CHAPTER 9

MATHEMATICAL EXPECTATION

9. 1 MEANING OF MATHEMATICAL EXPECTATION

Historically, the term **mathematical expectation** or **expected value** is derived from games of chances. In such games, the gamblers were concerned with how much, on the average, one would expect to win if the game is continued for a sufficiently long time. In statistical terminology, this term is associated with a random variable and in fact, is the average value of this random variable generated through a random experiment.

The computation of expected value of a random variable is straightforward. When the variable is discrete, it is simply the sum of the products of all possible values of the random variables multiplied by their respective probabilities. For continuous variable, it is analogously defined.

***Definition* 9.1**: *If X is a discrete random variable with the probability function f(x), then the expected value or the mathematical expectation of X, E(X), is defined as*

$$E(X) = \sum_{x} x\, f(x)$$

If X is continuous having a density function f(x), then

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

As we noted earlier, the probability distribution for a random variable is a theoretical model for the empirical distribution of the data associated with a real population. If the model is an accurate representation of nature, the theoretical and empirical distributions are equivalent. Consequently, all numerical descriptive measures of the probability distribution $f(x)$ will be consistent with those discussed in Chapter 3 and Chapter 4.

Definition 9.1 is completely consistent with the definition of arithmetic mean of a set of measurements that was given in chapter 3. Thus if $f(x)$ is an accurate characterization of the population frequency distribution, then $E(X) = \mu$, the population mean.

***Theorem* 9.1**: *Let X be a discrete random variable with probability function $f(x)$ and c be a constant. Then $E(c) = c$.*

Proof: By definition

$$E(c) = \sum_x cf(x) = c\sum_x f(x)$$

But $\sum f(x) = 1$ and hence

$$E(c) = c(1) = c$$

We are frequently interested in the mean or expected value of a function of a random variable X. The expressions of the forms X^2, $\sqrt{X}$, $2X+1$ etc, are all functions of the random variable X, and clearly they are random variables, which we denote by $w(X)$. Logically, we can speak of the expected values of these functions. We now define the expected value of $w(X)$ as follows:

***Definition* 9.2:** *Let X be a random variable with probability distribution $f(x)$. The expected value of the function $w(X)$ of the random variable X is*

$$E[w(X)] = \sum_x w(x)f(x), \qquad \text{if } X \text{ is discrete}$$

$$= \int_{-\infty}^{\infty} w(x)f(x)dx, \qquad \text{if } X \text{ is continuous}$$

We now seek the variance and standard deviation of this distribution. Which we define below in terms of the expected value

***Definition* 9.3:** *The variance of a random variable X is defined to be the expected value of the expected value of the differences between the values of X and their mean μ. That is*

$$V(X) = E(X-\mu)^2.$$

= E(X)^2 - (E(X))^2
= E(X^2) - (E(X))^2

CamScanner

Once again, if $f(x)$ is an accurate characterization of the population frequency distribution, then $V(X)=\sigma^2$, the population variance. The positive square root of this variance is the population standard deviation, i.e. $\sigma=\sqrt{E(X-M)^2}$

Example 9.1: A discrete random variable X has a probability function as shown in the following table:

x^2 : 9 4 0 1 4

Values of X: x :	-3	-2	0	1	2
$P(X=x)=f(x)$:	.10	.30	.15	.40	.05

Find $E(X)$ and $V(X)$.

or, $V(x)=E(x)^2-(E(x))^2$
$=9\cdot.10+4\cdot.30+0+1\cdot.40+4\cdot.05-(.4)^2$
$=2.54$

Solution: By definition

$$\mu=E(X)=\sum xf(x)=(-3)(.1)+(-2)\times(.3)+(0)\times(.15)$$
$$+(1)\times(.4)+(2)\times(.05)=-0.4$$

$$V(X)=E(X-\mu)^2=\sum_{x=-3}^{2}(X-\mu)^2 f(x)$$
$$=(-3+.4)^2(.1)+(-2+.4)^2(.3)+(0+.4)^2(.15)$$
$$+(1+.4)^2(.4)+(2+.4)^2(.05)=2.54$$

***Theorem* 9.2:** *Let X be a discrete random variable with probability mass function $f(x)$, $w(X)$ be a function of X, and let c be a constant. Then*

$$E[cw(X)]=cE[w(X)]$$

Proof: By definition 9.2,

$$E[cw(X)]=\sum_x cw(x)f(x)=c\sum_x w(x)f(x)=cE[w(X)]$$

In particular, if c=5 and $w(X)$, then $E(5X^2)=5E(X^2)$

***Theorem* 9.3**: *The expected value of a sum of several functions of a random variable X is equal to the sum of their respective expected values. That is if X be a random variable with probability mass function $f(x)$ and $w_1(X)$, $w_2(X)$, ..., $w_k(X)$ be k functions of X, then*

$$E[w_1(X)+w_2(X)+...+w_k(X)]=E[w_1(X)]+E[w_2(X)]+...+E[w_k(X)]$$

Proof: We will proof the theorem for k=2, but it can be extended to any finite k. By definition

$$E[w_1(X)+w_2(X)]=\sum_x [w_1(x)+w_2(x)]f(x)$$
$$=\sum_x w_1(x)f(x)+\sum_x w_2(x)f(x)$$
$$=E[w_1(X)]+E[w_2(X)]$$

This proves the theorem.

***Theorem* 9.4:** *Let X be a discrete random variable with probability mass function f(x); then*

$$V(X)=\sigma^2=E(X-\mu)^2=E(X^2)-\mu^2$$

Proof: Following definition 9.3

$$\sigma^2=E(X-\mu)^2=E(X^2-2X\mu+\mu^2)$$
$$=E(X^2)-2\mu E(X)+\mu^2,\quad \text{since } \mu \text{ is constant}$$
$$=E(X^2)-2\mu^2+\mu^2=E(X^2)-\mu^2$$

The above theorem often greatly reduces the labor in computing the variance of a random variable. Note that

$$E(X^2)=\sum_x x^2 f(x)$$

which follows from definition 9.2

Example 9.2: Using Theorem 9.4, find the variance and standard deviation of the random variable X in Example 9.1.

Solution: Employing the definition of the expected value of X^2, we have

$$E(X^2)=\sum x^2 f(x)$$
$$=(-3)^2(.1)+(-2)^2(.3)+(0)^2(.15)+(1)^2(.4)+(2)^2(.05)=2.7$$

Theorem 9.4 now yields

$$V(X)=E(X^2)-\mu^2=2.7-(-0.4)^2=2.54$$

which is consistent with our previous result.

The standard deviation of the random variable is the positive square of the variance so that

$$\sigma=\sqrt{2.54}=1.59.$$

Example 9.3: In a coin-tossing game, a man is promised to receive Tk. 5 if he gets all heads or all tails when three coins are tossed and he pays off (loses) Tk. 3 if either one or two heads appear. How much is he expected to gain in the long run?

Solution: The random variable here is the amount the man can win. If X is the random variable, then X will take on a value 5 when the coins show all heads and –3, otherwise. The table below shows the outcomes of the experiment, values of X and the associated probabilities:

Outcome:	*HHH*	*HTT*	*HTH*	*HHT*	*THH*	*THT*	*TTH*	*TTT*
x:	5	–3	–3	–3	–3	–3	–3	5
$f(x)$:	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

It appears from the above table that the variable X assumes values –3 and +5 with probabilities 6/8 and 2/8 respectively. Since the value –3 occurs 6 times and 5 occurs 2 times the expected value of X is

$$E(X)=\sum xf(x)=-3\left(\frac{6}{8}\right)+5\left(\frac{2}{8}\right)=-1$$

Thus the man is expected to lose Tk. 1 in the long run.

Let us now examine what happens if the man receives Tk. 5 for all heads or all tails, Tk. '0' for 2 heads and pays off Tk 3 for 1 head. The random variable X will now assume the values, 5, 0 and –3 with associated probabilities 2/8, 3/8 and 3/8 respectively. The expected value in this case will be

$$E(X)=\sum xf(x)==5\left(\frac{2}{8}\right)+0\left(\frac{3}{8}\right)+(-3)\left(\frac{3}{8}\right)=\frac{1}{8}=0.125$$

This shows that the man will be marginally gainer winning only 12.5 *paisa* if the payment is made as above.

Example 9.4: Let X have the uniform distribution on the first N natural numbers with the following probability mass function:

$$P(X=k)=\frac{1}{N},\quad k=1,2,\ldots,N$$

Find the mean and variance of this distribution.

