Unit – II
Operations on One Random Variable

Operations On Single & Multiple Random Variables – Expectations: Expected Value of a Random Variable, Function of a Random Variable, Moments about the Origin, Central Moments, Variance and Skew, Chebychev's Inequality, Characteristic Function, Moment Generating Function, Transformations of a Random Variable: Monotonic and Non-monotonic Transformations of Continuous Random Variable, Transformation of a Discrete Random Variable. Vector Random Variables, Joint Distribution Function and its Properties, Marginal Distribution Functions, Conditional Distribution and Density – Point Conditioning, Conditional Distribution and Density – Interval conditioning, Statistical Independence. Sum of Two Random Variables, Sum of Several Random Variables, Central Limit Theorem, (Proof not expected). Unequal Distribution, Equal Distributions. Expected Value of a Function of Random Variables: Joint Moments about the Origin, Joint Central Moments, Joint Characteristic Functions, Jointly Gaussian Random Variables: Two Random Variables case, N Random Variable case, Properties, Transformations of Multiple Random Variables, Linear Transformations of Gaussian Random Variables.

Introduction:

Concept of Random variable was introduced in the Unit-2 as a means of providing a systematic definition of events defined on a sample space. Specifically, it formed a mathematical model for describing characteristics of some real, physical world random phenomenon. In this unit some important operations that may be performed on a random variable are discussed. Most of these operations are based on a single concept – Expectation.

Expectation:

Definition:

Expectation is the name given to the process of averaging when a random variable is involved.
For a Random Variable X, the notation used to represent expectation of X is E(X). It can be said as the mathematical expectation of X (or) the expected value of X (or) the mean value of X (or) the statistical average of X. Another notation of E(X) is $\bar{X}$.

Expected value of a Random Variable:

In general, the expected value of a given random variable X is defined by

$$E[X] = \bar{X} = \int_{-\infty}^{\infty} x\, f_X(x)dx \quad \text{---- (3.1)}$$

If X happens to be discrete with N possible values xi having probabilities P(xi) of occurrence, then

$$f_X(x) = \sum_{i=1}^{N} P(x_i)\delta(x - x_i) \quad \text{------ (3.2)}$$

Upon substitution of eq (3.2) in eq(3.1), the expectation for discrete random variable can be

$$E[X] = \sum_{i=1}^{N} x_i\, P(x_i)$$

Example: Determine the mean value of the continuous, exponentially distributed random variable.

For exponentially distributed random variable, the density function is given by

$$f_X(x) = \begin{cases} \frac{1}{b}\, e^{-\frac{(x-a)}{b}} & ; x > a \\ 0 & ; x < a \end{cases}$$

Now, the expected value of the above function can be calculated using eq(3.1)

$$E[X] = \int_a^\infty \frac{x}{b}\, e^{\frac{-(x-a)}{b}}\, dx = \frac{e^{\frac{a}{b}}}{b} \int_a^\infty x\, e^{-\frac{x}{b}}\, dx = a + b$$

Expected value of a function of a Random Variable:

The Expected value of any given function says g(x) can be given by

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

In case of discrete, it is given by

$$E[g(x)] = \sum_{i=1}^{N} g(x_i) P(x_i)$$

Where N may be infinite for some random variables.

Note that if g(X) is a sum of N functions gn(X), n = 1,2,........, N, then the expected value of the sum of N functions of a random variable X is the sum of the N expected values of the individual functions of the random variable.

