UNIT 9 : THEORETICAL DISTRIBUTION II

UNIT STRUCTURE

9.1 Learning Objectives

9.2 Introduction

9.3 Continuous Probability Distribution

9.4 Normal Distribution

- 9.4.1 Properties of Normal Distribution
- 9.4.2 Standard Normal Variate

9.5 Let Us Sum Up

9.6 Further Readings

9.7 Answers To Check Your Progress

9.8 Model Questions

9.1 LEARNING OBJECTIVES

After going through this unit, you will able to-

- continuous probability distribution and probability density function
- normal distribution
- standard normal variate
- properties of normal distribution.

9.2 INTRODUCTION

Normal distribution is a continuous distribution. It is a limiting case of the binomial distribution. It has immense application. It is the most popular and commonly used distribution. It was discovered by Abraham de Moivre (1667–1754) in 1733.

9.3 CONTINUOUS PROBABILITY DISTRIBUTION

If a random variable is a continuous variable, its probability distribution is called a continuous probability distribution.

A continuous probability distribution cannot be expresed in tabular form. An equation or formula is used to describe a continuous probability

distribution. The equation used to describe a continuous probability distribution is called a probability density function (p.d.f.)

9.4 NORMAL DISTRIBUTION

Normal distribution is derived from the binomial distribution under two conditions, viz.

(i) n, the number of trials, is sufficiently large, i.e. $n \to \infty$

and (ii) neither p nor q is very small.

A random variable X is said to follow normal distribution, if and only if, its probability density function is given by

Where x is the real value of X, i.e. $-\infty < X < \infty$

The variable X is said to be distributed normally with mean μ and variance . symbolically we write $X \sim N(\mu, \alpha^2)$. It is read as 'X is distributed as $N(\mu, \alpha^2)$'.

Note :

1. The probability density function given by (1) has two parameters, viz., μ and .
   is any real number, i.e. whereas is a positive real number, i.e., .
2. We know that probability can never be negative.
3. If $\mu = 0,\ \alpha = 1$ the p.d.f is given by

In this case $X \sim N(0, 1)$. In this case X is called the standard normal variate, and the p.d.f given by (2) is called standard normal distribution.

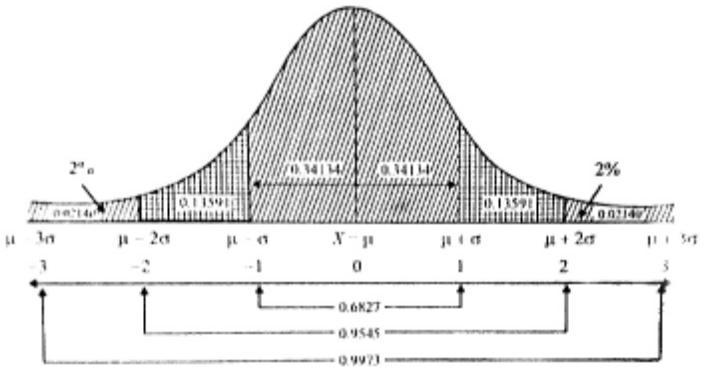
9.4.1 Properties of Normal Distribution

Normal distribution has many properties, some of which are given below.

1. The equation of the normal probability curve is

$$Y = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

It is of the shape given below. It is bell-shaped.



[Source : **Biostatistics** by Dr. P N Arora and Dr. P K Malhan (Himalya Publishing House, 2010)]

2. The curve is symmetrical about the line $x = \infty$, and x lies between and

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

3. At , the ordinate is maximum, and it is

4. The total area under the normal curve within the limits $-\infty$ to is 1, i.e.,

5. The percentage distribution of area under the normal curve given below, and is shown in the figure given above.
   (i) About 68% of the area falls between $\mu - \alpha$ and
   (ii) About 95.5% of the area falls between and .
   (iii) About 99.7% of the area falls between and .
6. The curve has a single peak, i.e., it is unimodal.
7. The distribution is symmctrical. Mean = median = mode = .
8. The curve is asymptotic, i.e., it meets the x-axis at both the

ends at infinity. Theoeretically the curve never touches the x-axis.