Solution: By definition

$$E(X)=\sum_{x=1}^{N}x\left(\frac{1}{N}\right)=\frac{1+2+\ldots+N}{N}=\frac{N+1}{2}.$$

$$E(X^2)=\sum_{x=1}^{N}x^2\left(\frac{1}{N}\right)=\frac{1^2+2^2+\ldots+N^2}{N}$$

$$=\frac{N(N+1)(2N+1)}{6N}=\frac{(N+1)(2N+1)}{6}$$

Hence

$$V(X)=E(X^2)-[E(X)]^2$$

$$=\frac{(N+1)(2N+1)}{6}-\left(\frac{N+1}{2}\right)^2$$

$$=\frac{N^2-1}{12}.$$

Example 9.5: A life insurance company in Bangladesh offers to sell a Tk. 25000 one-year term life insurance policy to a 25- year-old man for a premium of Tk. 2500. According to Bangladesh life table, the probability of surviving one year for a 25-year-old man is 0.97 and of his dying is 0.03. What is the company's expected gain in the long-run?

Solution: The gain X is a random variable that may take on the values Tk. 2500, if the man survives or 2500−25000 = −Tk.22500 if he dies. Consequently, the probability distribution of X is as follows:

x:	2500	−22500
$f(x:)$:	0.97	0.03

and consequently

$$E(X)=2500\times0.97-22500\times.03=1750$$

Thus the company's ultimate gain is Tk. 1750.

Example 9.6: Let X denote the number of spots showing on the face of a well-balanced die after it is rolled once. If $Y=X^2+2X$, find (a) $E(X)$, $E(Y)$, $E(X^2)$, $E(Y^2)$ and hence the variance of X and variance of Y.

Solution: The random variables X and Y together with their probability distributions are shown in the following table:

Values of X:	1	2	3	4	5	6
Values of Y:	3	8	15	24	35	48
$f(x)=f(y)$:	1/6	1/6	1/6	1/6	1/6	1/6

Hence

$$E(X)=\sum xf(x)=\frac{1}{6}(1+2+3+4+5+6)=3.5 \quad \text{and}$$

$$E(Y)=\sum yf(y)=\frac{1}{6}(3+8+15+24+35+48)=22.17$$

To find the variance of X, we find first $E(X^2)$:

$$E(X^2)=\sum x^2 f(x)=\frac{1}{6}\left(1^2+2^2+...+6^2\right)=15$$

so that

$$\sigma_x^2 = E(X^2)-[E(X)]^2 = 15-3.5^2 = 2.75$$

Similarly, to find the variance of Y, we compute $E(Y^2)$:

$$E(Y^2)=\sum y^2 f(x)=\frac{1}{6}\left(3^2+8^2+...+48^2\right)=4403$$

And hence

$$\sigma_y^2 = E(Y^2)-[E(Y)]^2 = 4403-22.17^2 = 3911.49.$$

Example 9.7: Two players A and B, play a coin-tossing game. A gives B one dollar if a head turns up; otherwise, B pays A one dollar. If the probability that the coin shows a head is p, find the expected gain of A.

Solution: Let X denote the gain of A. Then

$$P(X=1)=P(\text{Tails})=1-p, \quad P(X=-1)=p$$

and

$$E(X)=(1)(1-p)+(-1)(p)=1-2p\begin{cases} >0, & \text{If and only if } p<\frac{1}{2} \\ =0, & \text{if and only if } p=\frac{1}{2} \end{cases}$$

Example 9.8: Let X designate the number of heads obtained in the toss of three ideal coins. If two or more heads turn up, you receive Tk. 10; otherwise you pay Tk. 5. Find

(a) The expected number of heads to turn up.
(b) The expected amount of money you would win.

Solution: Here X is the random variable representing the number of heads and Y is the amount of money that is received on the outcome of the experiment. The outcomes of the experiment and the associated probabilities are shown in the table below:

Outcome	Values of X	Values of Y	$f(x)=f(y)$
HHH	3	10	1/8
HHT	2	10	1/8
HTH	2	10	1/8
HTT	1	–5	1/8
THH	2	10	1/8
THT	1	–5	1/8
TTH	1	–5	1/8
TTT	0	–5	1/8

The distribution of X can now be summarized as follows:

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Hence

$$E(X)=\sum x\,f(x)=0\left(\frac{1}{8}\right)+1\left(\frac{3}{8}\right)+2\left(\frac{3}{8}\right)+3\left(\frac{1}{8}\right)=1.5 \text{ heads}$$

Similarly the distribution of Y can be summarized as follows:

y:	–5	10
$f(y)$:	4/8	4/8

Hence

$$E(Y)=\sum y\,f(y)=-5\left(\frac{4}{8}\right)+10\left(\frac{4}{8}\right)=\frac{20}{8}=\text{Tk. } 2.5.$$

Hence on the average, you are expected to receive Tk.2.5 per toss in the long run.

Example 9.9: Given the following discrete distribution

$$f(x)=\frac{e^{-m}m^x}{x!},\qquad x=0,1,2,..\infty$$

Find $E(X)$.

Solution: By definition

$$E(X) = \sum_{x=0}^{\infty} xf(x) = \sum_{x=0}^{\infty} \frac{x e^{-m} m^x}{x!} = e^{-m} \sum_{x=0}^{\infty} \frac{x m^x}{x!}$$

$$= e^{-m} m \left(1 + m + \frac{m^2}{2!} + \frac{m^3}{3!} + \ldots\right) = e^{-m} m e^m = m$$

Example 9.10: A lot of 7 markers is sampled by a quality inspector; the lot contains 4 good markers and 3 defective markers. A sample of 3 is taken by the inspector. Find the expected value of the number of good markers in this sample.

Solution: Let X represent the number of good markers in the sample. It can be shown that the probability distribution of X is

$$f(x) = \frac{\left({}^4C_x\right)\left({}^3C_{3-x}\right)}{{}^7C_3}, \quad x = 0, 1, 2, 3.$$

Calculation shows that the probability distribution of X is as shown in the accompanying table:

Values of X	0	1	2	3
$f(x)$:	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$

Therefore

$$E(X) = (0)\left(\frac{1}{35}\right) + (1)\left(\frac{12}{35}\right) + (2)\left(\frac{18}{35}\right) + (3)\left(\frac{4}{35}\right) = 1.7$$

Thus if a sample of 3 markers is selected at random over and over again from a lot of 4 good markers and 3 defective markers, it would contain, on average 1.7 good markers.

Example 9.11: Let the random variable X have the following probability distribution

x:	-1	0	1
$f(x)$:	0.2	0.3	0.5

Compute the expected value of (i) $E(X)$ (ii) $E(2X)$ (iii) $E(X+2)$ (iv) $E(X^2)$ and(iii) $E(3X+1)$

Solution: (i) $E(X)=\sum xf(x)=(-1)\times(0.2)+(0)\times(0.3)+1\times0.5=0.3$

(ii) To compute $E(2X)$, we prepare the following table

Values of $2X$	-2	0	2
$f(x)$:	0.2	0.3	0.5

Since $2X=-2$ is equivalent to $X=-1$, the computation of $P(2X=-2)$ will be equivalent to the computation of $P(X=-1)$. Hence

$$P(2X=-2)=P(X=-1)=f(-1)=0.2.$$

Similarly,

$$P(2X=0)=P(X=0)=f(0)=0.3 \text{ and } P(2X=2)=P(X=1)=f(1)=0.5.$$

This shows that the probability distribution of the function of the random variable X remains the same as that of the probability distribution of X.

Thus

$$E(2X)=\sum 2x\ \times f(x)=(-2)\times(.2)+(0)\times(.3)+(2)\times(.5)=0.6$$
$$=(2)\times(0.3)=2\,E(X).$$

This result is true not only in this particular case, but is also true in general. This may be stated as follows:

If a is any arbitrary constant, then

$$E(aX)=aE(X)$$

(iii) To evaluate $E(X+2)$, we construct the following table:

Values of $X+2$:	1	2	3
$f(x)$:	0.2	0.3	0.5

Hence

$$E(X+2)=(1)\times(.2)+(2)\times(.3)+(3)\times(.5)=.2+.6+1.5=2.3=0.3+2=E(X)+2.$$

This result suggests that, for any arbitrary constant a,

$$E(X+a)=E(X)+a$$

(iv) To find $E(X^2)$, the accompanying table is constructed

Values of X^2:	1	0	1
$f(x)$:	0.2	0.3	0.5

Hence

$$E(X^2)=(1)\times(.2)+(0)\times(.3)+(1)\times(.5)=0.7$$

(v) To evaluate $E(3X+1)$, the following table is constructed

Values of $3X+1$:	-2	1	4
$f(x)$:	0.2	0.3	.0.5

Thus

$$E(3X+1)=(-2)\times(.2)+(1)\times(.3)+(4)\times(.5)=1.9=(3)\times(0.3)+1$$
$$=3E(X)+1,$$

which leads to the following general result:

$$E(aX+b)=aE(X)+b$$

where a and b are the arbitrarily chosen constants. We establish this relationship below:

***Theorem* 9.5:** *Let X be a random variable with a finite mean. Then for any numerical constants a and b,* $E(aX+b)=aE(X)+b$

Proof: *Case I: If X is discrete*

If X is a discrete random variable with the probability function $f(x)$, the following table can be prepared toward proving the theorem:

$w(x)$:	ax_1+b	ax_2+b	ax_3+b	...	ax_n+b
$f(x)$:	$f(x_1)$	$f(x_2)$	$f(x_3)$	...	$f(x_n)$

By definition

$$E(aX+b)=\sum(ax_i+b)f(x_i)$$
$$=(ax_1+b)f(x_1)+(ax_2+b)f(x_2)+\ldots+(ax_n+b)f(x_n)$$
$$=ax_1f(x_1)+ax_2f(x_2)+\ldots\ldots\ldots\ldots..+ax_nf(x_n)$$
$$+bf(x_1)+bf(x_2)+\ldots\ldots\ldots\ldots.+bf(x_n)$$
$$=a\big(x_1f(x_1)+x_2f(x_2)+\ldots\ldots\ldots\ldots.+x_nf(x_n)\big)$$
$$+b\big(f(x_1)+f(x_2)+\ldots\ldots\ldots\ldots\ldots+f(x_n)\big)$$

$$= a\sum x_i f(x_i) + b\sum f(x_i)$$

$$= aE(X) + b \qquad [\text{since} \sum f(x_i) = 1$$

Case II: If X is continuous

$$E(aX+b) = \int_{-\infty}^{\infty}(ax+b)f(x)dx = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$$

$$= aE(X) + b, (\text{since} \int_{-\infty}^{\infty} f(x)dx = 1)$$

Example 9.12: For Example 9.9, compute $E(X-1)^2$, $E(X^2)$ and $E(3X+1)$.

Solution: By definition

$$E(X-1)^2 = \sum_{x=-1}^{1}(x-1)^2 f(x)$$

$$= (-1-1)^2 \times 0.2 + (0-1)^2 \times 0.3 + (1-1)^2 \times 0.5 = 1.1$$

$$E(X^2) = \sum_{x=-1}^{1} x^2 f(x)$$

$$= (-1)^2 \times 0.2 + (0)^2 \times 0.3 + (1)^2 \times 0.5 = 0.7$$

$$E(3X+1) = \sum_{x=-1}^{1}(3X+1)f(x)$$

$$= (-3+1)\times 0.2 + (0+1)\times 0.3 + (3+1)\times 0.5 = 1.9$$

Example 9.13: Find the expected value of the random variable X and also of its square having the following density function:

$$f(x) = 2(1-x), \quad 0<x<1$$
$$=0, \quad \text{elsewhere}$$

Find also the variance of X.

Solution: By definition

$$E(X) = \int_{-\infty}^{\infty} xf(x) = 2\int_0^1 x(1-x)dx = 2\int_0^1 xdx - 2\int_0^1 x^2 dx = 1 - \frac{2}{3} = \frac{1}{3}$$

$$E(X^2) = \int_0^1 x^2 f(x)dx = 2\int_0^1 x^2(1-x)dx = 2\int_0^1 x^2 dx - 2\int_0^1 x^3 dx = \frac{1}{6}.$$

Hence the variance of X is

$$\sigma_x^2 = E(X^2) - (E(X))^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$

Example 9.14: X is a continuous random variable having the following probability density function

$$f(x) = e^{-x}, \quad x > 0$$
$$= 0, \quad \text{otherwise}$$

Find the expected value of $w(X) = e^{\frac{2x}{3}}$

Solution: Following the definition of $w(X)$

$$E[w(X)] = \int_0^{\infty} w(x) f(x) dx = \int_0^{\infty} e^{2x/3} e^{-x} dx = \int_0^{\infty} e^{-x/3} dx = -3(e^{-x/3})\Big|_0^{\infty} = 3$$

9.2 EXPECTED VALUE OF A FUNCTION OF TWO RANDOM VARIABLES

The notion of mathematical expectation can be extended to functions of two or more random variables. We will deal here with the case of two variables, which can analogously be extended for 3 or more variables. Let X and Y be two random variables with joint probability distribution $f(x,y)$ The expected value of the function $w(X,Y)$ is defined as

$$E[w(X,Y)] = \sum_x \sum_y w(x,y) f(x,y), \quad \text{if } X \text{ and } Y \text{ are discrete}$$

$$= \int_{-\infty}^{\infty} \int_{\infty}^{\infty} w(x,y) f(x,y) dx dy, \quad \text{if } X \text{ and } Y \text{ are continuous}$$

Further if $w(X, Y)$ is a function of the random variables X and Y, and c is a constant, then

$$E[cw(X,Y)] = cE[w(X,Y)] = c\sum_y \sum_x w(x,y) f(x,y)$$

And also if X and Y are two random variables and $w_1(X, Y), w_2(X, Y)$ are the functions of X and Y, then

$$E[w_1(X,Y) + w_2(X,Y)] = E[w_1(X,Y)] + E[w_2(X,Y)]$$

Theorem 9.6: *The expected value of the sum of two random variables X and Y is the sum of the expected values of the variables. Symbolically*

$$E(X+Y) = E(X) + E(Y)$$

Proof: We present the proof for both discrete and continuous cases

Case I: *When X and Y are discrete random variables*

By definition

$$E(X+Y) = \sum_x \sum_y (x+y) f(x,y)$$

$$= \sum_x \sum_y x f(x,y) + \sum_y \sum_x y f(x,y)$$

Now summing over y for the first term and x for the second term in the right hand side of the above expression, we have

$$\sum_x \sum_y x f(x,y) = \sum_x x g(x) \text{ and } \sum_y \sum_x y f(x,y) = \sum_y y h(y)$$

where $g(x)$ and $h(y)$ are the marginal probability mass functions of X and Y respectively.

Hence

$$E(X+Y) = \sum x g(x) + \sum y h(y) = E(X) + E(Y)$$

Case II: *When X and Y are continuous random variables*

$$E(X+Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f(x,y) dx dy$$

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

$$= \int_{-\infty}^{\infty} x g(x) dx + \int_{-\infty}^{\infty} y h(y) dy = E(X) + E(Y)$$

Note: The result also holds true for the difference $X-Y$, in which case, $E(X-Y) = E(X) - E(Y)$.

The addition theorem above can be established for any positive integer k by an induction argument, *i.e.*

$$E\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} E(X_i)$$

It is important to note that the above results are true regardless of whether or not the random variables are independent.

***Corollary* 9.1:** *If $X \geq Y$, then $E(X) \geq E(Y)$.*

***Corollary* 9.2:** *If a and b are two constants, then $E(aX + bY) = aE(X) + bE(Y)$*

***Corollary* 9.3:** *If $a_1, a_2, \ldots, a_k$ are k constants, then for k RVs $X_1, X_2, \ldots, X_k$*

$$E\left[\sum_{i=1}^{k} a_i X_i\right] = a_i \sum_{i=1}^{k} E(X_i)$$

In particular, for two *RVs* X_1 and X_2,

$$E(a_1X_1 + a_2X_2) = a_1\, E(X_1) + a_i\, E(X_2)$$

***Theorem* 9.7:** *Let X and Y be two independent random variables and g(X) and h(Y) be functions of only X and Y respectively. Then*

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

Proof: Let $f(x, y)$ denote the joint density function of X and Y. The product $g(X)\, h(Y)$ is a function of X and Y. Hence by definition

$$E[g(X)h(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x)f(y)dxdy$$

$$= \int_{-\infty}^{\infty} g(x)f(x)\left[\int_{-\infty}^{\infty} h(y)f(y)dy\right]dx$$

$$= \int_{-\infty}^{\infty} g(x)f(x)\,E[h(y)]dx$$

$$= E[h(y)]\int_{-\infty}^{\infty} g(x)f(x)\,dx$$

$$= E[g(X)]E[h(Y)]$$

This completes the proof.

The proof for the discrete case follows in an analogous manner.

***Theorem 9.8**: The expected value of the two random variables X and Y is equal to the product of their expected values, only when the variables are independent i.e.*

$$E(XY) = E(X)E(Y)$$

Or in other words,

The expected value of the product of two random variables is equal to the product of their expectations.

Proof: We present the proof for both discrete and continuous cases

***Case I**: When X and Y are discrete RVs*

Since X and Y are independent,

$$f(x, y) = g(x) \times h(y)$$

where $g(x)$ and $h(y)$ are the marginal probability functions of X and Y respectively. Thus

$$E(XY) = \sum_x \sum_y xy\, f(x,y) = \sum_x \sum_y xy\, g(x)h(y)$$

$$= \sum_x xg(x) \times \sum_y yh(y) = E(X)E(Y)$$

***Case II**: When X and Y are continuous RVs*

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

Since X and Y are independent,

$$f(x, y) = g(x) \times h(y)$$

where $g(x)$ and $h(y)$ are the marginal densities of X and Y respectively.

