Some Basic Probability Concept and Probability Distribution

OSRC

Odisha Social Science Research & Consultancy

Continuous Probability Distribution

Continuous variable:

- Assumes any value within a specified interval/range.
- Consequently any two values within a specified interval, there exists an infinite number of values.
- As the number of observation, n, approaches infinite and the width of the class interval approaches zero, the frequency polygon approaches smooth curve.
- Such smooth curves are used to represent graphically the distribution of continuous random variable

- This has some important consequences when we deal with probability distributions.
- The total area under the curve is equal to 1 as in the case of the histogram.
- The relative frequency of occurrence of values between any two points on the x-axis is equal to the total area bounded by the curve, the x-axis and the perpendicular lines erected at the two points.

Graph of a continuous distribution showing area between a and b

Definition of probability distribution:

- A density function is a formula used to represent the probability distribution of a continuous random variable.
- This is a nonnegative function f (x) of the continuous r.v, x if the total area bounded by its curve and the x-axis is equal to 1 and if the sub area under the curve bounded by the curve, the x-axis and perpendiculars erected at any two points a and b gives the probability that x is between the point a and b.

Normal Distribution(C.F.Gauss, 1777-1855)

- The distribution is frequently called the Gaussian distribution.
- It is a relative frequency distribution of errors, such errors of measurement. This curve provides an adequate model for the relative frequency distributions of data collected from many different scientific areas.
- The density function for a normal random variable

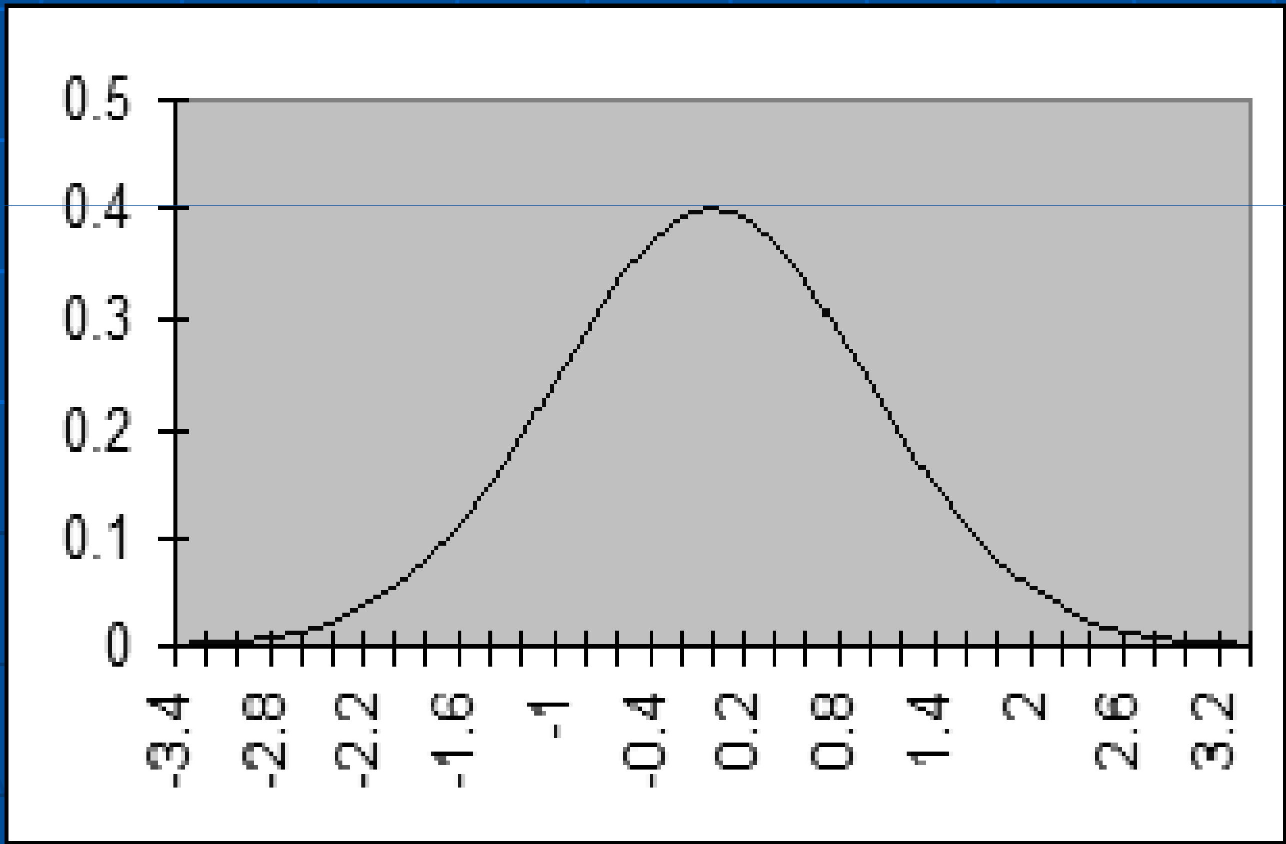
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

