

**Electronics & Communication Eng. Department, School of  
Electrical Engineering & Computing, Adama Science &  
Technology University, Adama, Ethiopia**



**Subject: Probability and Random Process**  
**( ECE 3103 )**

**Chapter 4**  
**Expectation**

**BY**

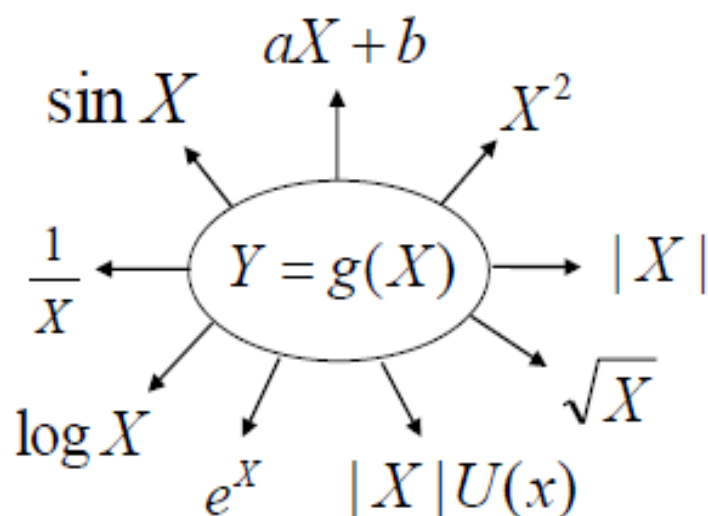
**Dr. Ram Sewak Singh**

# Functions of a Random Variable

- Let  $X$  be a continuous random variable with pdf  $f_X(\mathbf{x})$  and suppose  $g(\mathbf{x})$  is a function of the random variable  $X$  defined as:

$$Y = g(X)$$

- We can determine the cdf and pdf of  $Y$  in terms of that of  $X$ .
- Consider some of the following functions.



## Functions of a Random Variable Cont'd.....

- Steps to determine  $f_Y(y)$  from  $f_X(x)$ :

### Method I:

1. Sketch the graph of  $Y=g(X)$  and determine the range space of  $Y$ .
2. Determine the cdf of  $Y$  using the following basic approach.

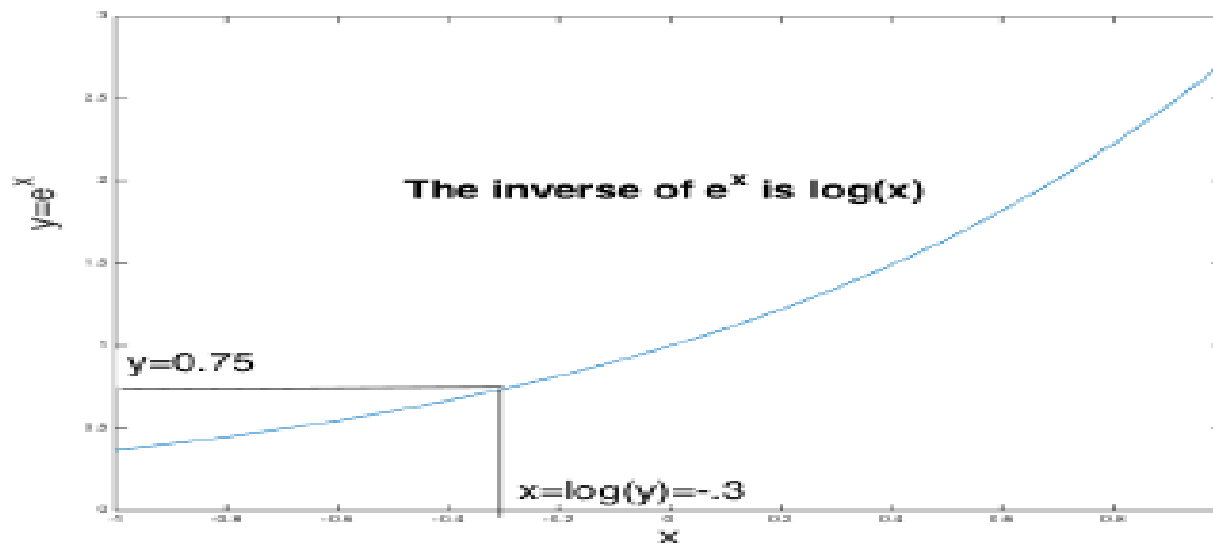
$$F_Y(y) = P(g(X) \leq y) = P(Y \leq y)$$