Hence

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, g(x)\, h(y) dxdy = \int_{-\infty}^{\infty} xg(x)dx \times \int_{-\infty}^{\infty} yh(y)dy = E(X)E(Y).$$

The result can be generalized for any number of independent variables. This implies that if $X_1, X_2, \ldots\ldots, X_n$ are n independent random variables, then

$$E(X_1 X_2 \ldots X_n) = E(X_1)\, E(X_2) \ldots E(X_n)$$

Note: The difference between the theorems 9.2 and 9.3 should be emphasized. If it is assumed that each expectation exists, the expectation of the sum of several random variables is always equal to the sum of their individual expectations. However, the expectation of the product of several random variables will be equal to the product of their individual expectations only when the variables are independent.

Example 9.15: Given the following joint distribution of X and Y. Find $E(X)$, $E(Y)$, $E(X+Y)$ and $E(XY)$.

		Y		
X	−3	2	4	Sum
1	0.1	0.2	0.2	0.5
3	0.3	0.1	0.1	0.5
Sum	0.4	0.3	0.3	1.0

Solution: The expected value of X can be obtained from the joint distribution as

$$E(X)=\sum_x\sum_y xf(x,y)=\sum_x xg(x),$$

where $g(x)$ is the marginal distribution of X.

Thus,

$$E(X)=x_1g(x_1)+x_2g(x_2)=(1)\times(0.5)+(3)\times(0.5)=2.0$$

Similarly, expected value of Y can be obtained as follows:

$$E(Y)=\sum_x\sum_y yf(x,y)=\sum_y yh(y),\ \text{summing over } x$$

where $h(y)$ is the marginal distribution of Y.

This gives,

$$E(Y)=y_1h(y_1)+y_2h(y_2)=(-3)\times(0.4)+(2)\times(0.3)+4\times0.3=0.6$$

The expected value of the sum $X+Y$ can be obtained as follows:

$$E(X+Y)=\sum_x\sum_y(x+y)f(x,y)$$

$$=(1-3)\times(0.1)+(1+2)\times(0.2)+(1+4)\times(0.2)+(3-3)\times(0.3)$$
$$+(3+2)\times(0.1)+(3+4)\times(0.1)=2.6$$

Note that the sum $X+Y$ has been obtained by adding all possible values of X and Y. A close examination of the entries in the above table shows that

Theorem 9.8: *The expected value of the two random variables X and Y is equal to the product of their expected values, only when the variables are independent i.e.*

$$E(XY) = E(X)E(Y)$$

Or in other words,

The expected value of the product of two random variables is equal to the product of their expectations.

Proof: We present the proof for both discrete and continuous cases

Case I: *When X and Y are discrete RVs*

Since X and Y are independent,

$$f(x, y) = g(x) \times h(y)$$

where $g(x)$ and $h(y)$ are the marginal probability functions of X and Y respectively. Thus

$$E(XY) = \sum_x \sum_y xy\, f(x,y) = \sum_x \sum_y xy\, g(x)h(y)$$

$$= \sum_x xg(x) \times \sum_y yh(y) = E(X)E(Y)$$

Case II: *When X and Y are continuous RVs*

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

Since X and Y are independent,

$$f(x, y) = g(x) \times h(y)$$

where $g(x)$ and $h(y)$ are the marginal densities of X and Y respectively.

Hence

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, g(x)\, h(y) dxdy = \int_{-\infty}^{\infty} xg(x)dx \times \int_{-\infty}^{\infty} yh(y)dy = E(X)E(Y).$$

The result can be generalized for any number of independent variables. This implies that if $X_1, X_2, \ldots\ldots, X_n$ are n independent random variables, then

$$E(X_1 X_2 \ldots X_n) = E(X_1)\, E(X_2) \ldots E(X_n)$$

Note: The difference between the theorems 9.2 and 9.3 should be emphasized. If it is assumed that each expectation exists, the expectation of the sum of several random variables is always equal to the sum of their individual expectations. However, the expectation of the product of several random variables will be equal to the product of their individual expectations only when the variables are independent.

Example 9.15: Given the following joint distribution of X and Y. Find $E(X)$, $E(Y)$, $E(X+Y)$ and $E(XY)$.

		Y		
X	−3	2	4	Sum
1	0.1	0.2	0.2	0.5
3	0.3	0.1	0.1	0.5
Sum	0.4	0.3	0.3	1.0

Solution: The expected value of X can be obtained from the joint distribution as

$$E(X)=\sum_x\sum_y xf(x,y)=\sum_x xg(x),$$

where $g(x)$ is the marginal distribution of X.

Thus,

$$E(X)=x_1g(x_1)+x_2g(x_2)=(1)\times(0.5)+(3)\times(0.5)=2.0$$

Similarly, expected value of Y can be obtained as follows:

$$E(Y)=\sum_x\sum_y yf(x,y)=\sum_y yh(y), \text{ summing over } x$$

where $h(y)$ is the marginal distribution of Y.

This gives,

$$E(Y)= y_1h(y_1)+y_2h(y_2)=(-3)\times(0.4)+(2)\times(0.3)+4\times0.3=0.6$$

The expected value of the sum $X+Y$ can be obtained as follows:

$$E(X+Y)=\sum_x\sum_y (x+y)f(x,y)$$

$$=(1-3)\times(0.1)+(1+2)\times(0.2)+(1+4)\times(0.2)+(3-3)\times(0.3)$$
$$+(3+2)\times(0.1)+(3+4)\times(0.1)=2.6$$

Note that the sum $X+Y$ has been obtained by adding all possible values of X and Y. A close examination of the entries in the above table shows that the sum $X+Y$ assumes the values −2, 0, 3, 5, and 7 with associated probabilities 0.1, 0.3, 0.2, 0.3 and 0.1. Hence the expected value of X can also be obtained from the distribution of the sum of the random variables X and Y. In tabular form, the distribution of $X+Y$ will be as follows

$x+y$:	−2	0	3	5	7
$f(x,y)$:	0.1	0.3	0.2	0.3	0.1

Hence

$$E(X+Y)=(-2)\times(0.1)+(0)\times(0.3)+(3)\times(0.2)+(5)\times(0.3)+(7)\times(0.1)=2.6,$$

a result, which is the same as obtained before.

Now to obtain $E(XY)$, we have, as before

$$E(XY)=(1)\times(-3)\times(0.1)+(1)\times(2)\times(0.2)+(1)\times(4)\times(0.2)+(3)\times(-3)\times(0.3)$$
$$+(3)\times(2)\times(0.1)+(3)\times(4)\times(0.1)=0$$

It is easy to verify that the joint distribution of X and Y is as follows:

xy:	−9	−3	2	4	6	12
$f(x, y)$:	0.3	0.1	0.2	0.2	0.1	0.1

Check that from this table too $E(XY)=0$

Example 9.16: Given the following joint density function of X and Y:

$$f(x,y)=4xy, \quad 0<x<1, \quad 0<x<1$$
$$=0, \quad \text{elsewhwre}$$

Obtain $E(X)$, $E(Y)$, $E(XY)$ and $E(X+Y)$

Solution:

$$E(X)=\int_0^1\int_0^1 x\,f(x,y)dxdy=4\int_0^1\int_0^1 x^2ydxdy$$

$$=\frac{4}{2}\int_0^1 (x^2y^2)\Big|_{y=0}^{y=1}dx=2\int_0^1 x^2dx=\frac{2}{3}x^3\Big|_0^1=\frac{2}{3}$$

$$E(Y)=\int_0^1\int_0^1 yf(x,y)dxdy=4\int_0^1\int_0^1 xy^2dxdy$$

$$=\frac{4}{3}\int_0^1 (xy^3)\Big|_{y=0}^{y=1}dx=\frac{4}{3}\int_0^1 xdx=\frac{4}{6}x^2\Big|_0^1=\frac{2}{3}$$

$$E(XY) = \int_0^1\int_0^1 xyf(x,y)dxdy = 4\int_0^1\int_0^1 x^2y^2\,dxdy$$

$$= \frac{4}{3}\int_0^1 x^2y^3\Big|_{y=0}^{y=1}dx = \frac{4}{9}x^3\Big|_0^1 = \frac{4}{9}$$

$$E(X+Y) = \int_0^1\int_0^1 (x+y)f(x,y)\,dxdy = \int_0^1\int_0^1 (x^2y+xy^2)dxdy$$

$$= 4\int_0^1\left(\frac{1}{2}x^2y^2 + \frac{1}{3}xy^3\right)\Bigg|_{y=0}^{y=1}dx = 4\int_0^1(\frac{1}{2}x^2 + \frac{1}{3}x)\,dx$$

$$= 4\left(\frac{1}{6}x^3 + \frac{1}{6}x^2\right)\Bigg|_0^1 = \frac{4}{3}.$$

Note: To find $E(X)$ and $E(Y)$ from the joint density function $f(x,y)$, first we evaluate the marginal densities $g(x)$ and $h(y)$ as follows:

$$E(X) = \begin{cases} \sum_x\sum_y xf(x,y) = \sum_x xg(x), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy = \int_{-\infty}^{\infty} xg(x)dx, & \text{if } X \text{ is continuous} \end{cases}$$

Again

$$E(Y) = \begin{cases} \sum_x\sum_y yf(x,y) = \sum_y yh(y), & \text{if } Y \text{ is discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy = \int_{-\infty}^{\infty} yh(y)dy, & \text{if } Y \text{ is continuous} \end{cases}$$

For discrete cases, evaluation of $E(X)$ and $E(Y)$ from $f(x,y)$ is rather easier, for, the marginal probability functions are readily available in most cases from the tabular presentation of the distribution.

