Econometrics

Course Calendar	
Week	Main Content
Week 1	**Introduction to Simple Regression**
Week 2	**Simple Regression**
Week 3	**Simple Regression: r^2 & Hands-on-Exercise**
Week 4	**Central Limit Theorem, Probability and Probability Density Function (PDF)**
Week 5	**Hypothesis Testing: Basics**
Week 6	**Simple Regression: Testing of Hypothesis**

Econometrics
Lecture 4: Central Limit Theorem, Probability and PDF

Geetha Rani Prakasam, Ph.D.
Professor

Recap

- Estimation of the simple regression parameters, intercept and slope and variance and standard errors
- The estimation done in excel with a numerical example
- Co-efficient of determination r^2 and its Estimation r^2 in excel
- Normality of U

outline

- Distribution of sample mean
- Std. deviation
- Distribution of Sample mean – examples
- Central Limit Theorem
- Confidence Intervals
- Sampling Distributions of Statistic
- Simulation of a Sampling Distribution of xbar
- Law of large numbers

outline

- Probability
- Probability distribution
- Probability density function
- Some important probability distribution

CENTRAL LIMIT THEOREM

- specifies a theoretical distribution
- formulated by the selection of all possible random samples of a fixed size n
- a sample mean is calculated for each sample and the distribution of sample means is considered

The Distribution of the Sample Mean

- The importance of the sample mean $\overline{X}$ comes from its use in drawing conclusions about the population mean μ.
- Some of the most frequently used inferential procedures are based on properties of the sampling distribution of $\overline{X}$.

The Distribution of the Sample Mean

- Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with mean value μ and standard deviation σ. Then

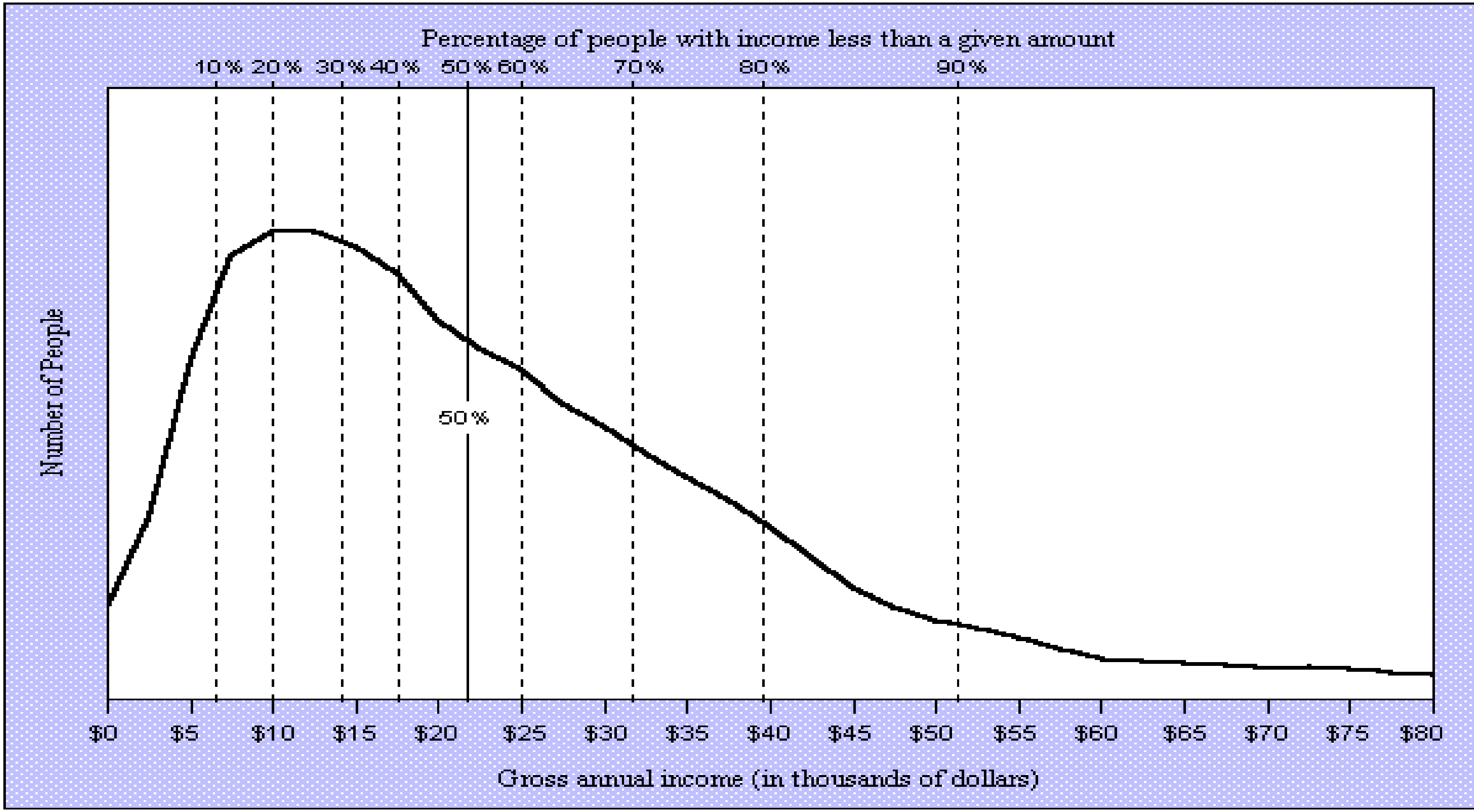
  1. $E(\overline{X}) = \mu_{\overline{X}} = \mu$