3. Obtain  $f_Y(y)$  from  $F_Y(y)$  by using direct differentiation, i.e.,

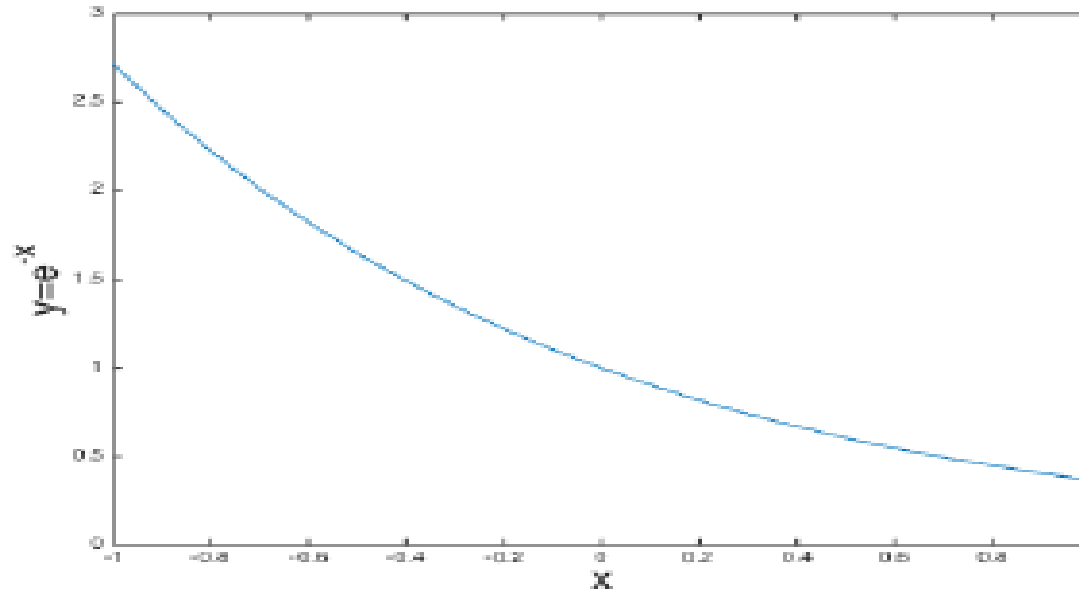
$$f_Y(y) = \frac{dF_Y(y)}{dy}$$

# Transformation of Random Variables

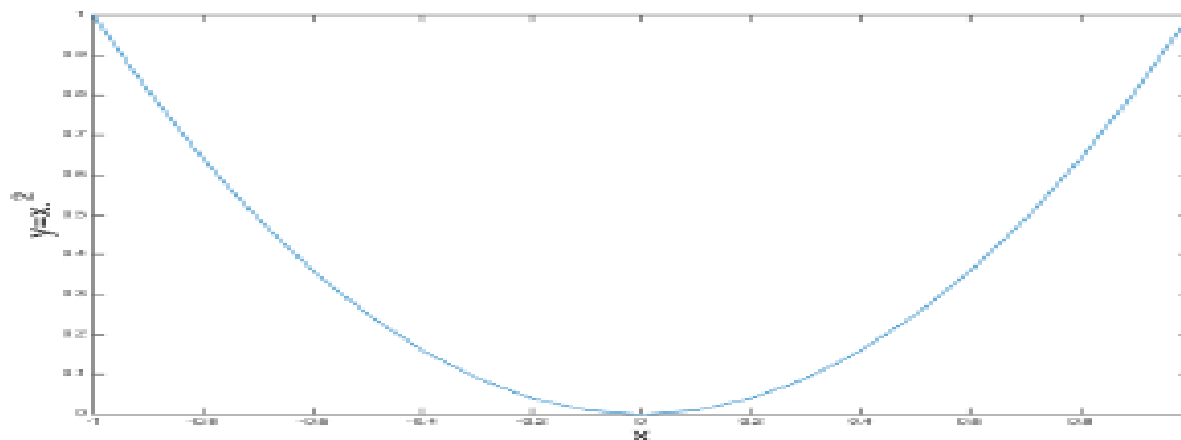
- ▶ Another special case is where  $g$  is a strictly monotonic function of  $X$ .
- ▶ Let  $X$  be a continuous random variable whose PDF is non-zero only in some range  $I$ . A function  $g(X)$  is said to be strictly monotonic over  $I$  if it is either:
  - ▶ **Monotonically increasing:** if  $x < x'$  then  $g(x) < g(x')$  for all  $x, x' \in I$ .