Example 9.17: Given the following density function

2014

$$f(x,y) = 2(x+y-2xy),\quad 0<x<1,\quad 0<y<1$$
$$= 0,\quad \text{elsewhere}$$

(a) Find $E(X)$, $E(Y)$, $E(X+Y)$ and $E(XY)$.

(b) Also verify whether $E(X+Y) = E(X)+E(Y)$.

(c) Are X and Y independent?

Solution: The marginal density of X is

$$g(x) = 2\int_0^1 (x+y-2xy)dy = 2\left(xy + \frac{y^2}{2} - \frac{2xy^2}{2}\right)\Bigg|_{y=0}^{y=1} = 2\left(x + \frac{1}{2} - x\right) = 1$$

Hence

$$E(X) = \int_0^1 xg(x)dx = \int_0^1 xdx = \frac{1}{2}$$

Similarly

$$h(y) = 2\int_0^1 (x+y-2xy)dx = 2\left(\frac{x^2}{2} + xy - \frac{2yx^2}{2}\right)\Bigg|_{x=0}^{x=1} = 2\left(\frac{1}{2} + y - y\right) = 1$$

and

$$E(Y) = \int_0^1 yh(y)dy = \int_0^1 ydy = \frac{1}{2}$$

$$E(X+Y) = 2\int_0^1\int_0^1 (x+y)(x+y-2xy)dxdy$$

$$= 2\int_0^1\int_0^1 (x^2+y^2+2xy-2x^2y-2xy^2)dxdy$$

$$= 2\int_0^1 \left(x^2y + \frac{y^3}{3} + xy^2 - x^2y^2 - \frac{2xy^3}{3}\right)\Bigg|_0^1 dx$$

$$= 2\int_0^1 \left(x^2 + \frac{1}{3} + x - x^2 - \frac{2x}{3}\right)dx$$

$$= 2\int_0^1 \left(\frac{1}{3} + \frac{x}{3}\right)dx = \frac{2}{3}\left(x + \frac{x^2}{2}\right)\Bigg|_0^1 = 1$$

Since $E(X) + E(Y) = \frac{1}{2} + \frac{1}{2} = 1$, we establish that $E(X+Y)=E(X)+E(Y)$.

$$E(XY) = \int_0^1\int_0^1 xyf(x,y)dxdy$$

$$= \int_0^1\int_0^1 xy(x+y-2xy)dxdy$$

$$= 2\int_0^1 \left(\frac{yx^3}{3} + \frac{y^2x^2}{2} - \frac{2y^2x^3}{3} \right) \Bigg|_{x=0}^{x=1}$$

$$= \int_0^1 \int_0^1 (x^2y + y^2x - 2x^2y^2) dxdy\, dy$$

$$= 2\left(\frac{y^2}{6} + \frac{y^3}{6} - \frac{2y^3}{9} \right) \Bigg|_0^1 = \frac{2}{9}$$

$$= 2\int_0^1 (\frac{y}{3} + \frac{y^2}{2} - \frac{2y^2}{3}) dy$$

Now

$$E(X)E(Y) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \text{ and } E(XY) = \frac{2}{9}$$

Since $E(XY)$ is not equal to $E(X) \times E(Y)$, the variables are not independent.

Example 9.18: The random variables X and Y have the following density function:

$$f(x, y) = \frac{1}{8}(6 - x - y), \qquad 0 < x < 2, \ 2 < y < 4$$

$$= 0, \qquad \text{otherwise.}$$

Find $E(X)$, $E(Y)$ and $E(XY)$. Are X and Y independent?

Solution: The marginal density of X is

$$g(x) = \frac{1}{8}\int_2^4 (6 - x - y)\, dy = \frac{1}{8}(6y - xy - \frac{y^2}{2}) \Bigg|_{y=2}^{y=4}$$

$$= \frac{1}{8}[6(4-2) - x(4-2) - \frac{1}{2}(16-4)] = \frac{1}{4}(3-x)$$

Hence

$$E(X) = \int_0^2 xg(x)dx = \frac{1}{4}\int_0^2 x(3-x)dx = \frac{1}{4}\int_0^2 (3x - x^2)dx$$

$$= \frac{1}{4}\left(\frac{3x^2}{2} - \frac{x^3}{3} \right) \Bigg|_0^2 = \frac{1}{4}\left[\frac{3}{2}(4-0) - \frac{1}{3}(8-0) \right] = \frac{5}{6}$$

Similarly, the marginal density of Y is

$$h(y)=\frac{1}{8}\int_0^2(6-x-y)dx=\frac{1}{8}\left(6x-\frac{x^2}{2}-xy\right)\Big|_0^2$$

$$=\frac{1}{8}\left[6(2-0)-\frac{1}{2}(4-0)-y(2-0)\right]=\frac{1}{4}(5-y)$$

$$E(Y)=\int_2^4 yh(y)dy=\frac{1}{4}\int_2^4 y(5-y)dy=\frac{1}{4}\int_2^4(5y-y^2)dy$$

$$=\frac{1}{4}\left(\frac{5}{2}y^2-\frac{y^3}{3}\right)\Big|_2^4=\frac{1}{4}\left[\frac{5}{2}(16-4)-\frac{1}{3}(64-8)\right]=\frac{17}{6}$$

$$E(XY)=\frac{1}{8}\int_0^2\int_2^4 xy(6-x-y)dxdy=\frac{1}{8}\int_0^2\left(3xy^2-\frac{x^2y^2}{2}-\frac{xy^3}{3}\right)\Big|_{y=2}^{y=4}dx$$

$$=\frac{1}{8}\int_0^2\left(36x-6x^2-\frac{56}{3}x\right)dx=\frac{1}{8}\left(18x^2-2x^3-\frac{28}{3}x^2\right)\Big|_0^2$$

$$=\frac{1}{12}\left(13x^2-3x^3\right)\Big|_0^2=\frac{1}{12}(52-24)=\frac{7}{3}$$

We note that

$$E(X)\times E(Y)=\frac{5}{6}\times\frac{17}{6}=\frac{85}{36}$$

But

$$E(XY)=\frac{7}{3}$$

showing that

$$E(XY)\neq E(X)E(Y)$$

Hence X and Y are not independent.