- As per result 1, the sampling (i.e., probability) distribution of $\overline{X}$ is centered precisely at the mean of the population from which the sample has been selected.

- **2.** $V(\overline{X}) = \sigma^2_{\overline{X}} = \sigma^2/n$ and $\overline{X}_{\overline{X}} = \sigma/\sqrt{n}$

- Result 2 shows that the distribution becomes more concentrated about μ as the sample size n increases.

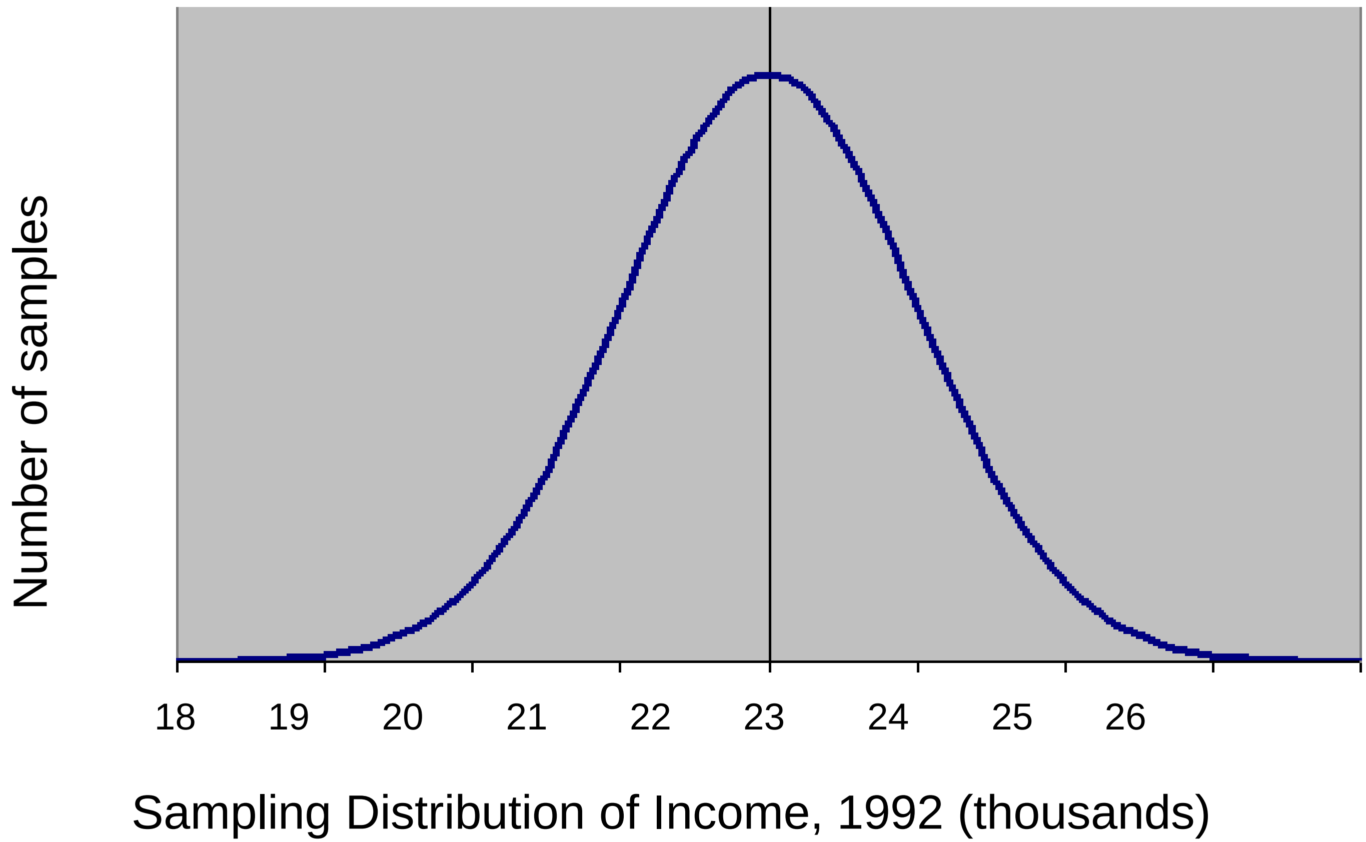
Most empirical distributions are not normal:



U.S. Income distribution 1992

Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

But the sampling distribution of mean income over many samples *is* normal



Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Standard Deviation

Measures how spread out a distribution is.