Example: A particular random voltage is represented by a Rayleigh random variable V having a density function given by $f_X(x) = \begin{cases} \frac{2}{b}\,(x-a) e^{-\frac{(x-a)^2}{b}} & ; x \geq a \\ 0 & ; x < a \end{cases}$, with a=0 and b = 5. The voltage is applied to a device that generates a voltage Y = g(V) = V^2 that is equal numerically to power in V. Find the average power in V.

$$power\ in\ V = E[g(V)] = E[V^2] = \int_0^\infty \frac{2v^3}{5}\ e^{-v^2/5}\, dv$$

Let $v^2 / 5 = \xi$, $d\xi = 2v\, dv/5$, we obtain

$$power\ in\ V = 5 \int_0^\infty \xi e^{-\xi}\, d\xi = 5\ W$$

Conditional Expected Value:

If in eq(1), $f_X(x)$ is replaced by the conditional density $f_X(x \backslash B)$, where B is any event defined on the sample space, we have the conditional expected value of X, denoted E[X\B]

$$E[X \backslash B] = \int_{-\infty}^{\infty} x\, f_X(x \backslash B) dx$$

If B is defined as an Event B = {x ≤ b}, -∞ < b < ∞, as discussed in unit – II and using the concepts discussed relating to it the Expected value can be given by

$$E[X\backslash X \leq b] = \frac{\int_{-\infty}^{b} x f_X(x)dx}{\int_{-\infty}^{b} f_X(x)dx}$$

Moments:

An immediate application of the expected value of a function g(x) of a random variable X is in calculating moments. Two types of moments are of interest are

1. Moments about origin
2. Moments about mean – central moments

Moments about Origin:

The function $g(X) = X^n$; n = 0, 1, 2... When used in $E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$ gives the moments about the origin of the random variable X. Denote the nth moment by m_n. Then

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

Clearly when n=0, $m_0 = 1$, which gives the area of the function $f_X(x)$

when n=1, $m_1 = E[X^1] = \int_{-\infty}^{\infty} X^1 f_X(x)dx = E[X] = \bar{X}$, the expected value of X.

Central Moments:

Moments about the mean value of X are called central moments and are given the symbol μ_n. They are defined as the expected value of the function

$$g(X) = (X - \bar{X})^n \quad ; n = 0,1,2......$$

which is

$$\mu_n = E[(X - \bar{X})^n] = \int_{-\infty}^{\infty} (X - \bar{X})^n f_X(x)dx$$

When n = 0, $\mu_0 = 1$, gives the area of $f_X(x)$

when n =1, $\mu_1 = 0$.

Variance:

The second central Moment μ_2 is so important and it is given the name **Variance** and the special notation is σ_X^2. Thus variance is given by

$$\sigma_X^2 = \mu_2 = E[(X - \bar{X})^2] = \int_{-\infty}^{\infty} (x - \bar{X})^2 f_X(x)dx \text{ ---- (A)}$$

The positive square root σ_X of variance is called the standard deviation of X; It is a measure of the spread in the function $f_X(x)$ about the mean.

Variance can be found from a knowledge of first and second moments. By expanding eq(A) we have

$$\sigma_X^2 = E[X^2 - 2X\bar{X} + \bar{X}^2] = E[X^2] - 2\bar{X}E[X] + \bar{X}^2$$

$$= E[X^2] - \bar{X}^2 = m_2 - {m_1}^2$$

Skew:

The third central moment $\mu_3 = E[(X - \bar{X})^3]$ is a measure of the asymmetry of $f_X(x)$ about $x = \bar{X} = m_1$. It is called the **Skew** of the density function.

If a density is symmetric about $x = \bar{X}$, it has zero skew. The normalized third central moment μ_3 / σ_X^3 is known as the **Skewness** of the density function or alternatively as the coefficient of skewness.

Problem: Let X have the exponential density function calculate the variance of X.

As X have exponential density function, it is given as

$$f_X(x) = \begin{Bmatrix} \frac{1}{b}\, e^{-\frac{(x-a)}{b}} & ; x > a \\ 0 & ; x < a \end{Bmatrix}$$

By substituting $f_X(x)$ in the expression for variance and evaluate, we have

$$\sigma_X^2 = \int_a^\infty (x - \bar{X})^2 \frac{1}{b}\, e^{-(x-a)/b}\, dx$$

On solving the above equation, we obtain, $\sigma_X^2 = (a + b - \bar{X})^2 + b^2$

From the previous problem, we know $a+b = \bar{X}$, hence $\sigma_X^2 = b^2$

Problem: Compute Skew and coefficient of skewness for the exponential density function.

