**CHAPTER-EIGHT**

**EXPECTATION**

**Expectation of a Random Variable**

The averaging process, when applied to a random variable is called expectation. It is denoted by E(X) or μ and is read as the expected value of X or the mean value of X.

**Expectation of a Random Variable: Discrete Case**

The expected value of a discrete random variable is defined by E(x) =; provided the sum exists. Since P(x)0, and, the expected value of a discrete random variable can also be straightforwardly interpreted as a **weighted** **average** of the possible outcomes (or range elements) of the random variable. In this context the weight assigned to a particular outcome of the random variable is equal to the probability that the outcome occurs (as given by the value of P(x))

**Example 1:** Consider the random variable representing the number of episodes of diarrhea in the first 2 years of life. Suppose this random variable has a probability mass function as below

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| X | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| P(X = x) | 0.129 | 0.264 | 0.271 | 0.185 | 0.095 | 0.039 | 0.017 |

What is the expected number of episodes of diarrhea in the first 2 years of life?

**Solution:** E (X) = = 0.P(X=0) +1.P (X=1) +2P(X=2) + … + 6P(X=6)

= 0 (0.129) + 1(0.264) +2(0.271) + … + 6(0.017) = 2.038

Thus, on the average a child would be expected to have 2 episodes of diarrhea in the first 2 years of life.

**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 0.2,0 .8, and 0.3, what is the firm’s expected total profit?

**Solution**: Letting Xi, i =1, 2, 3 denote the firm’s profit from job i, then

Total profit=. So

Therefore the firm’s expected total profit is 30 thousand dollars.

**Expected Value of a Random Variable: Continuous Case**

The expected value of the continuous random variable X is defined by

E(x), provided the integral exists.

**EXAMPLE**: Suppose that X is the time (in minutes) during which electrical equipment is used at maximum load in a certain specified time period. Suppose that X is a continuous random variable with the following pdf**;**

Solution: By definition, E(x) =

=1500 minutes

**Example:** A large domestic automobile manufacturer mails out quarterly customer satisfaction surveys to owners who have purchased new automobiles within the last 3 years. The proportion of surveys returned in any given quarter is the outcome of a random variable X having density function. What is the expected proportion of surveys returned in any given quarter?

**Solution**: By definition,

E(x) ===== 0.75

The expected proportion of surveys returned in any given quarter is 0.75

**Expectations of Functions of Random Variables**

Let X be a random variable having density function f(x).Then the expectation of

Y=g(x) is given by E(Y)=

**Example**:

1. A small rural bank has two branches located in neighboring towns in East Gojjam. The numbers of certificates of deposit that are sold at the branch in **Menkorere** and the branch in **Gozamen** in any given week can be viewed as the outcome of the bivariate random variable (X,Y) having joint probability density function
2. Are the random variables independent?
3. What is the expected number of certificate sales by the **Gozamen** Branch?
4. What is the expected number of combined certificate sales for both branches?

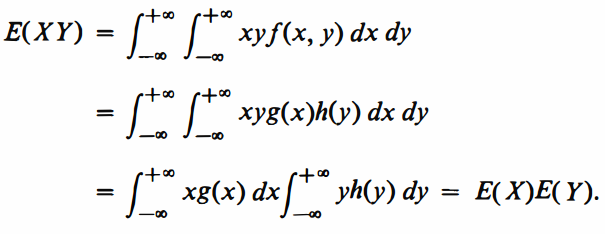
**Properties of Expectation**

If X and Y are random variables and a, b are constants then:

1. E(k) = k, where k is any constant
2. E (kX) = k E(X), where k is any constant
3. E (X + k) =E(X) + k
4. E(X + Y) = E(X) +E(Y)
5. E(X) ≥ 0, if X ≥ 0.
6. |E(X)| ≤ E(|X|)
7. |E(XY)2| ≤ E(X2) E(Y2).
8. E(XY) = E(X) E(Y), if X, Y are independent random variables

**Theorem:** Let (X, Y) be a two-dimensional random variable and suppose that X and Yare independent. Then E(XY) = E(X)E(Y).

Proof:



**Variance of random variable**

The variance of a random variable measures the variability of the random variable about its expected value.