9. The points of inflexion are and . [Inflexion points are where the function changes concavity.]
10. Coefficient of skewness = 0, and kurtosis = 3.
11. Mean deviation about mean
12. $Q.D. \simeq \frac{2}{3}\alpha$

9.4.2 Standard Normal Variate

Let a random variable X be distributed as $N(\mu, \alpha^2)$. The location and shape of the normal curve depends on μ, and , Where , and . Therefore, no master table for the area can be prepared. This hindrance can be removed if we introduce a new variable Z defined by.

In this case Z is called the standard normal variate.

Now we prove that the mean of Z is 0, and the standard deviation is 1. Before that we state some related properties.

We know that mean of a random variable X is given by

mean $= E(X) = \Sigma x_i p_i$

We prove that

(i) E(a) = a; a is a constant

(ii) E (aX + b) = aE(x) + b; a, b are constants

(iii) $E\left(X - \overline{X}\right) = 0$

Proof : (i) $E(a) = \Sigma a p_i$

$= a\ \Sigma p_i$

= a, $\because\ \Sigma p_i = 1$

(ii) $\Sigma(aX + b) = \Sigma(ax_i + b)p_i$

$= a\Sigma x_i p_i + b\Sigma p_i$

$= a\,E(X) + b,\ \because\ \Sigma p_i = 1$

(iii) $E\left(X-\overline{X}\right)$

$= E(X) - E(\overline{X})$

$= \overline{X} - \overline{X}, \quad \because \overline{X}$ is constant

= 0

Also, we note that

(iv) Var (aX) = a^2 Var (X), where a is a constant

(v) Var (aX $\pm b$) = a^2 Var (X), where a, b are constants.

Now we prove that the mean of Z is 0, and the standard deviation is 1.

Mean = E (Z)

$= E\left(\frac{X-\mu}{\alpha}\right)$

$= \frac{1}{\alpha} E(X-\mu),$ property (ii)

$= \frac{1}{\alpha}\left[E(X)-\mu\right],$ property (ii)

$= \frac{1}{\alpha}[\mu-\mu]$

= 0

Variance = Var (Z)

$= Var\left(\frac{X-\mu}{\alpha}\right)$

$= \frac{1}{\alpha^2}\text{var}(X-\mu),$ property (v)

$= \frac{1}{\alpha^2}\text{var } X,$ property (iv)

$= \frac{1}{\alpha^2} . \alpha^2$

= 1

Thus, $Z = \frac{X-\mu}{\alpha}$ is distributed as N (0, 1), i.e., Z ~ N(0, 1).

All the properties for a normal variable hold good for standard normal variable $(\mu = 0,\ \alpha = 1)$

Let Z be a standard normal variate N(0, 1). The area under the standard normal curve is given by a table which is addended at the end of this chapter.

We know that the total area under the normal curve within the limits to is 1. The standard normal curve is symmetrical about Z = 0. Therefore, the area to the left of z = 0 is 0.5, and the area to the right of z = 0 is 0.5.

Suppose, we are to find the area between z = 0, and z = 1.89, i.e, we are to find out .

First we find 1.8 at the left of the table and single out that row.

Next we find 0.09 at the top of the table and single out that column.

Then we obtain the entry where the row for 1.8 and the column for 0.09 meet. the entry is 0.47062.

Thus $P(0 \le Z \le 1.89)$

= (the are from z = 0 to z = 1.89)

= 0.47062.

Example 1 : Find the probability of the standard normal variate z when z is greater than 1.09.

Solution : We know that the total area under the normal curve within the limits $-\infty$ to is 1. Therefore, the area to the right of z = 0 is 0.5, as the curve is symmetrical about z = 0.

P(Z > 1.09) = 0.5 – 0.36214 = 0.13786

Example 2 : Find the probability of the standard normal variate z when z is less than 1.09.

Solution : We know that the total area under the normal curve within the limits to is 1. Therefore, the area to the left of z = 0 is 0.5, and the area to the right or z = 0 is 0.5.

Again, from the table, we get that the area from z = 0 to z = 1.09 is 0.36214.