Square root of the sum of the squared deviations of each case from the mean over the number of cases, or

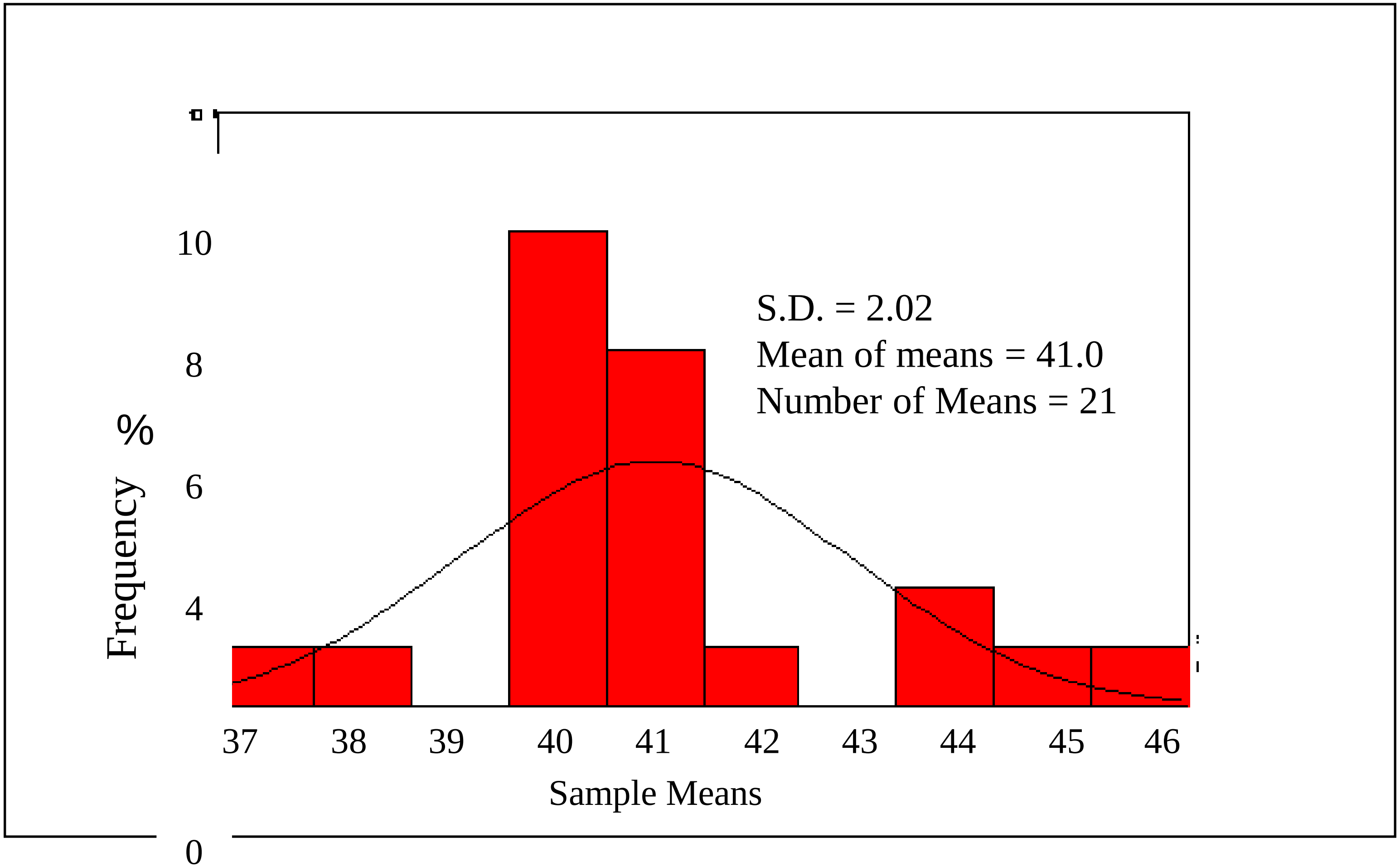
$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}}$$