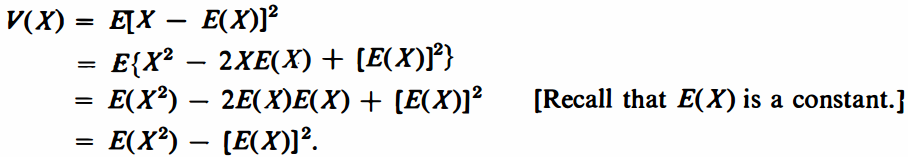
**Definition**: Let X be a random variable. We define the variance of X, denoted by V(X) or , as follows:

Mean of X = E(X)

Variance of X = =

The positive square root of V(X) is called the standard deviation of X and is denoted by .

**Proof:** Expanding and using the previously established properties for expectation, we obtain



**Case 1:** V**ariance for discrete random variable**

If X is a discrete random variable with expected value μ then the variance of X, denoted by Var (X), is defined by:

Var(X) = E(X-μ)2 = E(X2) – μ2=

Alternatively, Var(X) =

**Case 2:** V**ariance for continuous random variable**

If X is a continuous random variable, then var (X),

**Properties of Variances**

* For any random variable X and Y and constant a and b, it can be shown that
* if X1, X2 ……, Xn are independent random variables, then

i.e.,

Example: Compute the variance of f(x) = for 0 < x < 3

V(x) = E(x2) – [E(x)]2

Therefore, V(x) = E(x2) – [E(x)]2

**Moments of a Random Variable**

The expectations of certain power functions of a random variable have uses as measures of central tendency, spread or dispersion, and skewness of the density function of the random variable, and also are important components of statistical inference procedures that we will study in later chapters. These special expectations are called moments of the random variable (or of the density function). There are two types of moments that we will be concerned with moments about the origin and moments about the mean.

It tells us information about the shape of the distribution. The nature of the distribution can be identified by looking on various moment values.

**Moment about the Origin**

Let X be a random variable with density function f(x). Then the r**th** moment of X about the origin, denoted by, is defined for integers r>0 as

E() =

The value of r in the definition of moments is referred to as the order of the moment, so that one would refer to E() as the moment of order r. Note that 1 for any discrete or continuous random variable, since = E()= E(1) = 1.

The first moment about the origin is simply the expectation of the random Variable X,

i.e, = E()= E(x) a quantity that we have examined at the beginning of our discussion of mathematical expectation. This balancing point of a density function, or the weighted average of the elements in the range of the random variable, will be given a special name and symbol.

* The first moment about the origin of a random variable, X, is called the **mean of the random variable X**(or mean of the density function of X), and will be denoted by the symbol. Thus, the first moment about the origin characterizes the central tendency of a density function. Measures of spread and skewness of a density function are given by certain moments about the mean

**Moment about the Mean**

Let X be a random variable with density function f(x). Then the rth **central moment of X** (or the rth moment of X about the mean), denoted by , is defined as

) =

* Note that=1 for any discrete or continuous random variable, since

E()= E(1)=1 . Furthermore, = 0 for any discrete or continuous random variable for which E(X) exists, sinceE() =E(==0

* The second central moment is given a special name and symbol. The second central moment, E(), of a random variable, X, is called the **variance of the random variable X** (or the variance of the density function of X), and will be denoted by the symbol , or by Var(x).
* The nonnegative square root of the variance of a random variable, X, (i.e.,) is called the standard deviation of the random variable (or standard deviation of the density function of X) and will be denoted by the symbols,or by Std(X).

**The relationship between raw moment and central moment**

Exercise:

1. If X is a random variable having a probability mass function

Find the mean and variance of the random variable of X

1. If X is a random variable having a probability density functionFind the mean and variance of the random variable of X

**Moment Generating Functions**

The expectation ofresults in a function of t that, when differentiated with respect to the argument and then evaluated at ***t=0***, generates moments of X about the origin. The function is aptly called the moment-generating function of X.

**Definition**: The expected value of is defined to be the moment-generating function of X if the expected value exists for every value of t in some open interval containing 0,

i.e..The moment generating function of X will be denoted by Mx (t), and is represented by:

Mx(t) =E() =

Note that Mx(0) =E() =E(1)=1isalways defined, and from this property it is clear that a function of t cannot be a MGF unless the value of the function at t=0 is 1. The condition that Mx(t) must be defined t∈(-h, h) is a technical condition that ensures Mx(t) is differentiable at the point zero, a property whose importance will become evident shortly.

We now indicate how the MGF can be used to generate moments about the origin. In the following theorem, we use the notationto indicate the **rth** derivative of g(x) with respect to x evaluated at x=a.