P(z < 1.09)

= (area to the left of z = 0) + (area from z = 0 to z = 1.09)

= 0.5 + 0.36214

= 0.86214

Example 3 : Find P(0.5 < z < 1.09)

Solution : We know that the area to the right of z = 0 is 0.5. From the table, the area from z = 0 to z = 0.5 is 0.19146. Again, the area from z = 0 to z = 1.09 is 0.36214.

P(0.5 < z < 1.09)

= (Area from z = 0 to z = 1.09) – (area from z = 0 to x = 0.5)

= 0.36214 – 0.19146

= 0.17068

Example 4 : A large number of measurements is normally distributed with mean 65.5 cm and standard deviation 6.2 cm. Find the percentage of measurements that fall between 54.8 cm and 68.8 cm.

Solution : Let X be the normal variate showing the large number of measurements.

$\mu = 65.5, \alpha = 6.2$

$Z = \frac{X - \mu}{\alpha} = \frac{X - 65.5}{6.2}$ We have

Now,

When, X = 54.8 $Z = \frac{54.8 - 65.5}{6.2} = \frac{-10.7}{6.2} = -1.7$

When, X = 68.8, $Z = \frac{68.8 - 65.5}{6.2} = 0.5$

Thus, when X lies between 54.8 and 68.8, Z lies between –1.7 and 0.5

Now P(–1.7 < X < 0.5)

= P(–1.7 < Z < 0) + P(0 < Z < 0.5)

= 0.45543 + 0.19146

= 0.64689 = 64.69%

$\therefore$ the percentage of measurements is 64.69%

Example 5 : For the distribution of marks of candidates in an examination, mean is 14 and standard deviation is 2.5. Assuming the normality of the distribution, find the probability that a randomly

selected student will score above 15.

Solution : Let X be the normal variate showing the scores of candidates.

We have

Now,

When X,

$\therefore$ P(x > 15) = P(z > 0.4)

= 0.5 – (Area between z = 0 and Z = 0.4)

= 0.5 – 0.15542

= 0.34458.



CHECK YOUR PROGRESS 1

Q 1: Find (i) P (Z > 2.58), (ii) P(Z < 1.96), (iii) P (Z > –3)

Q 2: For a certain normal distribution, mean is 50, and standard deviation is 5. Find , X being a normal variate.

Q 3: A normal variate X has mean 100 and standard deviation 4. Find the probability that X lies between 92 and 10.8

Q 4: The distribution of scores of students in an examination is normal. If the mean is 45, and the standard deviation is 5, find the percentage of students who scored between 40 and 50.

Q 5: The marks obtained in a certain examination follow normal distribution with mean 45, and standard deviation 10. If 1000 students appear at the examination, calculate the number of students scoring (i) less than 30 marks, and (ii) more than 80 marks.



9.5 LET US SUM UP

- Normal distribution is a continous distribution. It is a limiting case of the binomial distribution.
- Normal distribution is derived from the binomial distribution under two conditions, viz.,
  - (i) n, the number of trials, is sufficiently large, i.e., $n \to \infty$ and
  - (ii) neither p nor q is very small.

  A random variable X is said to follow normal distribution, if and only if, its probability density function is given by

$$f(x) = \frac{1}{\alpha\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\alpha}\right)^2}$$

  Where x is the real value of X, i.e., $-\infty < X < \infty$

  The variable X is said to be distributed normally with mean and variance . Symbolically, we write $X \sim N(\mu,\ \alpha^2)$. It is read as 'X is distributed as $N(\mu,\ \alpha^2)$.
- The normal curve is bell-shaped, and the curve is symmetrical about the line $x = \mu$, and x lies between , and .
- The total area under the normal curve within the limits to is 1.
- If a new variable Z is defined by , Z is called the standard normal variate. The mean of Z is 0, and the standard deviation is 1.



9.6 FURTHER READINGS

1) Rao. G. Shanker, Probability and Statistics, Universitis Press.
2) Statistics, R.S.N. Pillai and Bagavathi S. Chand and Company Ltd. New Delhi-110055.