Example 9.19: Given the following joint density function of the random variables X and Y:

$$f(x,y)=\frac{x(3y^2+1)}{4},\quad 0<x<2,\ 0<y<1$$
$$=0,\qquad \text{elsewhere}$$

(i) Examine whether X and Y are independent.
(ii) Also check that $E(X+Y)=E(X)+E(Y)$

(iii) Find $E\left(\frac{Y}{X}\right)$

(iv) Verify if $E\left(\frac{Y}{X}\right)=\frac{E(Y)}{E(X)}$

Solution: We first compute the quantities $E(X)$, $E(Y)$, $E(X+Y)$ and $E(XY)$.

$$g(x)=\int_0^1 \frac{x(3y^2+1)}{4}dy=\int_0^1 \frac{(3xy^2+x)}{4}dy$$

$$=\frac{1}{4}\left(xy^3+xy\right)\Big|_0^1=\frac{x}{2},\qquad 0<x<2$$

$$h(y)=\int_0^2 \frac{x(3y^2+1)}{4}dx=\int_0^2 \frac{(3xy^2+x)}{4}dx$$

$$=\frac{1}{4}\left(\frac{3x^2y^2}{2}+\frac{x^2}{2}\right)\Big|_0^2=\frac{3y^2+1}{2},\quad 0<y<1$$

Hence

$$E(X)=\int_0^2 xg(x)dx=\int_0^2 \frac{x^2}{2}dx=\frac{x^3}{6}\Big|_0^2=\frac{4}{3}$$

$$E(Y)=\int_0^1 yh(y)dy=\int_0^1 \frac{(3y^3+y)}{2}dy=\frac{1}{2}\left(\frac{3y^4}{4}+\frac{y^2}{2}\right)\Big|_0^1=\frac{5}{8}$$

$$E(XY)=\int_0^1\int_0^2 xy\,f(x,y)dxdy$$

$$=\int_0^1\int_0^2 \frac{x^2y(3y^2+1)}{4}dxdy$$

$$=\int_0^1\left(\frac{x^3y^3}{4}+\frac{x^3y}{3}\right)\Big|_{x=0}^{x=2}dy=\int_0^1 \frac{2y(3y^2+1)}{3}dy$$

$$=\left(\frac{y^4}{2}+\frac{y^2}{3}\right)\Big|_0^1=\frac{5}{6}$$

Hence
$$E(X+Y)=\frac{47}{24}=\frac{4}{3}+\frac{5}{8}=E(X)+E(Y)$$

And since
$E(X)\,E(Y)=\frac{4}{3}\times\frac{5}{8}=\frac{5}{6}=E(XY)$, the variables X and Y are independent.

Finally
$$E\left(\frac{Y}{X}\right)=\int_0^1\int_0^2\frac{y}{x}\cdot\frac{x(1+3y^2)}{4}\,dx\,dy=\int_0^1\int_0^2\frac{y(1+3y^2)}{4}\,dx\,dy$$
$$=\int_0^1\frac{(y+3y^3)}{2}\,dy=\frac{5}{8}$$

And in general $E\left(\frac{Y}{X}\right)\neq\frac{E(Y)}{E(X)}$, since in this example $\frac{5}{8}\neq\left(\frac{5}{8}\div\frac{4}{3}=\frac{15}{32}\right)$

9.3 CONDITIONAL EXPECTATION

The idea of conditional expectation is analogous to the idea of the expected value of a random variable X (say) in terms of its probability distribution $f(x)$ defined as $\sum_x xf(x)$ or $\int_{-\infty}^{\infty}xf(x)dx$ according as X is discrete or continuous.

Suppose that X and Y are two random variables with a joint probability function $f(x, y)$. Let $g(x)$ stand for the marginal probability density function of X and $f(y|x)$ for the conditional probability function of Y given that $X=x$.

The conditional expectation of Y given X is written as $E(Y|X)$ and is defined as a function of the random variable X, whose value $E(Y|x)$, when $X=x$, is specified as follows:

$$E(Y\mid x)=\begin{cases}\int_{-\infty}^{\infty}y\,f(y\mid x)dy, & \text{when } X \text{ and } Y \text{ are contnuous}\\ \sum_y yf(y\mid x), & \text{when } X \text{ and } Y \text{ are discrete}\end{cases}$$

It seems obvious that $E(Y|x)$ is the mean of the conditional distribution of Y given that $Y=x$. The value of $E(Y|x)$ will not be defined for any value of x such that $g(x)=0$, where $g(x)$ is the marginal density of X.

We can define the conditional expectation of the random variable X given Y in a similar way:

$$E(X \mid y) = \begin{cases} \int_{-\infty}^{\infty} x f(x \mid y) dx, & \text{when } X \text{ and } Y \text{ are contnuous} \\ \sum_{x} x f(x \mid y), & \text{when } X \text{ and } Y \text{ are discrete} \end{cases}$$

In general, the conditional expectation of any function of Y, $G(Y)$, say, given $X=x$ is defined as

$$E[G(Y) \mid X = x] = \begin{cases} \int_{-\infty}^{\infty} G(Y) f(y \mid x) dy, & \text{when } X \text{ is contnuous} \\ \sum_{y} G(Y) f(y \mid x), & \text{when } X \text{ is discrete} \end{cases}$$

Siince $E(Y|X)$ is a function of the random variable X, it itself is a random variable. So we can logically evaluate the expected value of $E(Y|X)$. Furthermore, $E(Y|X)$ has its own probability distribution, which can be derived from the distribution of X. We will now prove an important theorem relating $E(Y|X)$ and $E(Y)$.

***Theorem* 9.9**: *For any random variables X and Y,*

$$E[E(Y \mid X)] = E(Y)$$

Proof: Let us assume that X and Y are continuous random variables. Then by definition

$$E(Y \mid x) = \int_{-\infty}^{\infty} y f(y \mid x) dy = \int_{-\infty}^{\infty} y \frac{f(x,y)}{g(x)} dy$$

where $f(x,y)$ is the joint pdf of X and Y, $f(y|x)$ is the conditional pdf of Y given $X=x$ and $g(x)$ is the marginal *pdf* of X.

Then

$$E[E(Y \mid X)] = \int_{-\infty}^{\infty} E(Y \mid x) g(x) dx = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y \frac{f(x,y)}{g(x)} dy \right] g(x) dx$$

$$= \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x,y) dx \right] dy = \int_{-\infty}^{\infty} y h(y) dy = E(Y)$$

When X is discrete, the proof of the theorem is as follows:

$$= y_1 \times \frac{f(1, y_1)}{g(1)} + y_2 \times \frac{f(1, y_2)}{g(1)}$$

$$= 0 \times \frac{f(1, 0)}{g(1)} + 1 \times \frac{f(1, 1)}{g(1)} = 1 \times \frac{2/8}{3/8} = \frac{2}{3}$$

When X=2

$$E(Y \mid X = 2) = \sum_{y=0}^{1} y f(y \mid x = 2) = \sum_{y=0}^{1} y \frac{f(2, y)}{g(2)}$$

$$= y_1 \times \frac{f(2, y_1)}{g(2)} + y_2 \times \frac{f(2, y_2)}{g(2)}$$

$$= 0 \times \frac{f(2,0)}{g(2)} + 1 \times \frac{f(2,1)}{g(2)} = 1 \times \frac{4/8}{5/8} = \frac{4}{5}$$

To find conditional variance of X for Y=0, the following distribution i constructed

Values of X^2:	1	4
$P(X \mid Y=0) = f(x \mid y=0)$:	1/2	1/2

Hence

$$E(X^2 \mid Y = 0) = 1 \times \frac{1}{2} + 4 \times \frac{1}{2} = \frac{5}{2}$$

and

$$V(X \mid Y = 0) = E(X^2 \mid Y = 0) - [E(X \mid Y = 0]^2 = \frac{5}{2} - \left(\frac{3}{2}\right)^2 = \frac{1}{4}$$

Example 9.21: Given the following conditional function

$$f(y \mid 2) = \begin{cases} 1/2, & y = 2 \\ 1/4, & y = 3 \\ 1/4, & y = 4 \end{cases}$$

Find $E(Y \mid X=2)$.

Solution: By definition

$$E[Y \mid X = 2] = \sum y f(y \mid 2)$$

$$= 2P(Y = 2 \mid X = 2) + 3P(Y = 3 \mid X = 2) + 4P(Y = 4 \mid X = 2)$$

$$= 2 \times \frac{1}{2} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = \frac{11}{4}$$

Example 9.22: Given the following joint density function of the random variables X and Y.

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

Find $E(X|Y)$ and $E(Y|X)$.

Solution:

$$E[X\,|\,Y] = \int_0^1 x f(x\,|\,y)dx = \int_0^1 x\frac{f(x,y)}{h(y)}dx$$

where

$$h(y) = \int_0^1 (x+y)dx = y + \frac{1}{2}, \qquad 0 \le y \le 1$$

Hence

$$E(X\,|\,Y) = \int_0^1 \frac{x(x+y)}{y+\frac{1}{2}}dx = \frac{\left[\frac{x^3}{3} + \frac{x^2 y}{2}\right]_{x=0}^{x=1}}{y+\frac{1}{2}} = \frac{2+3y}{3(2y+1)}$$

Again

$$E[Y\,|\,X] = \int_0^1 y f(y\,|\,x)dy = \int_0^1 y\frac{f(x,y)}{g(x)}dy$$

where

$$g(x) = \int_0^1 (x+y)dy = x + \frac{1}{2}, \qquad 0 < x < 1$$

Hence

$$E(Y\,|\,X) = \frac{\int_0^1 y(x+y)}{x+\frac{1}{2}}dy = \frac{\left[\frac{y^3}{3} + \frac{xy^2}{2}\right]_{y=0}^{y=1}}{x+\frac{1}{2}} = \frac{(2+3x)}{3(2x+1)}$$

9.5 MOMENT GENERATING FUNCTION

The moment of a distribution occupies a central position both in theoretical and applied statistics. Because of its role and importance, it would seem useful if a generalized function could be found that would give us a

Solution: The problem here is to find $P(16 < X < 24)$. Applying Chebychev's theorem

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

for $\mu = 20$ and $\sigma = 2$, it follows that

$$P(16 < X < 24) = P(\mu - 2\sigma < X < \mu - 2\sigma) \geq 1 - \frac{1}{2^2} = 0.75.$$

The result indicates that tomorrow's customer total will be between 16 and 24 with a fairly high probability of at least 3/4.