**Theorem**: Let X be a Random Variable for which the MGF, Mx(t),exists. Then

=E() =

Example: Suppose that X is binomially distributed with parameters n and p. Then the moment generating function is defined as Mx(t)= then find the 1st and the 2nd moment then find the mean and variance of the binomial distribution.

Solution:

M’x(t)=

M’’x(t)=

Therefore E(x)=M’x(0)=np and E()=M”x(0)=

Var(x) =

**Theorem**: Suppose that the random variable X has MGF Mx(t). Let Y= aX +b. Then My(t), the MGF of the random variable Y, is given by; My(t)=Mx(at)

**Example**: The random variable X has the density function. Find the MGF, and use it to define the mean and variance of X.

Solution: Mx(t)= = = =

= =, provided if t<1

Exercise:

1. Suppose that X has pdf given by.
2. Determine the MGF of X.
3. Using the MGF, evaluate E(X) and Var(X).
4. Suppose that the continuous random variable X has pdf
5. Obtain the MGF of X.
6. Using the MGF, find E(X) and Var(X).
7. Suppose that the MGF of a random variable X is of the form M x(t) = [0.4
   1. What is the MGF of the random variable Y = 3X + 2 ?
   2. Evaluate E(X).
   3. Can you check your answer to (b) by some other method?
8. The daily quantity of water demanded by the population of a large Addis Ababa city in the summer months is the outcome of a random variable, X, measured in millions of gallons and having a MGF of Mx(t) =,for t<2
9. Find the mean and variance of the daily quantity of water demanded.
10. Is the density function of water quantity demanded symmetric?

**The Markov and Chebychev’s inequalities**: It is often not possible to calculate exactly the probabilities associated with a random variable, and we will often look for bounds on these probabilities. Two of the most famous bounds (or inequalities) are Markov’s inequality and Chebychev’s inequality.

**Markov’s inequality:** Let *X* be a non-negative random variable and let ***a*** be any positive constant. Then

We shall prove this result in the continuous case. For a continuous random variable with probability density function, *f*

**Chebychev’s inequality:** Let *X* be a random variable with mean μ and variance σ2 and let *k* be any positive constant. Then indicates at least

(100%are found in the interval (andindicates at most 100%are found out of the interval (.**Where**

is the probability that the value of X lies at least k standard deviations from its mean is at most  where k is a positive number greater than 1.

Note:

Example:

1. a random variable X has the mean of 4 and variance of X is 2, then use chebyshev’s inequality obtain the upper bound for

Solution: By the definition of the chebyshev’s inequality

is the upper bound

But

,

So the probability that some value of the random variable X will be between 1 and 7 is at least 0.778

1. A random variable X has pdf, then use chebyshev’s inequality to estimate

Solution: By the definition of the chebyshev’s inequality

is the upper bound

Then

But

Then,

So the probability that some value of the random variable X will be between 0.5 and 4.5 is at least 0.8125

**Example:** Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.

* 1. What can be said about the probability that this week’s production will exceed 75?
  2. If the variance of a week’s production is known to equal 25, then what can be said about the probability that this week’s production will be between 40 and 60?

**Solution:** Let Xbe the number of items that will be produced in a week:

1. By Markov’s inequality;
2. By Chebyshev’s inequality;

Hence

So the probability that this week’s production will be between 40 and 60 is at least 0.75.

**Exercise:** From past experience, a professor knows that the test score of a student taking her final examination is a random variable with mean 75.

1. Give an upper bound to the probability that a student’s test score will exceed 85. Suppose in addition the professor knows that the variance of a student’s test score is equal to 25.
2. What can be said about the probability that a student will score between 65 and 85?
3. How many students would have to take the examination so as to ensure, with probability at least 0.9, that the class average would be within 5 of 75?

**Joint Moment about the Origin**

Let X and Y be two random variables having joint density function f(x,y). Then the joint moment of (X,Y) (or off(x,y)) about the origin is defined by

=

**Covariance and Correlation coefficient**

Regarding joint moments, our immediate interest is on a particular joint moment about the mean,, and the relationship between this moment and moments about the origin. The central momentis given a special name and symbol, and we will see thatis useful as a measure of “linear association” between X and Y.