Example of Standard Deviation

		Deviation from Mean	
Amount	$\bar{X}$	$(X - \bar{X})$	$(X - \bar{X})^2$
600	435	600 - 435 = 165	27,225
350	435	350 - 435 = -85	7,225
275	435	275 - 435 = -160	25,600
430	435	430 -435 = -5	25
520	435	520 - 435 = 85	7,225
		0	67,300

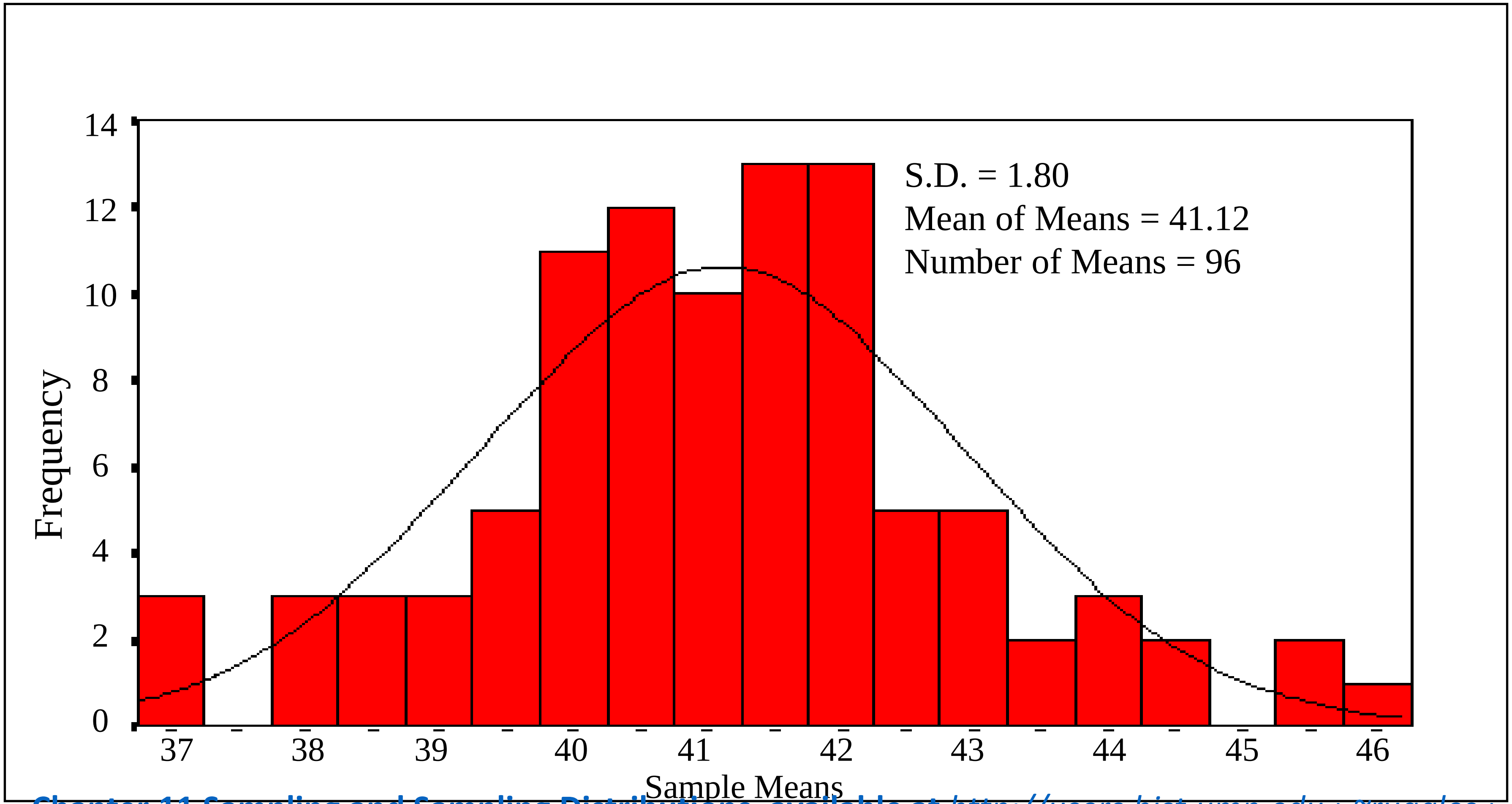
$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}} = \sqrt{\frac{67,300}{4}} = \sqrt{16,825} = 129.71$$