9.7 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

Ans to Q No 1: (i) We know that tha total area under the normal curve within the limits $-\infty$ to is 1. Therefore, the area to the left of z = 0 is 0.5, and the area to the night of z = 0 is 0.5. Again, from the table, we get that the area from z = 0 to z = 2.58 is 0.49506

P (z > 2.58)

= (area to the right of z = 0) – (area from z = 0 to z = 2.58)

= 0.5 – 0.49506

= 0.00494

(ii) We know that the total area under the normal curve within the limits to is 1. Therefore, the area to the left of z = 0 is 0.5, and the area to the right of z = 0 is 0.5. Again from the table, we get that the area from z = 0 to z = 1.96 is 0.47500.

P (z < 1.96)

= (area to the left of z = 0) + (area from z = 0 to z = 1.96)

= 0.5 + 0.47500

= 0.97500

(iii) We know that the total area under the normal curve within the limits to is 1. Therefore, the area to the left of z = 0 is 0.5, and the area to the right of z = 0 is 0.5.

Now, P(Z > –3)

= area from z = –3 to z = 0) + (area to the right of z = 0)

= (area from z = 0 to z = 3) + (area to the right of z = 0)

= 0.49865 + 0.5

= 0.99865

Ans to Q No 2: X is a normal variate.

We have

Now,

When X = 45

When X = 55, $Z = \frac{55-50}{5} = 1$

Thus, when X lies between 45 and 55, Z lies between –1 and 1.

Now $P(-1 \le Z \le 1)$

$= P(-1 \le Z \le 0) + P(0 \le Z \le 1)$

= 0.34134 + 0.34134

= 0.68268

Ans to Q No 3: Do yourself.

Ans to Q No 4: Do yourself.

Ans to Q No 5: Let X be the normal variate showing marks.

We have $\mu = 45,\ \alpha = 10$

Now

$Z = \frac{X-45}{10} = -1.5$

When X = 30,

When X = 80, $Z = \frac{80-45}{10} = \frac{35}{10} = 3.5$

(i) Now $P(X \le 30)$

= (area to the left of z = 0) – (area from z = – 1.5 to z = 0)

(area to the left of z = 0) – (area from z = 0) to z = 1.5)

= 0.5 – 0.43319

= 0.06681

$\therefore$ the required number of students

= 1000 x 0.06681

= 67 (approximately)

(ii)

$= P(Z \ge 3.5)$

= (area to the right of z = 0) – (area from z = 0 to z = 3.5)

= 0.5 – 0.49977

= 0.00023

Now 1000 x 0.00023 = 0.23

$\therefore$ there are no students scoring more than 80 marks.



9.8 MODEL QUESTIONS

Q 1: Define standard normal variate and write down the probability density function of this variate.

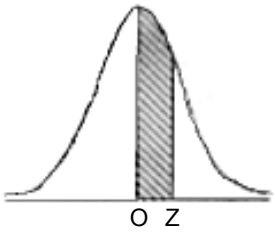
Q 2: State the conditions under which the normal distribution is obtained as a limiting case of the binomial distribution.

Q 3: Find (i) (ii) $P(Z \leq 1.96$ (iii) $P(Z > 2.58)$

Q 4: A normal variate X has mean 50 and standard deviation 5. Find $P(40 \leq X \leq 60)$.

Q 5: The mean of a normal distribution is 50 and the standard deviation is 4. What percentage of total frequency lies between 42.16 and 57.84?

Q 6: Let a continuous variable X follow a normal distribution with mean 12 and standard deviation 2. What is the probability that the value of X selected at random lies between 11 and 14 ?