- **Monotonically decreasing:** if  $x < x'$  then  $g(x) > g(x')$  for all  $x, x' \in I$ .



Non monotonic Functions:  $Y = X^2$



# Monotonically increasing

- ▶ We have random variable  $X \sim f_X$
- ▶ We have  $Y = e^X$ . What is  $f_Y(y)$ ?
- ▶ Start with the CDF .

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln y) = F_X(\ln y)$$

▶

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(\ln y)}{dy} = \frac{1}{y} f_X(\ln y)$$

# Functions of a Random Variable Cont'd.....

## Method II:

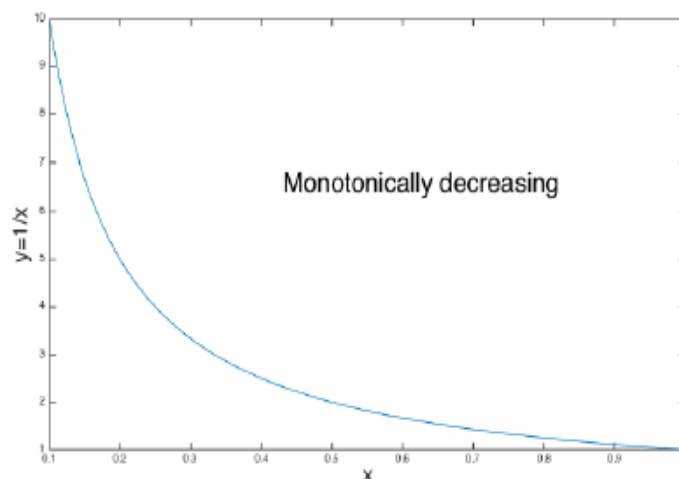
1. Sketch the graph of  $Y=g(X)$  and determine the range space of  $Y$ .
2. If  $Y=g(X)$  is one to one function and has an inverse transformation  $x=g^{-1}(y)=h(y)$ , then the pdf of  $Y$  is given by:

$$f_Y(y) = \left| \frac{dx}{dy} \right| f_X(x) = \left| \frac{dh(y)}{dy} \right| f_X[h(y)]$$

3. Obtain  $Y=g(x)$  is not one-to-one function, then the pdf of  $Y$  can be obtained as follows.
  - i. Find the real roots of the function  $Y=g(x)$  and denote them by  $x_i$

# Example: Monotonically decreasing function

- ▶ Let  $X$  be a uniform random variable on the interval  $(0, 1]$ , and let  $Y = g(X) = 1/X$ .
- ▶ So  $Y$  takes values in  $[1, \infty)$



- ▶ What is its inverse?  $h(y) = 1/y$
- ▶  $f_Y(y) = f_X(h(y)) \left| \frac{dh(y)}{dy} \right| \dots$  so we need to know  $f_X(h(y))$  and  $\left| \frac{dh(y)}{dy} \right|$ .
- ▶ So for  $y \geq 1$ ,  $f_X(h(y)) = f_X(1/y) = 1$ .
- ▶  $\left| \frac{dh(y)}{dy} \right| = \left| -\frac{1}{y^2} \right| = \frac{1}{y^2}$

### General functions of random variables

suppose  $Y = g(X)$  is a transformation (could be many-to-one)

to find  $f_Y(y)$  for a fixed  $y$ , we solve the equation  $y = g(x)$  to get  $x$

this may have multiple roots:

$$y = g(x_1) = g(x_2) = \dots = g(x_n) = \dots$$

it can be shown that

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \dots + \frac{f_X(x_n)}{|g'(x_n)|} + \dots$$

where  $g'(x)$  is the derivative (Jacobian) of  $g(x)$

## Functions of a Random Variable Cont'd.....

- ii. Determine the derivative of function  $g(x_i)$  at every real root  $x_i$ ,  
i.e. ,

$$g(x_i) = \frac{dx_i}{dy}$$

- iii. Find the pdf of Y by using the following formula.