9.11 SPECIAL MATHEMATICAL EXPECTATIONS

Many statistical measures can be expressed in terms of expected values. In this section, we will present some measures of central tendency, dispersion and correlation in terms of expected values. This presentation will be highly useful in subsequent chapters, especially for dealing with probability distributions.

To start with, we define the rth central moment μ_r of a random variable X as follows:

$$\mu_r = E(X-\mu)^r = \sum_x (x-\mu)^r f(x), \qquad \text{if } X \text{ is discrete}$$

$$= \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx, \qquad \text{if } X \text{ is continuous}$$

The rth raw moment about the origin is defined as

$$\mu_r' = E(X^r) = \sum_x x^r f(x), \qquad \text{if } X \text{ is discrete}$$

$$= \int_{-\infty}^{\infty} x^r f(x)dx, \qquad \text{if } X \text{ continuous}$$

Arithmetic Mean

When $r=1$ in the above expression, $\mu_1' = E(X)$, which is simply the **mean** of the random variable X. We frequently use μ_X or simply μ to designate this mean. It is also easy to check that $\mu_1=0$.

The second central moment is the variance. This is derived by putting $r=2$ in μ_r. If we use σ_X^2 or μ_2 to designate this variance

$$\sigma_X^2 = E(X-\mu)^2$$

The above expression can be written as follows

$$\sigma_X^2 = E(X-\mu)^2 = E(X^2) - \mu^2$$

Corollary **9.4:** *For any constant a,* $V(a)=\sigma_a^2=0$.
Corollary **9.5:** $V(aX)=a^2\sigma_X^2$.
Corollary **9.6:** $V(a\pm bX)=b^2\,\sigma_X^2$.

Example 9.34: Find the mean and variance of the random variable X, having the following density function

$$f(x) = 2(x-1), \quad 1<x<2$$
$$= 0, \quad \text{otherwise}$$

Solution: The first raw moment and hence the mean of the distribution is

$$\mu_1' = E(X) = 2\int_1^2 x(x-1)dx = \frac{5}{3}$$

and the second raw moment is

$$\mu_2' = E(X^2) = 2\int_1^2 x^2(x-1)dx = \frac{17}{6}$$

Hence the variance of the distribution is

$$\sigma_X^2 = E(X^2) - \mu_1'^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}$$

Geometric mean

The geometric mean G of the random variable X in terms of expected value is defined as

$$\log G = E(\log X)$$

$$= \sum_x \log x\, f(x), \quad \text{if } X \text{ is discrete}$$

$$= \int_{-\infty}^{\infty} \log x\, f(x)dx, \quad \text{if } X \text{ is continuous.}$$

provided the integral exists.

Harmonic mean

The harmonic mean H of the random variable X in terms of expected value is

$$\frac{1}{H} = E\left(\frac{1}{X}\right)$$

$$= \sum_x \frac{1}{x} f(x), \qquad \text{if } X \text{ is discrete}$$

$$= \int_{-\infty}^{\infty} \frac{1}{x} f(x)dx, \qquad \text{if } X \text{ is continuous}$$

Covariance

Let X and Y be two random variables having a specified joint probability distribution and let $E(X)=\mu_X$ and $E(Y)=\mu_Y$. Also let Var(X)$=\sigma_X^2$ and Var(Y)$=\sigma_Y^2$. Then the covariance of X and Y, denoted by σ_{XY} is defined as

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$$

$$= \sum_x \sum_y (x-\mu_X)(y-\mu_Y), \qquad \text{if } X \text{ and } Y \text{ arediscrete}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)dxdy, \qquad \text{if } X \text{ and } Y \text{ arediscrete}$$

It can easily be verified that

$$E(X-\mu_X)(Y-\mu_Y) = E(XY) - E(X)\,E(Y)$$

It is obvious from the above expression that if X and Y are independent, then the covariance is zero, but the converse is not generally true.

Example 9.35: Given the following density function of X. Calculate the geometric mean and the harmonic mean.

$$f(x) = \frac{1}{18}(3+2x), \qquad 2 \le x \le 4.$$

Solution: If G stands for the geometric mean of X, then

$$\log G = E(\log X)$$

$$= \frac{1}{18}\int_2^4 \log x(3+2x)dx$$

$$= \frac{1}{6}\int_2^4 \log x dx + \frac{1}{9}\int_2^4 x \log x dx$$

$$=\frac{1}{6}(x\log x-x]_2^4+\frac{1}{9}\left(\frac{x^2}{2}\log x-\frac{x^2}{2}\right]_2^4$$

$$=\frac{1}{6}(2.16)+\frac{1}{6}(6.7041)=1.1049$$

Taking antilog, G=3.0189

If H stands for the harmonic mean

$$\frac{1}{H}=E\left(\frac{1}{X}\right)=\frac{1}{18}\int_2^4\frac{3+2x}{x}dx$$

$$=\frac{1}{6}\int_2^4\frac{dx}{x}+\frac{1}{9}\int_2^4 dx=\frac{1}{6}\log x\Big|_2^4+\left(\frac{1}{9}x\right)\Big|_2^4$$

$$=0.1155+0.2222=0.3377$$

Hence

$$H=\frac{1}{0.3377}=2.96.$$

Correlation coefficient

If $0<\sigma_X^2<\infty$ and $0<\sigma_Y^2<\infty$, then the population correlation coefficient ρ_{XY} between X and Y is defined as follows:

$$\rho_{xy}=\frac{E(X-\mu_X)(Y-\mu_Y)}{\sqrt{\left[E(X^2)-\mu_X{}^2\right]\left[E(Y^2)-\mu_Y{}^2\right]}}=\frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}=\frac{\sigma_{XY}}{\sigma_x\sigma_Y}$$

The population regression coefficients may also be computed using the following formula:

(i) Regression coefficient of Y on X is

$$\beta_{Y|X}=\frac{E(X-\mu_X)(Y-\mu_Y)}{E(X-\mu_X)^2}=\frac{\sigma_{XY}}{\sigma_X{}^2}$$

(ii) Regression coefficient of X on Y is

$$\beta_{X|Y}=\frac{E(X-\mu_X)(Y-\mu_Y)}{E(Y-\mu_y)^2}=\frac{\sigma_{XY}}{\sigma_Y{}^2}$$

Example 9.36: The joint distribution of two discrete random variables X and Y is as follows:

	Y			
X:	−3	2	4	Total
1	0.1	0.2	.0.2	0.5
3	0.3	0.1	0.1	0.5
Total	0.4	0.3	0.3	1.0

Compute Cov (X, Y), ρ_{XY} and the regression coefficients..

Solution: To find Cov (X, Y) and ρ_{XY}, we compute the following quantities

$$E(XY) = \sum_x \sum_y xy\, f(x, y)$$

$$= (1)\times(-3)\times(0.1) + (1)\times(2)\times(0.2) + (1)\times(4)\times(0.2)$$

$$E(X) = \mu_X = \sum_x xg(x) = (1)\times(0.5) + (3)\times(0.5) = 2$$

$$E(Y) = \mu_Y = \sum_y yh(x) = (-3)\times(0.4) + (2)\times(0.3) + (4)\times(0.4) = 0.6$$

Hence

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 0 - (2)\times(0.6) = -1.2$$

$$E(X^2) = \Sigma x^2 g(x) = (1)\times(0.5) + (9)\times(0.5) = 5$$

$$E(Y^2) = \Sigma y^2 h(y) = (9)\times(0.4) + (4)\times(0.3) + (16)\times 5 \times(0.3) = 9.6$$

Hence

$$\rho_{xy} = \frac{E(XY) - E(X).E(Y)}{\sqrt{\left[E(X^2) - \mu_X{}^2\right]\left[E(Y^2) - \mu_Y{}^2\right]}}$$

$$= \frac{-1.2}{\sqrt{(5-4)\times(9.6-0.36)}} = \frac{-1.2}{\sqrt{9.24}} = -0.3947$$

(i) Regression coefficient of Y on X is

$$\beta_{Y|X} = \frac{E(X-\mu_X)(Y-\mu_Y)}{E(X-\mu_X)^2} = \frac{-1.2}{1} = -1.2$$

(ii) Regression coefficient of X on Y is

$$\beta_{X|Y} = \frac{E(X-\mu_X)(Y-\mu_Y)}{E(Y-\mu_y)^2} = \frac{-1.2}{9.24} = -.13$$