For n =3, we have

$$\mu_3 = E[(X - \bar{X})^3] = E[X^3 - 3\bar{X}X^2 + 3\bar{X}^2 X - \bar{X}^3]$$

$$\bar{X}^3 - 3\bar{X}\bar{X}^2 + 2\bar{X}^3 = \overline{X^3} - 3\bar{X}(\sigma_X^2 + \bar{X}^2) + 2\bar{X}^3 = \overline{X^3} - 3\bar{X}\sigma_X^2 - \bar{X}^3$$

Next we have, $\overline{X^3} = \int_a^\infty \frac{x^3}{b}\, e^{-\frac{x-a}{b}}\, dx = a^3 + 3a^2b + 6ab^2 + 6b^3$ (after evaluating using exponential functions)

On substituting $\bar{X} = a+b$ and $\sigma_X^2 = b^2$ and reducing the algebra, we find

Skewness = $\mu_3 = 2b^3$

and coefficient of skewness = $\mu_3 / \sigma_X^3 = 2$

Chebyshev's Inequality:

A useful tool in some probability problems is Chebyshev's in equality. For a random variable X with mean $\bar{X}$ and varianceσ_X^2, it states that

$$P\{|X - \bar{X}| \geq \epsilon\} \leq \sigma_X^2 / \epsilon^2 \quad \text{for any } \epsilon > 0$$

This expression can be demonstrated by integration of the probability density using $F_X(x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi$

$$P\{|X-\bar{X}|\ \geq \epsilon\} = \int_{-\infty}^{\bar{X}-\epsilon} f_X(x)dx + \int_{\bar{X}+\epsilon}^{\infty} f_X(x)dx = \int_{|X-\bar{X}|\,\geq\,\epsilon}^{\infty} f_X(x)dx$$

But since

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{X})^2\ f_X(x)dx\ \geq \int_{|X-\bar{X}|\,\geq\,\epsilon}^{\infty} (x-\bar{X})^2 f_X(x)dx$$

$$\geq \epsilon^2\ \int_{|X-\bar{X}|\,\geq\,\epsilon}^{\infty} f_X(x)dx = \epsilon^2\ P\{|X-\bar{X}|\ \geq \epsilon\}$$

must be true. Alternative form of Chebyshev's inequality is

$$P\{|X-\bar{X}| < \epsilon\}\ \geq 1-(\sigma_X^2\ /\epsilon^2)$$

Functions that give Moments:

We discussed about moments and its types in previous section. Here we are discussing two functions that allow moments to be calculated for a random variable X. They are

1. Characteristic function
2. Moment generating function

Characteristic function:

The characteristic function of a random variable X is defined by

$$\emptyset_X(\omega) = E[e^{j\omega x}]\ \text{--- (1)}$$

Where j =sqrt(-1). It is a function of the real variable $-\infty < \omega < \infty$.

If Characteristic function $\emptyset_X(\omega)$ is written in terms of density function, it seen to be the Fourier transform of $f_X(x)$

$$\emptyset_X(\omega) = \int_{-\infty}^{\infty} f_X(x)e^{j\omega x}\,dx\ \text{--- (2)}$$

Because of this fact, if $\emptyset_X(\omega)$ is known, $f_X(x)$ can be found from the inverse Fourier Transform.

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \emptyset_X(\omega)e^{-j\omega x}\,d\omega\ \text{--- (3)}$$

Hence it is observed that $\emptyset_X(\omega)$ and $f_X(x)$forms Fourier transform Pair.

By formal differentiation of eq(2) n times with respect to ω and setting ω = 0in the derivative, it can be shown that the nth moment of X is given by

$$m_n = (-j)\frac{d^n\emptyset_X(\omega)}{d\omega^n}\ at\ \omega = 0$$

A major advantage of using characteristic function to find moments is that $\emptyset_X(\omega)$always exists. So the moments can always be found if $\emptyset_X(\omega)$ is known, provided of course both the moments and derivatives of $\emptyset_X(\omega)$ exist.

