute the expected number of different types that are contained in a set for 10 coupons.

SOLUTION Let X denote the number of different types in the set of 10 coupons. We compute $E[X]$ by using the representation

$$X = X_1 + \cdots + X_{20}$$

where

$$X_i = \begin{cases} 1 & \text{if at least one type } i \text{ coupon is contained in the set of 10} \\ 0 & \text{otherwise} \end{cases}$$

Now

$$\begin{aligned} E[X_i] &= P\{X_i = 1\} \\ &= P\{\text{at least one type } i \text{ coupon is in the set of 10}\} \\ &= 1 - P\{\text{no type } i \text{ coupons are contained in the set of 10}\} \\ &= 1 - \left(\tfrac{19}{20}\right)^{10} \end{aligned}$$

when the last equality follows since each of the 10 coupons will (independently) not be a type i with probability $\frac{19}{20}$. Hence,

$$E[X] = E[X_1] + \cdots + E[X_{20}] = 20\left[1 - \left(\tfrac{19}{20}\right)^{10}\right] = 8.025 \quad \blacksquare$$

Problem 15

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 (i) E(X) (ii) E(Y) (iii) E(XY) (iv) E(2X + 3Y) (v) Var(X)

(vi) Cov(X,Y).

Solution:

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

$= \int_1^5\int_0^4 x\left(\dfrac{xy}{96}\right)dxdy$

$= \dfrac{8}{3}$

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

$= \int_1^5\int_0^4 y\left(\dfrac{xy}{96}\right)dxdy$

$= \dfrac{31}{9}$

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

$= \int_1^5\int_0^4 xy\left(\dfrac{xy}{96}\right)dxdy$

$= \dfrac{248}{27}$

iv) $E[2X + 3Y] = 2E(X) + 3E(Y)$ $= 2.\dfrac{8}{3} + 3.\dfrac{31}{9} = \dfrac{47}{3}$

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

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

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

= 8

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

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

vi) $Cov(X,Y) = E(XY) - E(X).E(Y)$

= $\frac{248}{27}$ - $\left(\frac{8}{3}\right)\left(\frac{31}{9}\right)$

= 0

Moment generating function (M.G.F.)

The m.g.f of a random variable X(about origin) whose probability function $f(x)$ is given by

$$M_X(t) = E(e^{tX})$$

$$= \begin{cases} \int_{-\infty}^{\infty} e^{tx} f(x)dx, \text{for a continuous probability distribution} \end{cases}$$

$$= \begin{cases} \sum_{x}^{\infty} e^{tx} p(x), \text{for a discrete probability distribution} \end{cases}$$

Where t is real parameter and the integration or summation being extended to the entire range of x.

To find the rth moment of X about origin,

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r \qquad [\text{ using } \mu'_r = E(X^r))$$

This gives the m.g.f interms of moments. Thus the coefficient of $\frac{t^r}{r!} in\, M_X(t)$ gives the rth moment of the r.v.'X' about origin(μ'_r).

Since $M_X(t)$ generates moments, it is known as moment generating function.

Characteristic Function

In few cases the m.g.f. does not exist. In this case we use characteristic function which is more serviceable function than m.g.f.

The characteristic function of a random variable X whose probability function $f(x)$ is given by

$$\Phi_X(t) = E(e^{itX})$$

$$= \begin{cases} \int_{-\infty}^{\infty} e^{itx} f(x)dx, \text{for a continuous probability distribution} \end{cases}$$

$$= \begin{cases} \sum_{x}^{\infty} e^{itx} p(x), \text{for a discrete probability distribution} \end{cases}$$

Properties of Characteristic Functions:

Property:1

$$\Phi_X(0) = 1$$

2. $$|\Phi_X(t)| \le 1 = \Phi_X(0)$$

3. $\Phi_X(-t) = \overline{\Phi_X(t)}$ where $\overline{\Phi_X(t)}$ is the complex conjugate of $\Phi_X(t)$.

Problems:

1. Find the m.g.f of the R.V. with the probability law P(X=x)= $q^{x-1}p$, x=1,2,3…..Find the Mean and Variance.

Solution:

We know that,

$$M_X(t) = E(e^{tX})$$

$$= \sum_x^{\infty} e^{tx} P(x) \quad \text{[by definition}$$

$$= \sum_{x=1}^{\infty} e^{tx} q^{x-1}.p$$

$$= \sum_{x=1}^{\infty} e^{tx} q^x q^{-1}.p$$

$$= q^{-1}.p \sum_{x=1}^{\infty} (e^t q)^x.$$

$$= \frac{p}{q}.qe^t \sum_{x=1}^{\infty} (e^t q)^{x-1}$$

$$= pe^t[1 + qe^t + (qe^t)^2 + \ldots]$$

$$= pe^t(1 - qe^t)^{-1}$$

$$= \frac{pe^t}{(1 - qe^t)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Mean:

Differentiating (1) w.r.t. 't', we get

$$\frac{d}{dx}\{M_X(t)\} = M'_X(t) = \frac{pe^t}{(1-qe^t)^2}$$

$$\mu_1'(about\,\text{origin}) = M'_X(0) = \frac{p}{(1-q\ \ )^2} = \frac{1}{p}$$

Variance:

$$= \mu_2' - \mu_1'^2$$

$$\frac{d^2}{dx^2}\{M_X(t)\} = M''_X(t) = \frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

$$\mu_2'(about\,\text{origin}) = M''_X(0) = \frac{p(1+q)}{(1-q\ \ )^{32}}$$

$$= \frac{q}{p^2}$$

2. Find the m.g.f. of the R.V. whose moments are $\mu_r' = (r+1)!\,2^r$.

Solution:

We know that the m.g.f interms of moments is given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu'_r$$

$$= \sum_{r=0}^{\infty} \frac{t^r}{r!}(r+1)!2^r$$

$$= \sum_{r=0}^{\infty} \frac{(2t)^r(r+1)!}{r!}$$

$$= 1 + 2(2t) + 3(2t)^2 + \ldots.$$

$$= (1-2t)^{-2} = \frac{1}{(1-2t)^2}$$

Problem:

EXAMPLE: Suppose that you have a fair 4-sided die, and let X be the random variable representing the value of the number rolled.

(a) Write down the moment generating function for X.
(b) Use this moment generating function to compute the first and second moments of X.

Solution:

(a)

$$\begin{aligned} m_X(t) &= \mathbb{E}\left[e^{tX}\right] = e^{1\cdot t}\frac{1}{4} + e^{2\cdot t}\frac{1}{4} + e^{3\cdot t}\frac{1}{4} + e^{4\cdot t}\frac{1}{4} \\ &= \frac{1}{4}\left(e^{1\cdot t} + e^{2\cdot t} + e^{3\cdot t} + e^{4\cdot t}\right) \end{aligned}$$

(b)

$$\begin{aligned} m_X'(t) &= \frac{1}{4}\left(e^{1\cdot t} + 2e^{2\cdot t} + 3e^{3\cdot t} + 4e^{4\cdot t}\right), \\ m_X''(t) &= \frac{1}{4}\left(e^{1\cdot t} + 4e^{2\cdot t} + 9e^{3\cdot t} + 16e^{4\cdot t}\right), \end{aligned}$$

$$\mathbb{E}X = m_X'(0) = \frac{1}{4}(1+2+3+4) = \frac{5}{2}$$

$$\mathbb{E}X^2 = m_X''(0) = \frac{1}{4}(1+4+9+16) = \frac{15}{2}.$$

EXAMPLE: Let X be a random variable whose probability density function is given by

$$f_X(x) = \begin{cases} e^{-2x} + \frac{1}{2}e^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}.$$

(a) Write down the moment generating function for X.

(b) Use this moment generating function to compute the first and second moments of X.

(a)

for $t < 1$ we have

$$\begin{aligned} m_X(t) &= \mathbb{E}\left[e^{tX}\right] = \int_0^\infty e^{tx}\left(e^{-2x} + \frac{1}{2}e^{-x}\right)dx \\ &= \frac{1}{t-2}e^{tx-2x} + \frac{1}{2(t-1)}e^{tx-x}\Big|_{x=0}^{x=\infty} = \\ &= 0 - \frac{1}{2-t} + 0 - \frac{1}{2(t-1)} \\ &= \frac{1}{t-2} + \frac{1}{2(1-t)} = \frac{t}{2(2-t)(1-t)} \end{aligned}$$

(b)

$$\begin{aligned} m'_X(t) &= \frac{1}{(2-t)^2} + \frac{1}{2(1-t)^2} \\ m''_X(t) &= \frac{2}{(2-t)^3} + \frac{1}{(1-t)^3} \end{aligned}$$

and so $\mathbb{E}X = m'_X(0) = \frac{3}{4}$ and $\mathbb{E}X^2 = m''_X = \frac{5}{4}$.

EXAMPLE: Let X and Y be two independent random variables with respective moment generating functions

$$m_X(t) = \frac{1}{1-5t}, \quad \text{if } t < \frac{1}{5}, \quad m_Y(t) = \frac{1}{(1-5t)^2}, \quad \text{if } t < \frac{1}{5}.$$

Find $\mathbb{E}(X+Y)^2$.

Solution let $W = X + Y$, and using that X, Y

are independent, then we see that

$$m_W(t) = m_{X+Y}(t) = m_X(t) m_Y(t) = \frac{1}{(1-5t)^3},$$

recall that $\mathbb{E}[W^2] = m''_W(0)$, which we can find from

$$\begin{aligned} m'_W(t) &= \frac{15}{(1-5t)^4}, \\ m''_W(t) &= \frac{300}{(1-5t)^5}, \end{aligned}$$

thus

$$\mathbb{E}\left[W^2\right] = m''_W(0) = \frac{300}{(1-0)^5} = 300.$$