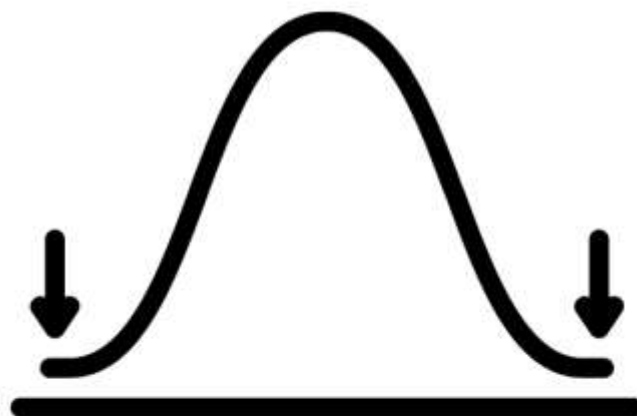




Probability Distribution

$$P(A) = \frac{n(A)}{n(S)}$$



WWW.TYCHR.COM

By the end of this chapter you should be familiar with:

- Continuous and discrete variables
- Mean and variance of a binomial distribution
- Mean and variance of Poisson distribution
- Mean and variance of a normal distribution
- Linear transformation
- Combination of data
- Markov chains

RANDOM VARIABLES

They can be **discrete or continuous**. All the probability distributions that we are going to discuss come under **discrete random variable (DRV)**.

The assignment of a value to every possible outcome of a random experiment is called random variable. Mathematically random variable is a function from the set of all sample space to the set of **all real numbers**.

DISCRETE RANDOM VARIABLE

If a random variable takes an integer value then it is called a discrete random variable or a random variable is only discrete if and only if its **range** is countable.

If X is a discrete random variable with range $R_x = \{x_1, x_2, \dots, x_n\}$ then the function $P_x(x_k) = P(X = x_k)$ for all $k = 1, 2, 3, \dots$ is called a probability mass function. The properties are:

- $0 \leq P_x(x_i) \leq 1$
- $\sum_{i=1}^n P_x(x_i) = 1$

If x is a random variable then **mean or expected value** of discrete

random variable is denoted by
$$\mu = \sum_{i=1}^n x_i P(X = x_i)$$

The **variance** of a discrete random variable measures the spread or the variability of the distribution and is denoted by

$$\sigma_x^2 = \sum_{i=1}^n (x_i - \mu)^2 P(X = x_i)$$

Standard deviation is σ_x

As we already know, if $F(x) = P(X \leq x)$

$\sum_{i=1}^n P_x(x_i)$ and $P(X > x) = 1 - F(x)$

Example: Let x be a DRV with range $R_x = \{1, 2, 3, \dots\}$. Suppose PMF of x is given by $P_x(k) = 1/2^k$ for $k = 1, 2, 3, \dots$. Find

a. $P(2 < X \leq 5)$

b. $P(X > 4)$

Solution:

a. $P(2 < X \leq 5) = P(3) + P(4) + P(5)$
 $= 1/2^3 + 1/2^4 + 1/2^5$
 $= 7/32$

b. $P(X > 4) = 1 - P(X \leq 4)$
 $= 1 - [P(1) + P(2) + P(3) + P(4)]$
 $= 1 - [1/2 + 1/2^2 + 1/2^3 + 1/2^4]$
 $= 1/16$

CONTINUOUS RANDOM VARIABLE

If a random variable takes all possible values between two limits or if it represents a **real number** then the random variable is called a **continuous random variable**.

BINOMIAL DISTRIBUTION

A random experiment which has only two possible outcomes is called the **Bernoulli's trial**. If **p** represents the **probability of success** and **q** the **probability of failure** then

the **mean** and **variance** and the probability distribution is given by:

$$E(X) = \mu = np$$

$$\text{Var}(X) = \sigma^2 = npq$$

$$P(x) = b(n,p,x) = {}_n C_x p^x (1-p)^{n-x}$$

Example: If the mean and variance of correctly answered questions in a test given to 4096 students are 2.5 and 1.875. Find and estimate of the number of students answering

- a. 2 or less correctly
- b. 5 questions correctly

Solution: $\mu = np$, $2.5 = np$, $\sigma^2 = npq$ $1.875 = 2.5q$

$$q = 0.75, p = 1 - q = 0.25, n = 10$$

- a. $P(X \leq 2) = P(1) + P(2) + P(3)$
 $= {}^{10}C_1 (0.25)^1 (0.75)^{10-1} + {}^{10}C_2 (0.25)^2 (0.75)^{10-2} + {}^{10}C_3 (0.25)^3 (0.75)^{10-3}$
 $= 0.719$
- b. $P(X = 5) = {}^{10}C_5 (0.25)^5 (0.75)^{10-5} = 0.0583$