**Covariance**: The central joint moment = is called the covariance betweenXandY, and is denoted by the symbol, or by.

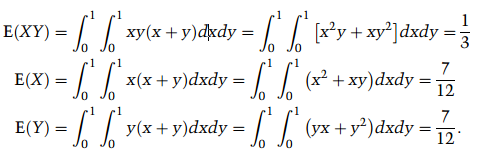
Note that there is a simple relationship between and moments about the origin that can be used for the calculation of the covariance.

Proof: This result follows directly from the properties of the expectation operation. In particular, by definition

Example: Let the bivariate random variable (X,Y) have a joint density function

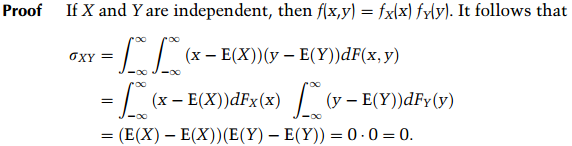
Find.

**Solution**: from the definition of covariance it is computed as



Therefore by the definition of covariance

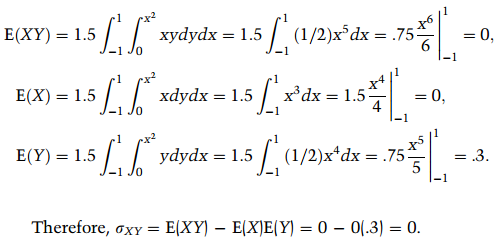
* If X and Y are independent random variable then, (assuming the covariance exists)otherwise X and Y are dependent variable.



Example: Let X and Y be two random variables having a joint density function given by

.

Note this density implies that (x,y) points are equally likely to occur on and below the parabola represented by the graph of. There is a direct functional dependence between Xand the range ofY, so thatwill change as x changes and thus X and Y must be dependent random variables. Nonetheless;. To see this, note that



So that the X and Y are independent random variable

**Properties covariance:**

If X and Y are either continuos or discrete random variable and if a and b are any constant number, then covariance has the following properties.

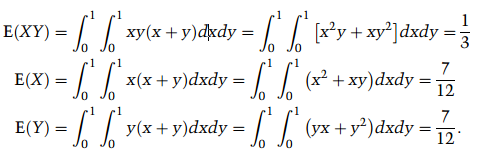
**Correlation coefficient:** The correlation coefficient between two random variables X and Y is defined by

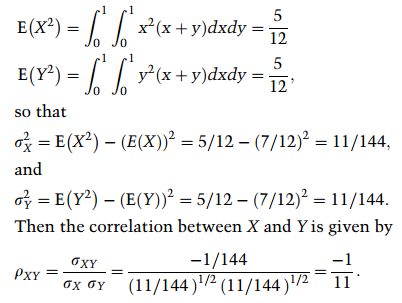
* Correlation coefficient tells us the degree of association and the direction of the linear relationship between the random variables.
* The correlation coefficient computed from the sample data measures the strength and direction of a linear relationship between two variables.
* The symbol for the sample correlation coefficient is r.
* The symbol for the population correlation coefficient is ρ
* The range of the correlation coefficient is from −1 to +1.
* If there is a strong positive linear relationship between the variables, the value of r will be close to +1.
* If there is a strong negative linear relationship between the variables, the value of r will be close to −1.
* When there is no linear relationship between the variables or only a weak relationship, the value of r will be close to 0.

Example: Let the bivariate random variable (X,Y) have a joint density function

, then compute the correlation coefficient of X and Y?

**Solution:**





**Properties of correlation:** Let *X* and *Y* be random variables with correlation equal to ρ.Then-1≤ ρ ≤1. Furthermore, ρ equals 1 or -1 if and only if *Y* is a linear function of *X*. In fact:

1. If *Y* = *a* +*bX* for some constants *a* and *b* then ρ=1 if *b*>0, and ρ=-1 if *b*<0.
2. If *Y* ≠ *a* +*bX* for all *a* and *b* then -1< ρ<1.

If *X* and *Y* are independent then ρ=0

Exercise:

* 1. A Seattle newspaper intends to administer two different surveys relating to two different anti-tax initiatives on the ballot in November. The proportion of surveys mailed that will actually be completed and returned to the newspaper can be represented as the outcome of a bivariate random variable (X,Y) having the density function

Where X is the proportion of surveys relating to initiative It hat are returned, and Y refers to the proportion of surveys relating to initiative II that are returned. Compute the **covariance** and **correlation coefficient** between the X and Y and also interpret the values of **correlation coefficient** between the X and Y