It can be shown that the maximum magnitude of a characteristic is unity and occurs at ω that is

$$|\emptyset_X(\omega)| \leq \emptyset_X(0) = 1$$

Example: Consider the random variable with the exponential density and hence find its characteristic function and first moment using characteristic function.

Exponential density function is given by $f_X(x) = \begin{Bmatrix} \frac{1}{b}e^{-\frac{(x-a)}{b}} & ; x > a \\ 0 & ; x < a \end{Bmatrix}$

and it is known that if density function is known, characteristic function can be calculated by taking Fourier transform of density function. Hence,

$$\emptyset_X(\omega) = \int_a^{\infty} \frac{1}{b} e^{-\frac{(x-a)}{b}} e^{j\omega x}\, dx = \frac{e^{a/b}}{b} \int_a^{\infty} e^{-\left(\frac{1}{b-j\omega}\right)x}\, dx$$

By evaluating the integral, we can calculate characteristic function

$$\emptyset_X(\omega) = \frac{e^{\frac{a}{b}}}{b}\left[\frac{e^{-\left(\frac{1}{b-j\omega}\right)x}}{-\left(\frac{1}{b-j\omega}\right)}\right]_a^{\infty} = \frac{e^{j\omega a}}{1-j\omega b}$$

To calculate first moment, calculate the derivative of $\emptyset_X(\omega)$,

$$\frac{d\emptyset_X(\omega)}{d\omega} = e^{j\omega a}\left[\frac{ja}{1-j\omega b} + \frac{jb}{(1-j\omega b)^2}\right]$$

and hence first moment is given by

$$m_1 = (-j)\frac{d^1\emptyset_X(\omega)}{d\omega^1}\ at\ \omega = 0 = a + b$$

Moment Generating Function:

Another statistical average closely related to the characteristic function is the moment generating function, defined by

$$M_X(v) = E[e^{vx}]$$

where v is a real number -∞ < v < ∞. Thus $M_X(v)$ is given by

$$M_X(v) = \int_{-\infty}^{\infty} f_X(x) e^{vx}\, dx$$

The main advantage of the moment generating function derives from its ability to give moments. Moments are related to $M_X(v)$ by the expression

$$m_n = \frac{d^n M_X(v)}{dv^n}\ at\ v = 0$$

The main disadvantage of the moment generating function as opposed to the characteristic function is that it may not exist for all random variables and values of v.

Example: Consider the random variable with the exponential density and hence find its characteristic function and first moment using moment generating function

Exponential density function is given by $f_X(x) = \begin{cases} \frac{1}{b}\, e^{-\frac{(x-a)}{b}} & ; x > a \\ 0 & ; x < a \end{cases}$

on using moment generating function

$$M_X(v) = \int_a^{\infty} \frac{1}{b}\, e^{-\frac{(x-a)}{b}} e^{jvx}\, dx = \frac{e^{\frac{a}{b}}}{b} \int_a^{\infty} e^{\left(v-\left(\frac{1}{b}\right)\right)x}\, dx = \frac{e^{av}}{1-bv}$$

To calculate first moment, calculate the derivative of $M_X(v)$, and hence first moment is given by

$$m_1 = \frac{dM_X(v)}{dv}\; at\; v = 0 = \frac{e^{av}\,[a(1-bv)+b]}{(1-bv)^2} at\; v = 0 = a + b$$

Transformations of a Random Variable:

In some cases, one may wish to transform one random variable X into a new random variable Y by means of a transformation

Y = T(X)

If density function $f_X(x)$ or distribution function $F_X(x)$ of X is known, the problem is to determine either density function $f_Y(y)$ or distribution function $F_Y(y)$ of random variable Y. Then we can use transformation.

The problem can be viewed as a black box with input X, output Y and transfer function Y = T(X) is illustrate0d in figure below.