Distribution of Sample Means with 21 Samples



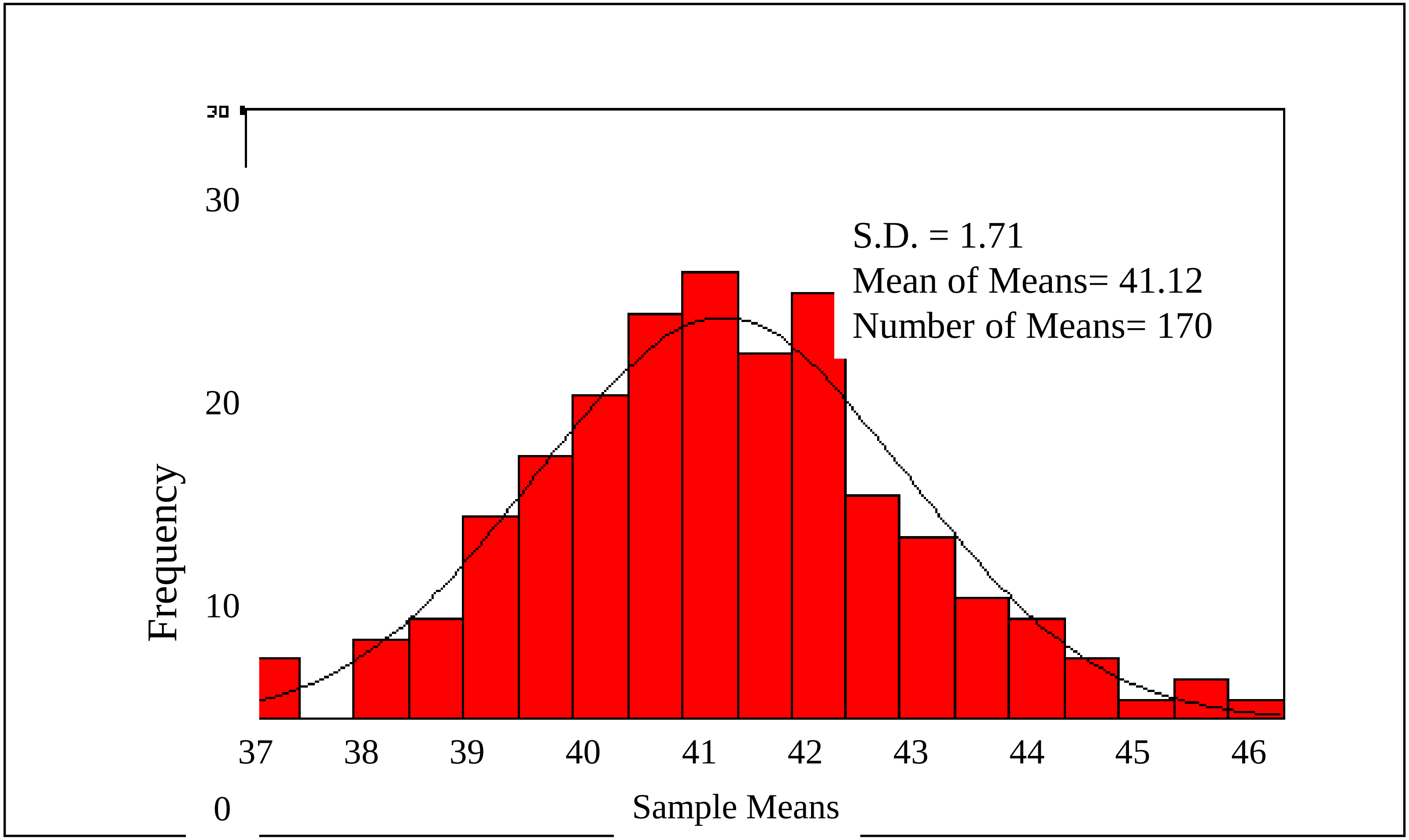
Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Distribution of Sample Means with 96 Samples



Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Distribution of Sample Means with 170 Samples



Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Standard Error

The standard deviation of the sampling distribution is called the standard error.

Standard error can be estimated from a single sample:

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Where

s is the sample standard deviation (i.e., the sample based estimate of the standard deviation of the population), and

n is the size (number of observations) of the sample.