POISSON DISTRIBUTION

Poisson distribution gives the probability of happening of the number of events occurring in a fixed interval of time, space or any other parameter with a known average. The mean, variance and the probability distribution is given by:

$$E(X) = \mu = np$$

$$\text{Var}(X) = \sigma^2 = \mu$$

$$P(x) = p(\mu, x) = \frac{e^{-\mu} \mu^x}{x!}$$

Example: The number of accidents in a year to the taxi drivers in the city follow Poisson distribution with $\mu = 3$. Out of 1000 taxi drivers find approximately the number of drivers with

- a. No accident in a year
- b. More than 3 accidents in a year

Solution:

- a. $P(0) = e^{-3} = 0.0497 \times 1000 = 50$
- b. $P(X > 3) = 1 - P(X \leq 3)$
 $= 1 - [P(0) + P(1) + P(2) + P(3)]$
 $= 1 - e^{-3} [1 + 3 + 4.5 + 4.5]$
 $= 0.35276 \times 1000$
 $= 353$

NORMAL DISTRIBUTION

A **normal distribution** is a binomial distribution for a **very large n** and the **probability success is close to half**. Given set of data is said to be normally distributed if the weight at which the frequencies fall off is proportional to the distance from the mean and frequencies themselves. The distribution is given by:

$$N(\mu, \sigma) = f_x(x) = \frac{e^{-\frac{1}{2\sigma^2}(x-\mu)^2}}{\sigma\sqrt{2\pi}} \quad \text{where } x \text{ can be any value}$$

between $-\infty$ to ∞ , σ is always positive and mean can also be any value between $-\infty$ to ∞ . The graph of a normal distribution is called a **bell curve** which is **symmetric about the mean**.

The normal distribution for which $\mu = 0$ and $\sigma = 1$ is called standard normal distribution or the **z distribution** and $z = (x - \mu)/\sigma$

Example: If x is a random variable which is distributed normally with $\mu = 60$ and $\sigma = 5$. Find the probability of

1. $60 \leq X \leq 70$
2. $X \leq 50$
3. $X > 40$

Solution:

1. $z = (x - \mu)/\sigma = (x - 60)/5$
 $p(0 \leq z \leq 2) = A(2) = 0.4772$
2. $z = -2$
 $p(z \leq 0) - p(-2 \leq z \leq 0)$
 $= 0.5 - 0.4772 = 0.0228$
3. $z = -4$
 $p(z \geq 0) + p(0 \leq z \leq 4) = 0.5 + 0.4996 = 0.996$

LINEAR TRANSFORMATION

A **linear transformation** is a change to a variable characterized by one or more of the following operations: adding a constant to the variable, subtracting a constant from the variable, multiplying the variable by a constant, and/or dividing the variable by a constant.

When a linear transformation is applied to a random variable, a new random variable is created. To illustrate, let X be a random variable, and let m and b be constants. Each of the following examples show how a linear transformation of X defines a new random variable Y .

- Adding a constant: $Y = X + b$
- Subtracting a constant: $Y = X - b$
- Multiplying by a constant: $Y = mX$
- Dividing by a constant: $Y = X/m$
- Multiplying by a constant and adding a constant: $Y = mX + b$
- Dividing by a constant and subtracting a constant: $Y = X/m - b$

EFFECT ON MEAN AND VARIANCE

Suppose a linear transformation is applied to the random variable X to create a new random variable Y . Then, the mean and variance of the new random variable Y are defined by the following equations.

$$\bar{Y} = m\bar{X} + b \text{ and } \text{Var}(Y) = m^2 \text{Var}(X)$$

where m and b are constants, \bar{Y} is the mean of Y , \bar{X} is the mean of X , $\text{Var}(Y)$ is the variance of Y , and $\text{Var}(X)$ is the variance of X .

Example: The average salary for an employee at Acme Corporation is \$30,000 per year. This year, management awards the following bonuses to every employee.

- A Christmas bonus of \$500.
- An incentive bonus equal to 10 percent of the employee's salary.

What is the mean bonus received by employees?

Solution: $Y = mX + b$

$Y = 0.10 X + 500$ where Y is the transformed variable (the bonus), X is the original variable (the salary), m is the multiplicative constant 0.10, and b is the additive constant 500.