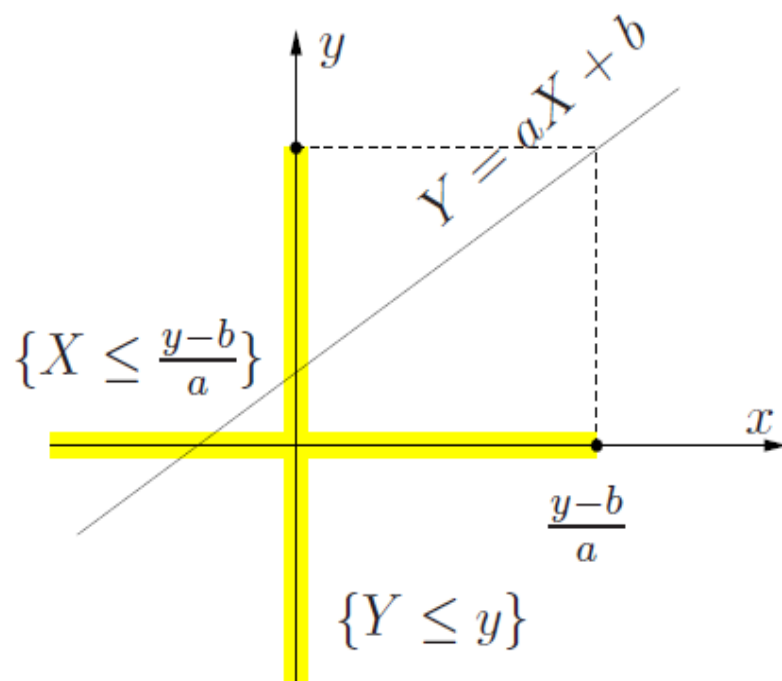
$$f_Y(y) = \sum_i \left| \frac{dx_i}{dy} \right| f_X(x_i) = \sum_i g(x_i) f_X(x_i)$$

**Example 1:**

define  $Y = aX + b$ ,  $a > 0$ . Find CDF and PDF of  $Y$

If  $a > 0$

$$dy = a \, dx$$



$$\begin{aligned} P(Y \leq y) &= P(aX + b \leq y) \\ &= P(X \leq (y - b)/a) \end{aligned}$$

thus,

$$F_Y(y) = F_X\left(\frac{y - b}{a}\right)$$

PDF of  $Y$  is obtained by differentiating the CDF wrt. to  $y$

$$f_Y(y) = \frac{dF_Y}{dy}$$

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y - b}{a}\right)$$

## Example 2:

let  $X \sim \mathcal{N}(m, \sigma^2)$  :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x - m)^2}{2\sigma^2}$$

let  $Y = aX + b$ , with  $a > 0$

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y - b}{a}\right) = \frac{1}{\sqrt{2\pi(a\sigma)^2}} \exp - \frac{(y - b - am)^2}{2(a\sigma)^2}$$

- $Y$  has also a Gaussian distribution with mean  $b + am$  and variance  $(a\sigma)^2$
- thus, a linear function of a Gaussian is also a Gaussian

### Example :3

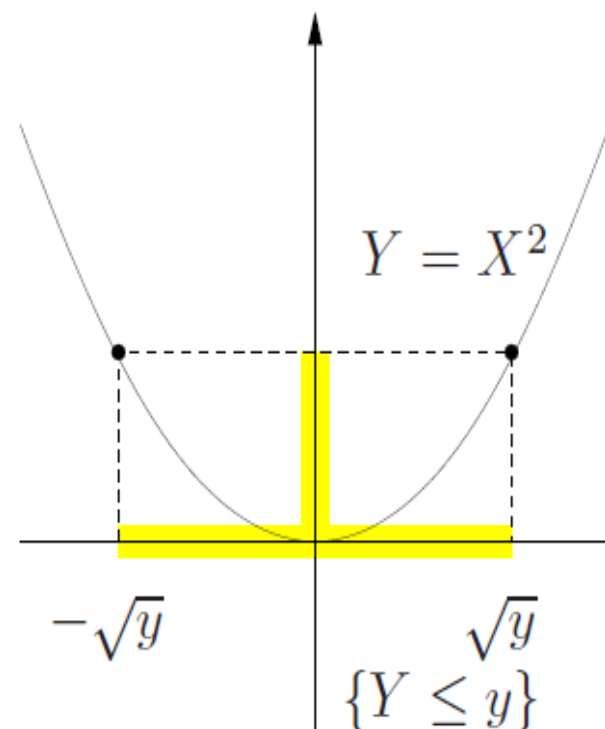
define  $Y = X^2$ . find CDF and PDF of  $Y$

for a positive  $y$ , we have

$$\{Y \leq y\} \iff \{-\sqrt{y} \leq X \leq \sqrt{y}\}$$

thus,

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y > 0 \end{cases}$$



differentiating wrt. to  $y$  gives

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

## BY USING GENERAL FUNCTIONS OF RANDOM VARIABLES

**Affine transformation:**  $Y = aX + b$ ,  $g'(x) = a$

the equation  $y = ax + b$  has a single solution  $x = (y - b)/a$  for every  $y$ , so