The Distribution of the Sample Mean

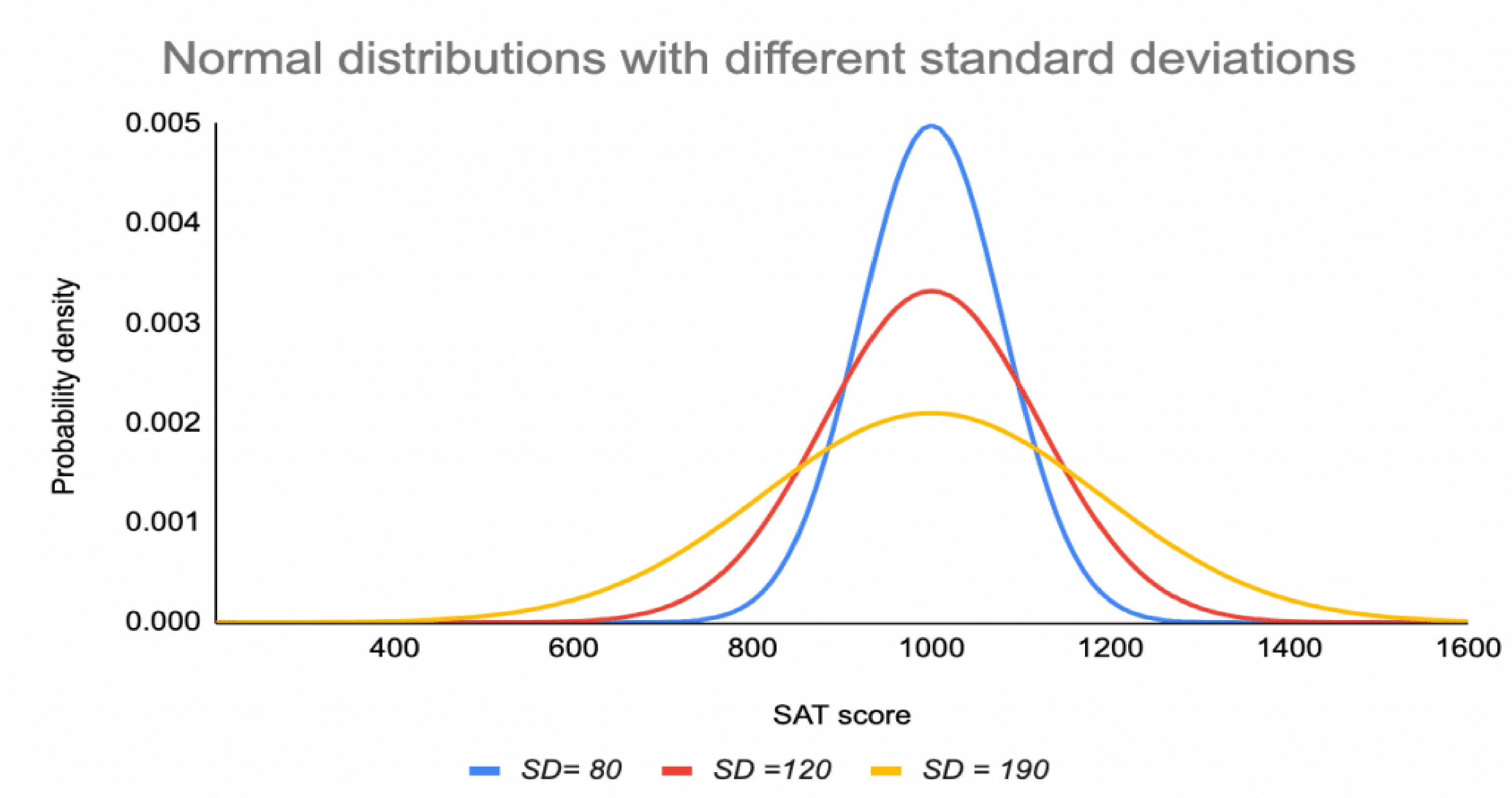
- The importance of the sample mean $\overline{X}$ comes from its use in drawing conclusions about the population mean μ.
- Some of the most frequently used inferential procedures are based on properties of the sampling distribution of $\overline{X}$.

The Distribution of the Sample Mean

$$\sigma_{\overline{X}} = \sigma/\sqrt{n}$$

- The standard deviation is often called the *standard error of the mean;* it describes the magnitude of a typical or representative deviation of the sample mean from the population mean.