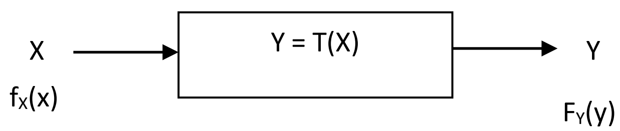


Fig 1: Transformation of a random variable X to a new random variable Y

In general, X can be a discrete, continuous or mixed random variable. In turn, the transformation T can be linear, non-linear segmented, staircase etc. Clearly, there are many cases to consider in a general study, depending on the form of X and T. Here we will consider three cases.

1. X continuous and T continuous and either monotonically increasing or decreasing
2. X continuous and T continuous but nonmonotonic
3. X discrete and T continuous

Monotonic Transformation of a continuous Random variable:

A transformation is called monotonically increasing if $T(x_1) < T(x_2)$ for any $x_1 < x_2$. It is monotonically decreasing if $T(x_1) > T(x_2)$ for any $x_1 < x_2$

Case 1:

Consider first the increasing transformation. We assume T is continuous and differentiable at all values of x for which $f_X(x) \neq 0$. Let Y have a particular value Y_o corresponding to the particular value X_o of X as shown in the figure below.

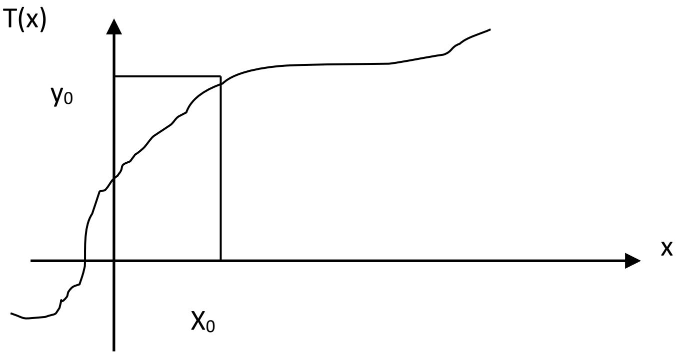


Fig 2: Monotonic Transformation and Increasing

The two numbers are related by $y_0 = T(x_0)$ or $x_0 = T^{-1}(y_0)$

Where T^{-1} represents the inverse of the transformation T. Now the probability of the event $\{Y \leq y_0\}$ must equal to the probability of the event $\{X \leq x_0\}$ because of one to one correspondence between X and Y. Thus

$$F_Y(y_0) = P\{Y \leq y_0\} = P\{X \leq x_0\} = F_X(x_0)$$

$$\text{or} \quad \int_{-\infty}^{y_0} f_Y(y)dy = \int_{-\infty}^{x_0=T^{-1}(y_0)} f_X(x)dx$$

Next we differentiate both sides of the above equation with respect to y_0 using Leibniz's rule we get,

$$f_Y(y_0) = f_X\left[T^{-1}(y_0)\right] \frac{dT^{-1}(y_0)}{dy_0}$$

Since the result applies for any y_0, we may now drop the subscript and write

$$f_Y(y) = f_X\left[T^{-1}(y)\right] \frac{dT^{-1}(y)}{dy}$$

or more compactly

$$f_Y(y) = f_X(x)\frac{dx}{dy}$$

Case 2:

Consider the second case where the function T is continuous and monotonically decreasing