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y - b}{a}\right)$$

**Quadratic transformation:**  $Y = aX^2$ ,  $a > 0$ ,  $g'(x) = 2ax$

if  $y \leq 0$ , then the equation  $y = ax^2$  has no real solutions, so  $f_Y(y) = 0$

if  $y > 0$ , then it has two solutions

$$x_1 = \sqrt{y/a}, \quad x_2 = -\sqrt{y/a}$$

and therefore

$$f_Y(y) = \frac{1}{2\sqrt{ay}} \left( f_X(\sqrt{y/a}) + f_X(-\sqrt{y/a}) \right)$$

### Example:

verify that if  $X$  has a standard uniform distribution  $\mathcal{U}(0, 1)$ , then

$$Y = -\ln(X)/\lambda$$

has an exponential distribution with parameter  $\lambda$

for  $Y = y$ , we can solve  $X = x = e^{-\lambda y} \Rightarrow$  unique root

- the Jacobian is  $g'(x) = -1/\lambda x = -e^{\lambda y}/\lambda$
- $f_Y(y) = 0$  when  $x = e^{-\lambda y} \notin [0, 1]$  or when  $y < 0$
- when  $y \geq 0$  (or  $e^{-\lambda y} \in [0, 1]$ ), we will have

$$f_Y(y) = \frac{f_X(e^{-\lambda y})}{|-1/\lambda x|} = \lambda e^{-\lambda y}$$

## Example

let  $Y = \cos X$  where  $X \sim \mathcal{U}(0, 2\pi]$ , find the pdf of  $Y$

for  $|y| > 1$  there is no solution of  $x \Rightarrow f_Y(y) = 0$

for  $|y| < 1$  the equation  $y = \cos x$  has two solutions:

$$x_1 = \cos^{-1}(y), \quad x_2 = 2\pi - x_1$$

the Jacobians are

$$g'(x_1) = -\sin(x_1) = -\sin(\cos^{-1}(y)) = -\sqrt{1-y^2}, \quad g'(x_2) = \sqrt{1-y^2}$$

since  $f_X(x) = 1/2\pi$  in the interval  $(0, 2\pi]$ , so

$$f_Y(y) = \frac{1}{\pi\sqrt{1-y^2}}, \quad \text{for } -1 < y < 1$$

note that although  $f_Y(\pm 1) = \infty$  the probability that  $y = \pm 1$  is 0

3. The random variable  $X$  is uniform in the interval  $[0, 1]$ . Find the pdf of the random variable  $Y$  if  $Y = -\ln X$ .
- 4 If  $X$  has Exponential Distribution with parameter 1, find the p.d.f of  $Y = \sqrt{X}$ .
- 5 If  $X$  is a Normal R.V with mean zero and variance  $\sigma^2$  Find the p.d.f of  $Y = e^X$ .

# Expectation of Function of the Random Variable

Let  $X$  be a random variable with a probability distribution  $f(x)$ , and let  $g(X)$  be a function of the random variable  $X$ .

The mean (or expected value) of the random variable  $g(X)$  is denoted by  $\mu_{g(X)}$  or  $E[g(X)]$  and is defined by:

$$E[g(X)] = \mu_{g(X)} = \begin{cases} \sum_{\text{all } x} g(x) f(x); & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x) f(x) dx; & \text{if } X \text{ is continuous} \end{cases}$$

## Theorem

If  $X$  is a random variable with mean  $\mu = E(X)$ , and if  $a$  and  $b$  are constants, then:

$$E(aX \pm b) = a E(X) \pm b$$

$$\Leftrightarrow$$

$$\mu_{aX \pm b} = a \mu_X \pm b$$

**Corollary 1:**  $E(b) = b$  ( $a=0$  in Theorem)

**Corollary 2:**  $E(aX) = a E(X)$  ( $b=0$  in Theorem)

## Corollary 3:

If  $X$ , and  $Y$  are random variables, then:

$$E(X \pm Y) = E(X) \pm E(Y)$$

## Discrete Random Variable

### Example:

Let  $X$  be a discrete random variable with the following probability distribution