Ex: Normal Population Distribution with varying n



A normal population distribution and sampling mean distributions

https://www.scribbr.com/statistics/normal-distribution/

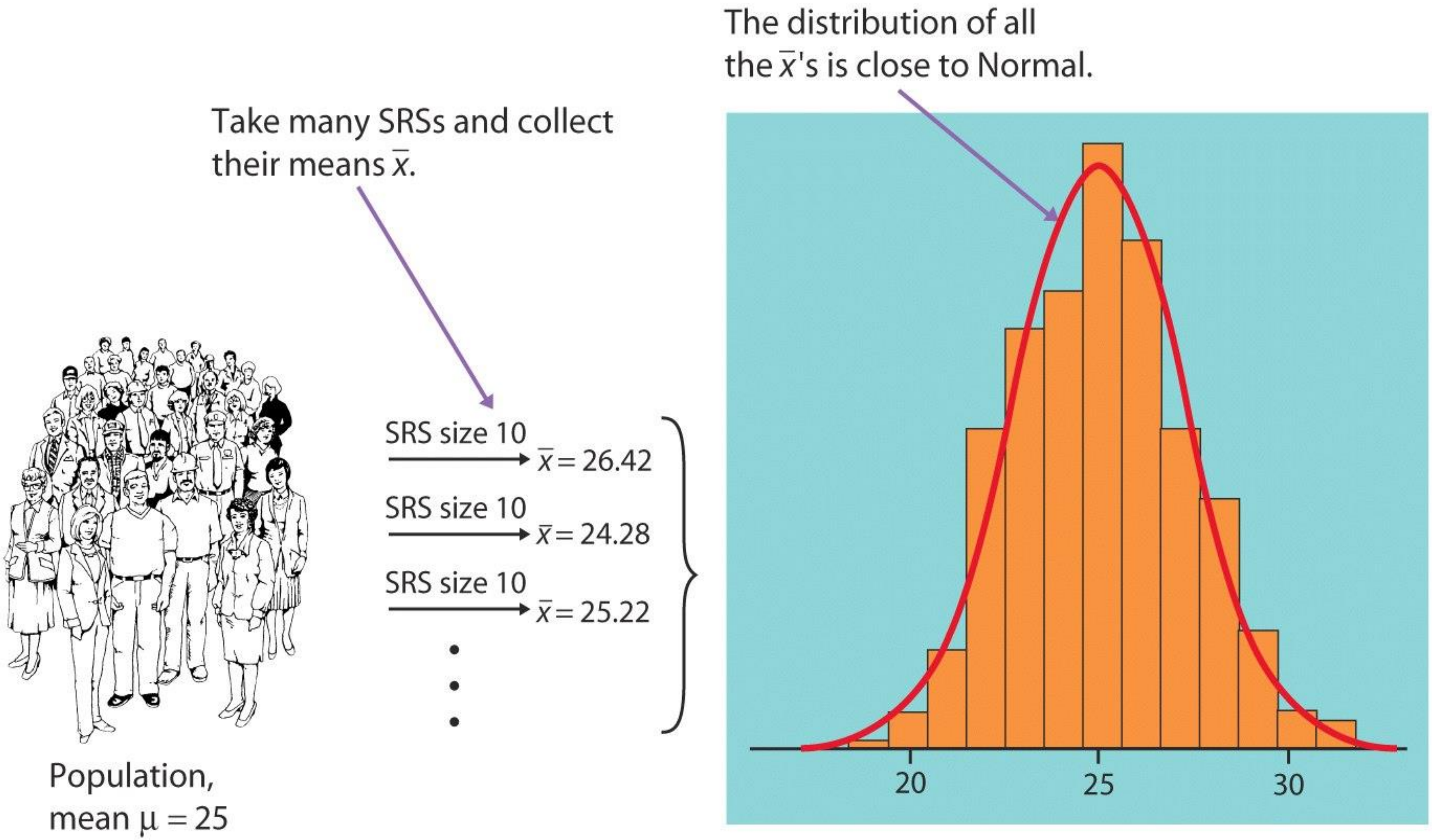
Parameters and Statistics

- **Parameter** ≡ a constant that describes a **population** *or* **probability model**, e.g., μ from a Normal distribution
- **Statistic** ≡ a random variable calculated from a sample e.g., "x-bar"
- These are related but are not the same!
- For example, the average age of the UNITECH student population $\mu = 27.5$ (parameter), but the average age in any sample x-bar (statistic) is going to differ from μ

Sampling Distributions of Statistic

- The sampling distribution of a statistic predicts the behavior of the statistic in the long run
- The next slide show a simulated sampling distribution of mean from a population that has X~N(25, 7). We take 1,000 samples, each of *n* =10, from population, calculate x-bar in each sample and plot.
- **SAMPLING DISTRUBTUION of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population.**

Simulation of a Sampling Distribution of xbar



https://slidetodoc.com/chapter-11-sampling-distributions-hs-67-sampling-distributions/