- The parameters μ and σ^2 are the mean and the variance , respectively, of the normal random variable

Characteristics Of The Normal Distribution

- It is symmetrical about its mean, μ.
- The mean, the median, and the mode are all equal.
- The total area under the curve above the x-axis is one square unit.
- This characteristic follows that the normal distribution is a probability distribution.
- Because of the symmetry already mentioned, 50% of the area is to the right of a perpendicular erected at the mean, and 50% is to the left.

- If $\mu = 0$ and $\sigma = 1$ then . The distribution with this density function is called the standardized normal distribution. The graph of the standardized normal density distribution is shown in Figure



- If 'x' is a normal random variable with the mean μ and variance σ then

  1) the variable $$z = \frac{x-\mu}{\sigma}$$

  is the standardized normal random variable.

The equation of pdf for standard normal distribution

$$\mathbf{f\,(z) = \frac{1}{\sqrt{2\prod}}}\ \mathrm{e}^{\mathbf{-z^2/2}},\ -\infty < z < \infty$$

- Area properties of normal distribution

$$P(|x-\mu| \leq 2\sigma) = 0.9544$$

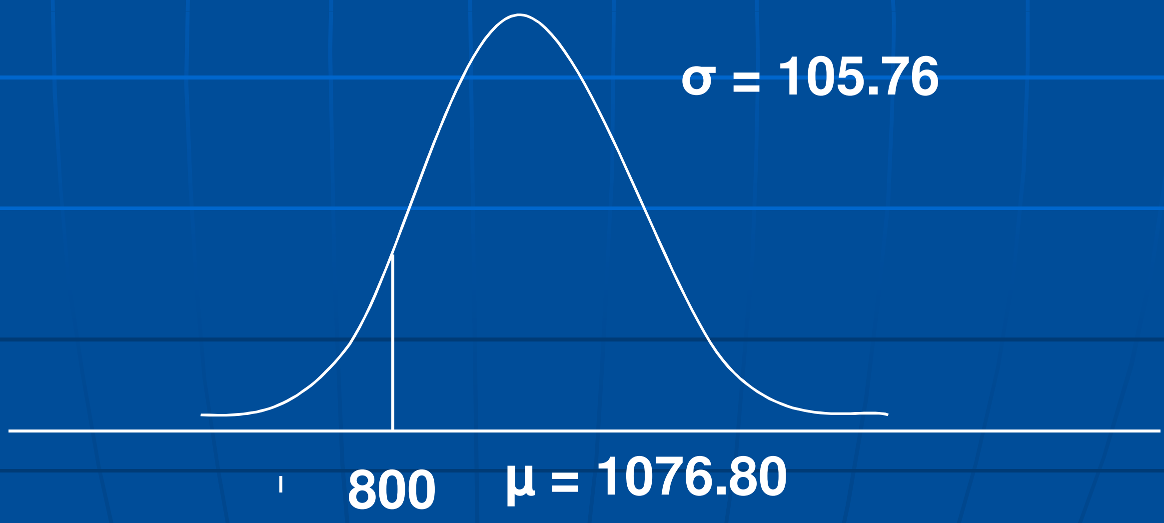
$$P(|x-\mu| \leq 2\sigma) = 0.9544$$

$$P(|x-\mu| \leq 3\sigma) = 0.9973$$

- Namely, if a population of measurements has approximately a normal distribution the probability that a random selected observation falls within the intervals $(\mu - \sigma, \mu + \sigma)$, $(\mu - 2\sigma, \mu + 2\sigma)$, and $(\mu - 3\sigma, \mu + 3\sigma)$, is approximately 0.6826, 0.9544 and 0.9973, respectively.