Find  $E[g(X)]$ , where  $g(X) = (X - 1)^2$ .

| $x$             | 0               | 1               | 2              |
|-----------------|-----------------|-----------------|----------------|
| $F(x) = p(X=x)$ | $\frac{10}{28}$ | $\frac{15}{28}$ | $\frac{3}{28}$ |

## Solution:

$$g(X) = (X - 1)^2$$

$$E[g(X)] = \mu_{g(X)} = \sum_{x=0}^2 g(x) f(x) = \sum_{x=0}^2 (x - 1)^2 f(x)$$

$$= (0 - 1)^2 f(0) + (1 - 1)^2 f(1) + (2 - 1)^2 f(2)$$

$$= (-1)^2 \frac{10}{28} + (0)^2 \frac{15}{28} + (1)^2 \frac{3}{28}$$

$$= \frac{10}{28} + 0 + \frac{3}{28} = \frac{10}{28}$$

## Continuous Random Variable

**Example.** Let  $X$  be a random variable with the following probability density function:

$$f(x) = \begin{cases} \frac{1}{3}x^2 & ; -1 < x < 2 \\ 0 & ; \text{elsewhere} \end{cases}$$

Find  $E(4X + 3)$ .

### Solution:

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{-1}^2 x \left[ \frac{1}{3} x^2 \right] dx = \frac{1}{3} \int_{-1}^2 x^3 dx = \frac{1}{3} \left[ \frac{1}{4} x^4 \right]_{x=-1}^{x=2} = 5/4$$

$$E(4X+3) = 4 E(X) + 3 = 4(5/4) + 3 = 8$$

Another solution:

$$E[g(X)] = \int g(x) f(x) dx \quad ; g(X) = 4X+3$$

$$E(4X+3) = \int_{-\infty}^{\infty} (4x+3) f(x) dx = \int_{-1}^2 (4x+3) \left[ \frac{1}{3} x^2 \right] dx = 8$$

### Example

Probability of  $X$ , the number of cars passing through a car wash in one hour on a sunny Friday afternoon, is given by

| $x$        | 4    | 5    | 6   | 7   | 8   | 9   |
|------------|------|------|-----|-----|-----|-----|
| $P(X = x)$ | 1/12 | 1/12 | 1/4 | 1/4 | 1/6 | 1/6 |

Let  $g(X) = 2X - 1$  represent the amount of money paid to the attendant by the manager.

What can the attendant expect to earn during this hour on any given sunny Friday afternoon?

**Solution:**

$$E[g(X)] = \sum g(x) f(x) = \sum (2X-1) f(x)$$

$$= (2*4-1)(1/12) + (2*5-1)(1/12) \dots + (2*9-1)(1/6) = 12.67$$

# Covariance

- The expected value of  $X$  identifies the center of mass of the distribution of  $X$ . The variance of the expected value of provides a measure of the spread of the distribution.
- In the case of two random variables we are interested in how  $X$  and  $Y$  vary together. In particular, we are interested in whether the variation of  $X$  and  $Y$  are correlated . For example, if  $X$  increases does  $Y$  tend to increase or to decrease?
- Covariance measure the nature of the association between two variables and describes a potential linear relationship:
- Positive relationship
  - \*Large values of  $X$  result in large values of  $Y$
- Negative relationship
  - \*Large values of  $X$  result in small values of  $Y$

# Covariance between two random variables is Defined as

## For Continuous

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

## For Discrete

$$\sigma_{xy} = \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) P(x_i, y_i)$$

$$\begin{aligned}\text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY - XE(Y) - E(X)Y + E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

## Covariance between two random variables:

1.  $\text{Cov}(X, Y) > 0$       X and Y are positively correlated
2.  $\text{Cov}(X, Y) < 0$       X and Y are inversely correlated
3.  $\text{Cov}(X, Y) = 0$       X and Y are independent

## Correlation

- Correlation analysis is a method of statistical evaluation used to study the strength of a relationship between two, numerically measured, continuous variables.
- Values between -1 and 1, and independent random variables have a correlation of zero.
- If two random variables have a covariance of zero, they are said to be uncorrelated

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Question: If random variable  $X$  and  $Y$  have linear relationship, that is,

$$Y = aX + b$$

$$\begin{aligned} \text{Cov}(X, Y) &= E[XY] - E[X]E[Y] \\ &= E[X(aX + b)] - E[X]E[aX + b] \\ &= aE[X^2] + bE[X] - aE^2[X] - bE[X] \\ &= a(E[X^2] - E^2[X]) = a\text{Var}(X) \end{aligned}$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{a\text{Var}(X)}{\sqrt{\text{Var}(X)a^2\text{Var}(X)}}$$