Since we know that the mean salary is \$30,000, we can compute the mean bonus from the following equation.

$$\bar{Y} = m\bar{X} + b$$

$$\bar{Y} = 0.10 \$30,000 + \$500 = \$3,500$$

RANDOM VARIABLE COMBINATIONS EFFECT ON MEAN

Suppose you have two variables: X with a mean of μ_x and Y with a mean of μ_y . Then, the mean of the sum of these variables μ_{x+y} and the mean of the difference between these variables μ_{x-y} are given by the following equations.

$$\mu_{x+y} = \mu_x + \mu_y \text{ and } \mu_{x-y} = \mu_x - \mu_y$$

The above equations for general variables also apply to random variables. If X and Y are random variables, then

$$E(X + Y) = E(X) + E(Y) \text{ and } E(X - Y) = E(X) - E(Y)$$

where $E(X)$ is the expected value (mean) of X, $E(Y)$ is the expected value of Y, $E(X + Y)$ is the expected value of X plus Y, and $E(X - Y)$ is the expected value of X minus Y.

EFFECT ON VARIANCE

Suppose X and Y are independent random variables. Then, the variance of $(X + Y)$ and the variance of $(X - Y)$ are described by the following equations

$$\text{Var}(X + Y) = \text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$$

where $\text{Var}(X + Y)$ is the variance of the sum of X and Y, $\text{Var}(X - Y)$ is the variance of the difference between X and Y, $\text{Var}(X)$ is the variance of X, and $\text{Var}(Y)$ is the variance of Y.

Note: The standard deviation (SD) is always equal to the square root of the variance (Var). Thus,
 $\text{SD}(X + Y) = \text{sqrt}[\text{Var}(X + Y)]$ and $\text{SD}(X - Y) = \text{sqrt}[\text{Var}(X - Y)]$

Example: Suppose X and Y are independent random variables. The variance of X is equal to 16; and the variance of Y is equal to 9. Let $Z = X - Y$. What is the standard deviation of Z ?

Solution: The solution requires us to recognize that Variable Z is a combination of two independent random variables. As such, the variance of Z is equal to the variance of X plus the variance of Y .
$$\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y) = 16 + 9 = 25$$

The standard deviation of Z is equal to the square root of the variance. Therefore, the standard deviation is equal to the square root of 25, which is 5.

MARKOV CHAINS

A **Markov chain** is a mathematical system that experiences transitions from one state to another according to certain probabilistic rules.

A Markov chain is a **stochastic process**, but it differs from a general stochastic process in that a Markov chain must be “memory-less.” That is, (the probability of) future actions are not dependent upon the steps that led up to the present state. This is called the **Markov property**.

A **transition matrix T** for Markov chain at time t is a matrix containing information on the probability of transitioning between states.

$$X_n = X_0 T^n$$

Lets take an example to understand better

Example: Consider the Markov chain with three states, $S = \{1,2,3\}$,

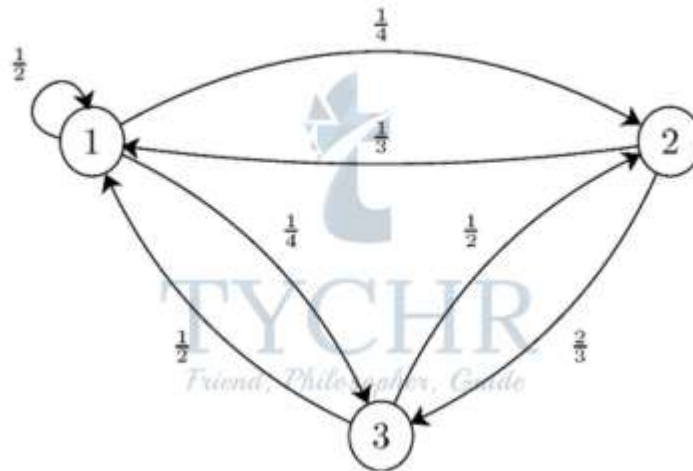
$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/3 & 0 & 2/3 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

that has the following transition matrix

1. Draw the state transition diagram for this chain.

2. If we know $P(X_1 = 1) = P(X_1 = 2) = 1/4$, find $P(X_1 = 3, X_2 = 2, X_3 = 1)$

Solution:



1. First, we obtain

$$P(X_1 = 3) = 1 - P(X_1 = 1) - P(X_1 = 2) = 1 - 1/4 - 1/4 = 1/2$$

We can now write

$$\begin{aligned} P(X_1 = 3, X_2 = 2, X_3 = 1) &= P(X_1 = 3) \cdot p_{32} \cdot p_{21} \\ &= (1/2) (1/2) (1/3) = 1/12. \end{aligned}$$

TYCHR
Friend, Philosopher, Guide



WWW.TYCHR.COM



+91 9540653900