Normal Distribution Application

Example:1 As a part of a study of Alzeheimer's disease, reported data that are compatible with the hypothesis that brain weights of victims of the disease are normally distributed. From the reported data, we may compute a mean of 1076.80 grams and a standard deviation of 105.76 grams. If we assume that these results are applicable to all victims of Alzeheimer's disease, find the probability that a randomly selected victim of the disease will have a brain that weighs less than 800 grams.



Solution:

R.V x 'Brain weights' follows a Normal distribution with μ=1076.80 and σ = 105.76)

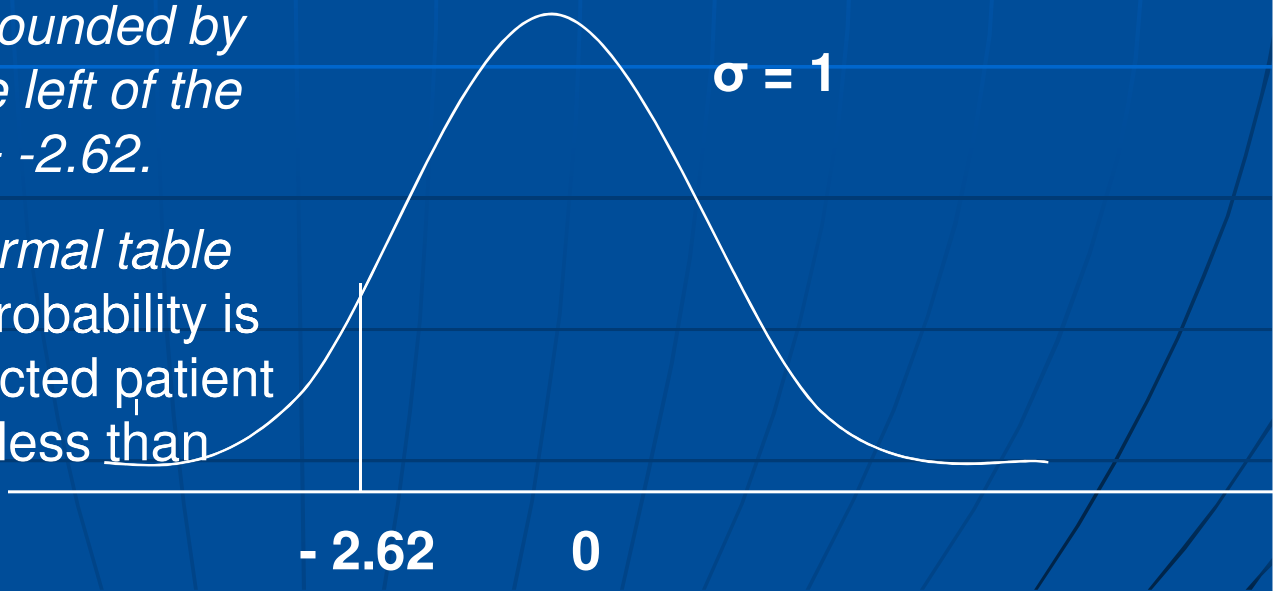
The Corresponding Standard Normal Variate

$$z = \frac{x - \mu}{\sigma}$$

= $$z = \frac{x - 1076.80}{105.76}$$

We have to find out P (x < 800) i.e P (z < -2.62). This is the area bounded by the curve, x axis and to the left of the perpendicular drawn at z = -2.62.

Thus from the standard normal table this prob., p= .0044. The probability is .0044 that a randomly selected patient will have a brain weight of less than 800 grams.