**That is,  $\text{Corr}(X,Y) = 1$  if  $a > 0$ ;  
 $= -1$  if  $a < 0$ .**

**Question:** If  $X$  and  $Y$  are independent random variables then

**Prove that**  $\text{cov}(X, Y) = \rho(X, Y) = 0$

**proof:** (for discrete random variables) Note that, since  $X$  and  $Y$  are independent we have

$$f_{XY}(x, y) = f_X(x)f_Y(y) .$$

Now

$$\begin{aligned}\mathbb{E}[XY] &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} xy f_{XY}(x, y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} xy f_X(x) f_Y(y) \\ &= \sum_{x \in \mathcal{X}} \left( x f_X(x) \sum_{y \in \mathcal{Y}} y f_Y(y) \right) = \sum_{x \in \mathcal{X}} (x f_X(x) \mathbb{E}[Y]) \\ &= \mathbb{E}[Y] \sum_{x \in \mathcal{X}} x f_X(x) = \mathbb{E}[Y] \mathbb{E}[X]\end{aligned}$$

$$\text{Then } \text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 \quad \Rightarrow \rho(X, Y) = 0.$$

## Expected Value of a Function of Two Random Variables

- The problem of finding the expected value of a function of two or more random variables is similar to that of finding the expected value of a function of a single random variable.
- For  $Z = g(x, y)$ , Expected value is calculated as,

$$E[Z] = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy & X, Y \text{ jointly continuous} \\ \sum_i \sum_n g(x_i, y_n) p_{X,Y}(x_i, y_n) & X, Y \text{ discrete.} \end{cases}$$

## Sum of Random Variables

Let  $Z = X + Y$ . Find  $E[Z]$ .

$$\begin{aligned} E[Z] &= E[X + Y] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x' + y') f_{X,Y}(x', y') dx' dy' \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x' f_{X,Y}(x', y') dy' dx' + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y' f_{X,Y}(x', y') dx' dy' \\ &= \int_{-\infty}^{\infty} x' f_X(x') dx' + \int_{-\infty}^{\infty} y' f_Y(y') dy' = E[X] + E[Y]. \end{aligned}$$

The expected value of a sum of “n” random variables is equal to the sum of the expected values:

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + \cdots + E[X_n].$$

## Product of Functions of Independent Random Variables

Suppose that  $X$  and  $Y$  are independent random variables, and let  $g(X, Y) = g_1(X)g_2(Y)$ . Find  $E[g(X, Y)] = E[g_1(X)g_2(Y)]$ .

$$\begin{aligned} E[g_1(X)g_2(Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x')g_2(y')f_X(x')f_Y(y') dx' dy' \\ &= \left\{ \int_{-\infty}^{\infty} g_1(x')f_X(x') dx' \right\} \left\{ \int_{-\infty}^{\infty} g_2(y')f_Y(y') dy' \right\} \\ &= E[g_1(X)]E[g_2(Y)]. \end{aligned}$$

## CHARACTERISTIC FUNCTIONS

The characteristic function of a random variable  $X$  is defined as

$$\begin{aligned} C_X(t) &= \mathbb{E}[e^{itX}] \\ &= \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)], \end{aligned} \quad t \in \mathbb{R}$$

which can also be written as

$$C_X(t) = \int e^{itx} d\mathbb{P}_X.$$

If  $X$  is a continuous random variable with density function  $f_X(x)$ , then

$$C_X(t) = \int e^{itx} f_X(x) dx.$$

The advantage with the characteristic function is that it always exists, unlike the moment generating function, which can be infinite everywhere except  $s = 0$ .

## PROPERTIES OF CHARACTERISTIC FUNCTIONS

- 1) If  $Y = aX + b$ ,  $C_Y(t) = e^{ibt}C_X(at)$ .
- 2) If  $X$  and  $Y$  are independent random variables and  $Z = X + Y$ , then  $C_Z(t) = C_X(t)C_Y(t)$ .
- 3) If  $M_X(s) < \infty$  for  $s \in [-\epsilon, \epsilon]$ , then  $C_X(t) = M_X(it)$  for all  $t \in \mathbb{R}$ .
- 4)  $C_X(0) = 1$  and  $|C_X(t)| \leq 1, \forall t \in \mathbb{R}$ .