$$F_Y(y_0) = P\{Y \leq y_0\} = P\{X \geq x_0\} = 1 - F_X(x_0)$$

The following figure represents such function

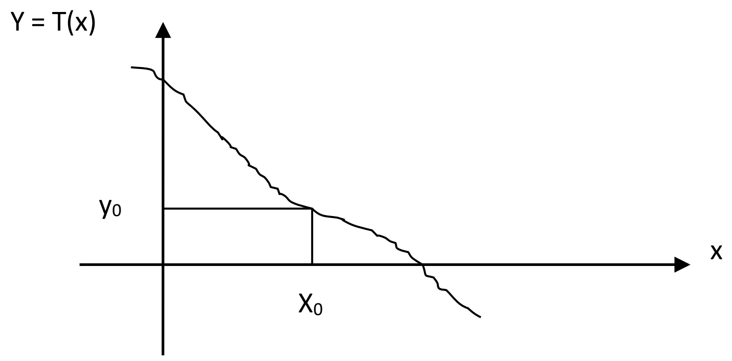


Fig 3: Monotonic Transformation and decreasing

Repeat the same as for monotonic increasing except that the right side is negative. However since the slope $T^{-1}(y)$ is also negative, it can be concluded that for either type of monotonic transformation

$$f_Y(y) = f_X\,[T^{-1}(y)]\,\frac{dT^{-1}(y)}{dy}$$

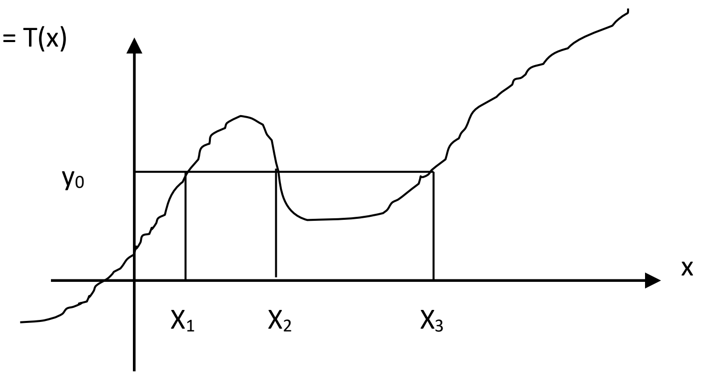
or simply

$$f_Y(y) = f_X\,(x)\frac{dx}{dy}$$

Note: Linear transformation of a Gaussian random variable produces another Gaussian random variable.

Nonmonotonic Transformations of a continuous Random variable:

A transformation may not be monotonic in the more general case. The following figure illustrates such example.



There may now be more than one interval of values of X that correspond to the event $\{Y \le y_0\}$. For the value of y_0 shown in the figure, the event $\{Y \le y_0\}$ corresponds to the event $\{X \le x_1$ and $x_2 \le x \le x_3\}$. Thus the probability of the event $\{Y \le y_0\}$ now equals the probability of the event { x values yielding $Y \le y_0$ }, which can be written as $\{x \backslash Y \le y_0\}$. In other words

$$F_Y(y_0) = P\{Y \le y_0\} = P\{x \backslash Y \le y_0\} = \int_{\{x \backslash Y \le y_0\}} f_X(x)dX$$

Formally, one may differentiate to obtain the density function of Y

$$f_Y(y_0) = \frac{d}{dy} \int_{\{x \backslash Y \le y_0\}} f_X(x)dX$$

Transformation of a Discrete Random Variable:

If X is discrete random variable while Y = T(X) is a continuous transformation, the problem is especially simple. Here

$$f_X(x) = \sum_n P(x_n)\delta(x - x_n)$$

$$F_X(x) = \sum_n P(x_n)u(x - x_n)$$

Where the sum is taken to include all the possible values x_n, n =1,2,...... of X.

If the transformation is monotonic, there is a one to one correspondence between X and Y so that a set $\{y_n\}$ corresponds to the set $\{x_n\}$ through the equation $y_n = T(x_n)$. The probability $P(y_n)$ equals $P(x_n)$. Thus

$$f_Y(y) = \sum_n P(y_n)\delta(y - y_n)$$

$$F_Y(y) = \sum_n P(y_n)u(y - y_n)$$

Where $y_n = T(x_n)$ and $P(y_n) = P(x_n)$.

If T is not monotonic, the above procedure remains valid except there now exists the possibility that more than one value x_n corresponds to a value y_n. In such a case $P(y_n)$ will equal the sum of the probabilities of the various x_n for which $y_n = T(x_n)$