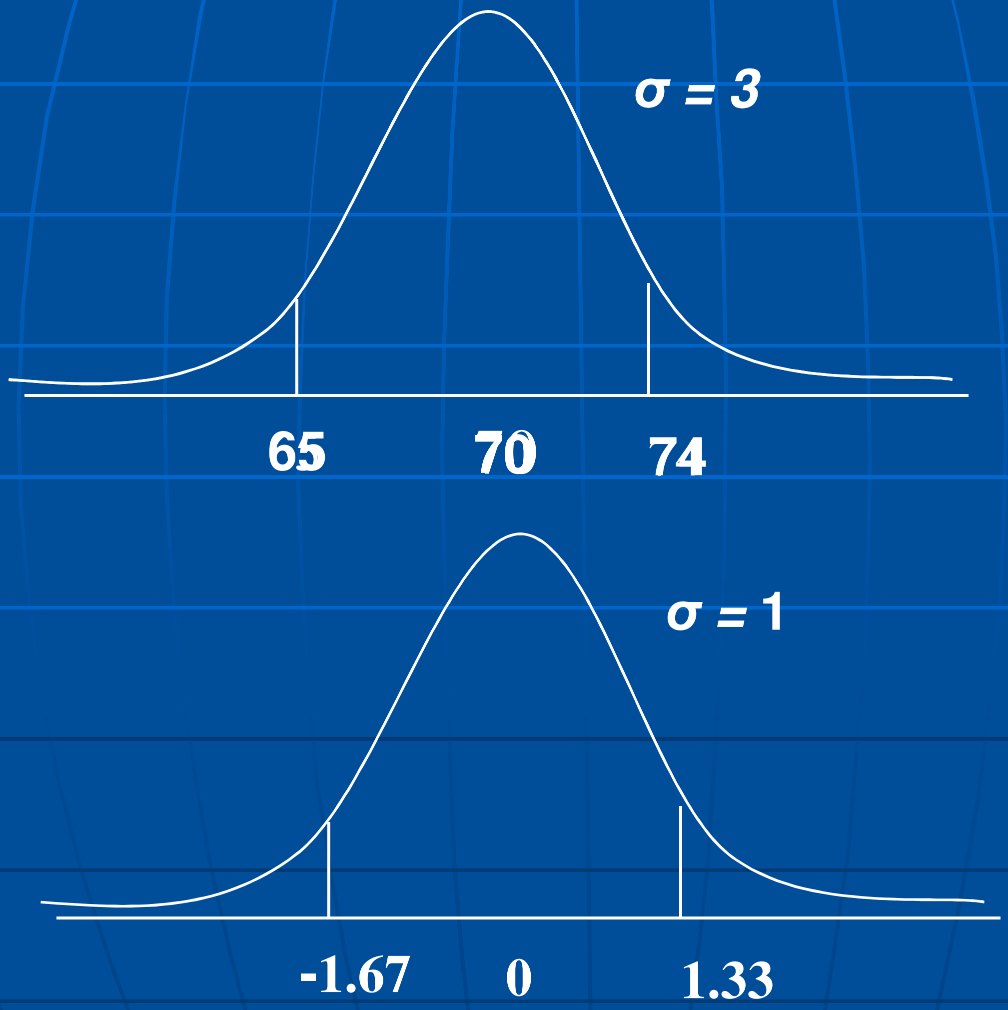


Example: 2

Suppose it is known that the heights of a certain population of individuals are approximately normally distributed with a mean of 70 inches and a standard deviation of 3 inches. What is the probability that a person picked at random from this group will be between 65 and 74 inches tall.

Solution: In fig are shown that the distribution of heights and the z distribution to which we transform the original values to determine the desired probabilities. We find the value corresponding to an x of 65 by

$$z = \frac{65-70}{3} = -1.76$$



Similarly, for x= 74 we have

$$z = \frac{74-70}{3} = 1.33$$

The area between -∞ and -1.76 to be .0475 and the area between -∞ and 1.33 to be .9082. The area desired is the difference between these, .9082- .0475=.8607

To summarize,

$$P(65 \le x \le 74) = P\left(\frac{65-70}{3} \le z \le \frac{74-70}{3}\right)$$

$$= P(-1.76 \le z \le 1.33)$$

$$= P(-\infty \le z \le 1.33) - P(-\infty \le z \le -1.67)$$

$$= .9082 - .0475$$

$$= .8607$$

The probability asked for in our original question, then, is .8607

Example: 3

In a population of 10,000 of the people described in previous example how many would you expect to be 6 feet 5 inches tall or taller?

Solution:

we first find the probability that one person selected at random from the population would be 6 feet 5 inches tall or taller. That is,

$$\mathbf{P\,(\,x \geq 77\,) = P\left(z \geq \frac{77\text{-}70}{3}\right)}$$

$$\mathbf{= P\,(\,z \geq 2.33\,) = 1 - .9901 = .0099}$$

Out of 10,000 people we would expect 10,000 (.0099) = 99 to be 6 feet 5 inches (77 inches tall or taller).

Exercise:

1. Given the standard normal distribution, find the area under the curve, above the z-axis between $z=-\infty$ and $z = 2$.
2. What is the probability that a z picked at random from the population of z's will have a value between -2.55 and + 2.55?
3. What proportion of z values are between -2.74 and 1.53?
4. Given the standard normal distribution, find $P(z \geq 2.71)$
5. Given the standard normal distribution, find $P(.84 \leq z \leq 2.45)$.