**Example 1:** Let  $X$  be an exponential random variable with parameter  $\mu$ . Find its characteristic function.

**Solution:** Recall that for an exponential random variable with parameter  $\mu$ ,  $f_X(x) = \mu e^{-\mu x}$ . Thus, we have

$$\begin{aligned} C_X(t) &= \int_{x=0}^{\infty} \mu e^{-\mu x} e^{itx} dx \\ &= \frac{\mu}{\mu - it}. \end{aligned}$$

We have evaluated the above integral essentially by pretending that  $\mu - it$  is a real number. Although this happens to produce the correct answer in this case, the correct method of evaluating a characteristic function is by performing contour integration. Indeed, in the next example, it is not possible to obtain the correct answer by pretending that  $it$  is a real number (which is not).

**Example 2:** Let  $X$  be a Cauchy random variable. Find its characteristic function.

**Solution:** The density function for a Cauchy random variable is

$$f_X(x) = \frac{1}{\pi(1+x^2)}.$$

Therefore,

$$\begin{aligned} C_X(t) &= \int_{x=-\infty}^{\infty} \frac{e^{itx}}{\pi(1+x^2)} dx \\ &= e^{-|t|}. \end{aligned}$$

The above expression is not entirely trivial to obtain. Indeed, it requires considering two separate contour integrals for  $t > 0$  and  $t < 0$ , and invoking Cauchy's residue theorem to evaluate the contour integrals.

Example:

1. [Papoulis] Use characteristic function definition to find the distribution of  $Y = aX^2$ , if  $X$  is Gaussian with zero mean and variance  $\sigma^2$ .

Suppose  $X$  is  $N(0;\sigma)$  and  $Y = aX^2$ . Then

$$C_Y(t) = E(e^{itY}) = E(e^{itaX^2}) = \int_{-\infty}^{\infty} (e^{itaX^2}) f_X(x) dx$$

$$= \frac{2}{\sqrt{2\pi}\sigma} \int_0^{\infty} (e^{itaX^2}) e^{-x^2/2\sigma^2} dx.$$

For  $0 \leq x < \infty$ , note that the transformation  $y = ax^2$  is one-to-one. Hence, make the change of variable  $y = ax^2$ ,  $dy = (2ax)dx = 2\sqrt{ay} dx$  to obtain

$$= \frac{2}{\sqrt{2\pi}\sigma} \int_0^{\infty} e^{ity} e^{-y/2a\sigma^2} \frac{dy}{2\sqrt{ay}} = \int_0^{\infty} e^{ity} \left[ \frac{e^{-y/2a\sigma^2}}{\sigma\sqrt{2\pi ay}} \right] dy.$$

$$f_Y(y) = \frac{e^{-y/2a\sigma^2}}{\sigma\sqrt{2\pi ay}} U(y).$$

2. [Papoulis] Use characteristic function definition to find the distribution of  $Y = \sin(X)$ , if  $X$  is uniformly distributed in  $(-\pi/2, \pi/2)$ .

## □ Central Limit Theorem (CLT)

CLT states that if we sum a sufficient number of random variables, the resulting probability is a Gaussian.

Let  $\{X_n\}$  be a sequence of  $n$  independent

identically distributed random variables with mean  $\mu_X$  and variance  $\sigma_X^2$ . Their sum  $S_n$  is given by

$$S_n = X_1 + X_2 + \cdots + X_n$$

The mean  $\mu$  and the variance  $\sigma^2$  of  $S_n$  are

$$E[S_n] = \mu = n\mu_X$$

$$\text{var}[S_n] = \sigma^2 = n\sigma_X^2$$

We will now define a normalized random variable  $Z_n$  with mean 0 and variance 1 as follows:

$$Z_n = \frac{S_n - \mu}{\sigma} = \frac{S_n - n\mu_X}{\sigma_X \sqrt{n}}$$

The central limit theorem states that as  $n$  tends to  $\infty$ ,  $Z_n$  tends to a standard normal distribution:

$$\lim_{n \rightarrow \infty} P(Z_n \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-(x^2/2)} dx$$

We can show this result heuristically by using moment generating functions (MGFs) in

$$Z_n = \frac{S_n - n\mu_X}{\sigma_X \sqrt{n}} = \frac{1}{\sigma_X \sqrt{n}} \sum_{i=1}^n (X_i - \mu_X)$$