Example 9.37: Find the coefficient of correlation between the variables X and Y given the following density function

$$f(x,y) = 2, \quad 0 < x < y, \quad 0 < y < 1$$
$$= 0, \quad \text{otherwise}$$

Solution: Here

$$\mu_X = \int_0^1 \int_0^y 2x dx dy = \int_0^1 (x^2) \Big|_0^y dy = \int_0^1 y^2 \, dy = \frac{1}{3}$$

and

$$\mu_Y = \int_0^1 \int_0^y 2y dx dy = 2\int_0^1 y(x) \Big|_0^y dy = 2\int_0^1 y^2 \, dy = \frac{2}{3}$$

Also

$$E(XY) = \int_0^1 \int_0^y 2xy \, dx dy = \frac{1}{4}$$

Thus

$$\sigma_{XY} = E(XY) - \mu_X \cdot \mu_Y = \frac{1}{4} - \left(\frac{1}{3}\right) \times \left(\frac{2}{3}\right) = \frac{1}{36}$$

Again

$$E(X^2) = \int_0^1 \int_0^y 2x^2 dx dy = \frac{2}{3} \int_0^1 (x^3) \Big|_{x=0}^{x=y} dy = \frac{2}{3} \int_0^1 (y^3) dy = \frac{1}{6}$$

and

$$E(Y^2) = \int_0^1 \int_0^y 2y^2 dx dy = 2\int_0^1 y^2(x) \Big|_{x=0}^{x=y} dy = 2\int_0^1 y^3 dy = \frac{1}{2}$$

$$\sigma_X^2 = E(X^2) - [E(X)]^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18},$$

$$\sigma_Y^2 = E(Y^2) - [E(Y)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

Hence

$$\rho_{XY} = \frac{E(XY) - E(X)E(Y)}{\sqrt{\left[E(X^2) - \{E(X)\}^2\right] \times \left[E(Y^2) - \{E(Y)\}^2\right]}} = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{1/36}{1/18} = \frac{1}{2}.$$

Example 9.38: A production process in an industry yields a product containing two types of defectives. For a specified sample from this process, let X denote the proportion of defectives in the sample and let Y denote the proportion of type I defectives among all impurities found. Suppose that the joint distribution of X and Y can be modeled by the following probability density function:

$$f(x,y)=\begin{cases}2(1-x), & 0\le x\le 1,\quad 0\le y\le 1\\ 0 & \text{elsewhere}\end{cases}$$

Find the expected value of the proportion of type I defectives in the sample. Test if X and Y are independent.

Solution: As it follows from the problem, XY is the proportion of type I defectives in the entire sample. Thus we want to evaluate $E\,(XY)$:

$$E(X\,Y)=\int_0^1\int_0^1 2xy(1-x))dydx=2\int_0^1 x(1-x)\left(\frac{1}{2}\right)dx$$

$$=\int_0^1(x-x^2)dx=\left(\frac{x^2}{2}-\frac{x^3}{3}\right)=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$$

Therefore, one would expect 1/6 of the samples to be made up of type I defectives.

To test the independence of the random variables, we would find $E(X)$ and $E(Y)$. From the joint density function, the marginal density of X is

$$g(x)=\begin{cases}2\int_0^1(1-x)dy=2(1-x), & 0\le x\le 1\\ 0 & \text{elsewhere}\end{cases}$$

Similarly, the marginal density of Y is

$$h(y)=\begin{cases}2\int_0^1(1-x)dx=1, & 0\le x\le 1\\ 0 & \text{elsewhere}\end{cases}$$

Thus

$$E(X)=2\int_0^1 x(1-x)dx=\frac{1}{3}$$

And

$$E(Y)=\int_0^1 ydy=\frac{1}{2}$$

It follows that

$$E(XY) = E(X)E(Y)$$

Hence X and Y are independent.

Example 9.39: Given the following density function of X and Y. Find the covariance of X and Y.

$$f(x,y) = \begin{cases} 3x, & 0 < y < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Solution: We have

$$E(XY) = \int_0^1\int_0^x xy(3x)dydx = \int_0^1 3x^2\left(\frac{y^2}{2}\right)_0^x dx$$

$$= \int_0^1 \frac{3}{2}x^4 dx = \frac{3}{2}\left(\frac{x^5}{5}\right)_0^1 = \frac{3}{10}$$

Again

$$E(X) = \int_0^1\int_0^x x(3x)dydx = \int_0^1 3x^3 dx = \frac{3}{4}x^4\Big|_0^1 = \frac{3}{4}$$

$$E(Y) = \int_0^1\int_0^x y(3x)dydx = \int_0^1 3x\left(\frac{y^2}{2}\Big|_0^x\right)dx = \int_0^1 \frac{3}{2}x^3 dx = \frac{3}{8}\left(x^4\right)_0^1 = \frac{3}{8}$$

Thus we obtain

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{3}{10} - \left(\frac{3}{4}\right)\left(\frac{3}{8}\right) = 0.02$$

Example 9.40: Let X and Y have a joint density function as follows:

$$f(x,y) = \begin{cases} 2x, & 0 < x < 1,\ 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Show that the covariance between X and Y is 0.

Solution: By definition

$$E(XY) = \int_0^1\int_0^1 xy(2x)f(x,y)dxdy$$

$$= \int_0^1 y\left(\frac{2x^3}{3}\Big|_0^1\right)dy = \frac{2}{3}\int_0^1 ydy$$

$$= \frac{2}{3}\left(\frac{y^2}{2}\right)_0^1 = \frac{1}{3}$$

$$E(X) = \int_0^1\int_0^1 x(2x)dxdy$$

$$= \int_0^1\left(\frac{2x^3}{3}\bigg|_0^1\right)dy = \frac{2}{3}\int_0^1 dy = \frac{2}{3}$$

$$E(Y) = \int_0^1\int_0^1 y(2x)dxdy$$

$$= \int_0^1 y\left(\frac{2x^2}{2}\bigg|_0^1\right)dy = \int_0^1 ydy = \frac{1}{2}$$

Hence

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = 0$$

The above result furnishes specific example of the general result given in the following theorem.

***Theorem* 9.24** *If X and Y are independent random variable, then*

$$\text{Cov}(X,Y) = 0$$

Proof: We have earlier shown that

$$\text{Cov}(X,Y) = E(XY) - \mu_x\mu_y$$

We have also established that, when X and Y are independent

$$E(XY) = E(X)E(Y) = \mu_x\mu_y$$

from which the desired result follows.

Example 9.41: Let X and Y be two discrete random variables with joint probability distribution as shown in the accompanying table. Show that X and Y are not independent but have zero covariance.

	x values			
y values	–1	0	1	Total
–1	1/16	3/16	1/16	5/16
0	3/16	0	3/16	6/16
1	1/16	3/16	1/16	5/16
Total	5/16	6/16	5/16	1

Solution: You can easily verify that

$$f(-1,-1) \neq g(-1)h(-1)$$
$$f(-1,0) \neq g(0)h(-1)$$
$$f(-1,1) \neq g(-1)h(1)$$

Similar results hold for other cross-products. Hence X and Y are not independent.

Also

$$E(XY) = \sum_x \sum_y xyf(x,y)$$
$$= (-1)(-1)(1/16) + (0)(-1)(3/16) + ... + (0)(1)(3/16) + (1)(1)(1/16) = 0$$

Further

$$E(X) = (-1)(5/16) + (0)(6/16) + (1)(5/16) = 0$$
$$E(Y) = (-1)(5/16) + (0)(6/16) + (1)(5/16) = 0$$

Thus,

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 - (0)(0) = 0$$

This example shows that the converse of Theorem 9.24 is not true. If the covariance of two random variables is zero, the variables need not be independent.

EXERCISES 9

1. What is mathematical expectation? What relation does it have with the mean of a frequency distribution? Why do you call it population mean?
2. Define mathematical expectation of a random variable. Can this value be negative? Justify your answer with an example.
3. What are the properties of the expected value of a random variable? Let X denote the number of spots showing on the face of a well-balanced die after it is rolled. Given that $Y=X^2+2X+3$, find $E(X)$ and $E(Y)$. [**Ans.**: 3.5, 25.17]
4. Let X be the number of heads obtained in the toss of three ideal coins. If two or more heads turn up, you win \$2, otherwise you lose \$1. What is your expected gain per game? [**Ans.**: ½]
5. Suppose that an ideal coin is tossed twice. If it turns up tails both times, you win nothing. If head turns up only once, you win \$1. If both tosses result in heads, you win \$4. Find(a) the expected number of heads to turn up (b) the expected amount of money you would win [**Ans.**: 1, 1.5]