Area under the Standard Normal Curve

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.00000	.00399	.00798	.01197	.01595	.01994	.02392	.02790	.03188	.03586
0.1	.03983	.04383	.04776	.05172	.05567	.05962	.06356	.06749	.07142	.07535
0.2	.07926	.08317	.08706	.09095	.09483	.09871	.10257	.10642	.11026	.11409
0.3	.11791	.12172	.12552	.12930	.13307	.13683	.14058	.14431	.14803	.15173
0.4	.15542	.15910-	.16276	.16640	.17003	.17364	.17724	.18082	.18439	.18793
0.5	.19146	.19497	.19847	.20194	.20540	.20884	.21226	.21566	.21904	.22240
0.6	.22575	.22907	.23237	.23565	.23891	.24215	.24537	.24857	.25175	.25490
0.7	.25804	.26115	.26424	.26730	.27035	.27337	.27637	.27935	.28230	.28524
0.8	.28814	.29103	.29389	.29673	.29955	.30234	.30511	.30785	.31057	.31372
0.9	.31594	.31859	.32121	.32381	.32639	.32894	.33147	.33398	.33646	.33891
1.0	.34134	.34375	.34614	.34849	.35083	.35314	.35543	.35769	.35993	.36214
1.1	.36433	.36650	.36864	.37076	.37286	.37493	.37698	:37900	.38100	.38298
1.2	.38493	.38686	.38877	.39065	.39251'	.39435	.39617	.39796	.39973	.40147
1.3	.40320	.40490	.40658	.40824	.40988	.41198	.41309	.41466	.41621	.41774
1.4	.41924	.42073-	.42220	.42364	.42507	.42647	.42785	.42922	.43056	.43189
1.5	.43319	.43448	.43574	.43699	.43822	.43943	.44062	.44179	.44295	.44408
1.6	.44520	.44630	.44738	.44845	.44950	.45053	.45154	.45254	.45352	.45449
1.7	.45543	.45637	.45728	.45818	.45907	.45994	.46080	.46164	.46246	.46327
1.8	.46407	.46485	.46562	.46638	.46712	.46784	.46856	.46926	.46995	.47062
1.9	.47128	.47193	.47257	.47320	.47381	.47441	.47500	.47558	.47615	.47670
2.0	.47725	.47784	.47831	.47882	.47932	.47982	.48030	.48077	.48124	.48169
2.1	.48214	.48257	.48300	.48341	.48382	.48422	.48461	.48500	.48537	.48574
2.2	.48610	.48645	.48679	.48713	.48745	.48778	.48809	.48840	.48870	.48899

2.3	.48928	.48956	.48983	.49010	.49036	.49061	.49086	.49111	.49134	.49158
2.4	.49180	.49202	.49224	.49245	.49266	.49286	.49305	.49324	.49343	.49361
2.5	.49379	.49396	.49413	.49430	.49446	.49461'	.49477	.49492	.49506	.49520
2.6	.49534	.49547	.49560	.49573	.49585	.49598	.49609	.49621	.49632.	.49643
2.7	.49653	.49664	.49674	.49683	.49693	.49702	.49711	.49720	.49728	.49736
2.8	.49744	.49752	.49760	.49767	.49774	.49781	.49788	.49795	.49801	.49807
2.9	.49813	.49819	.49825	.49831	.49836	.49841	.49846	.49851	.49856	.49861
3.0	.49865	.49869	.49874	.49878	.49882	.49886	.49889	.49893	.49896	.49900
3.1	.49903	.49906	.49910	.49913	.49916	.49918	.49921	.49924	.49926	.49929
3.2	.49931	.49934	.49936	.49938	.49940	.49942	.49944	.49946	.49948	.49950
3.3	.49952	.49953	.49955	.49957	.49958	.49960	.49961	.49962	.49964	.49965
3.4	.49966	.49968	.49969	.49970	.49971	.49972	.49973	.49974	.49975	.49976
3.5	.49977	.49978	.49978	.49979	.49980	.49981	.49981	.49982	.49983	.49983
3.6	.49984	.49985	.49985	.49986	.49986	.49987	.49987	.49988	.49988	.49989
3.7	.49989	.49990	.49990	.49990	.49991	.49991	.49992	.49992	.49992	.49992
3.8	.49993	.49993	.49993	.49994	.49994	.49994	.49994	.49995	.49995	.49995
3.9	.49995	.49995	.49996	.49996	.49996	.49996	.49996	.49996	.49997	.49997
4.0	.49997	-	-	-	-	-	-	-	-	-
4.5	.499997	-	-	-	-	-			-	-
5.0	.4999997	-	-	-	-	-	-	-	-	-

[Source : Basic statisties by B. L. Agarwal (New Age International Publishers, 2009)]

*** ***** ***