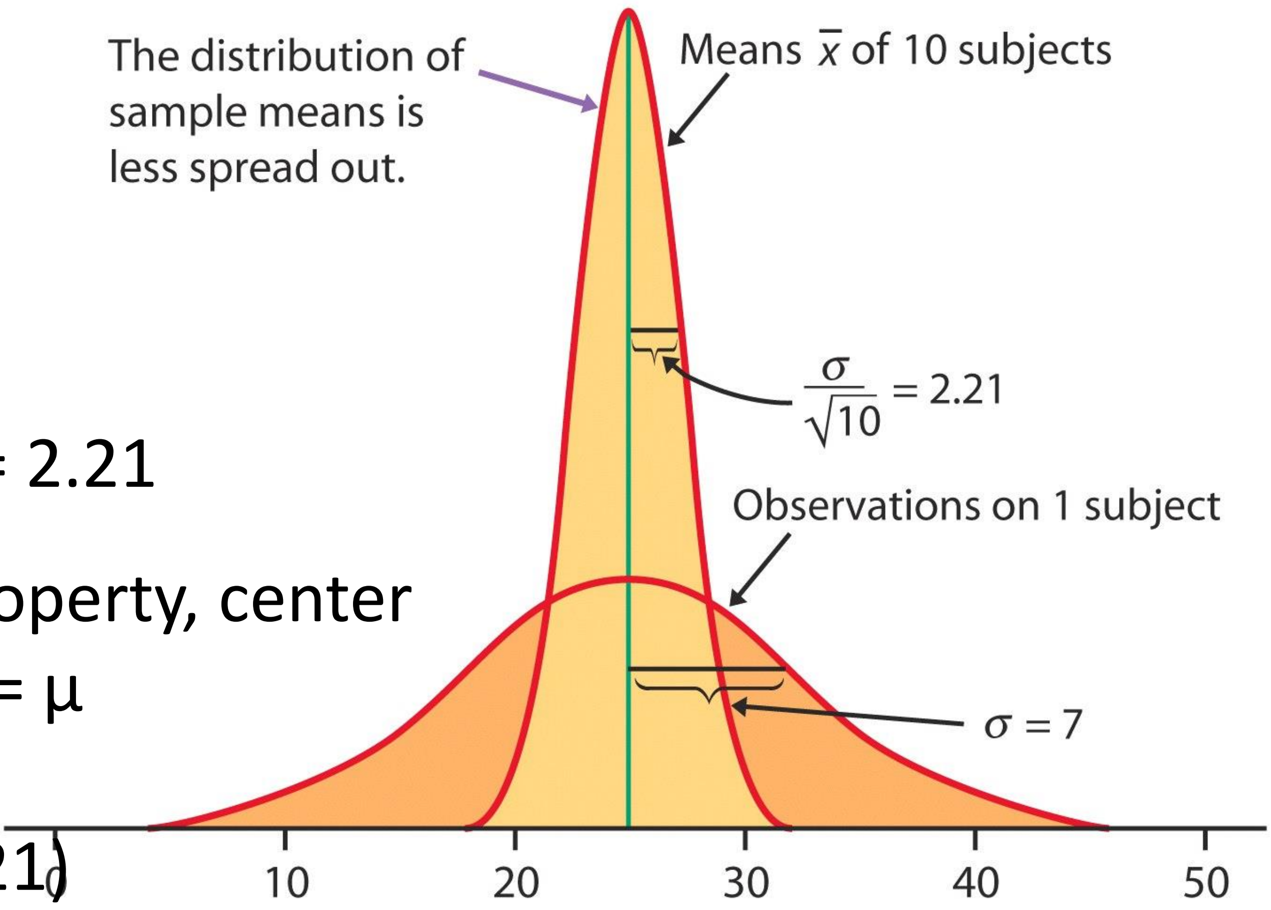
Sampling Distribution of Mean: Example

Population
X~N(25,7)

Sample *n* = 10

By sq. root law,
$\sigma_{xbar} = 7 / \sqrt{10} = 2.21$

By unbiased property, center of distribution = μ
Thus
x-bar~N(25, 2.21)



https://slidetodoc.com/chapter-11-sampling-distributions-hs-67-sampling-distributions/

The Central Limit Theorem

- The CL theorem concerns the distribution of all possible sample means, *x bar of sample size n from a population* with mean μ.
- It can also be applied to the distribution of all possible sample pro*portions, p hat, o*f sample size *n from a population with proportion p.*
- *Note that a distribution of sample values always* contains all possible samples of the same size, *n.*
- The central limit theorem has three main concepts:

The Central Limit Theorem

- 1. As the sample size*, n, increases, the distribution* of all possible *x bar (p hat)* values becomes increasingly closer to a normal distribution, regardless of the population shape.
- 2. The mean of all the possible sample values *x bar (p hat)* of sample size *n equals the population* parameter, (or *p).*
- 3. The standard deviation of the distribution of the sample statistics decreases as n increases.

Law of Large Numbers

- Draw observations at random from any population with finite mean μ.
- As the number of observations drawn increases, the mean xbar of the observed values get closer and closer to the mean μ of the population.

The Central Limit Theorem

- When the X_i's are normally distributed, so is $\overline{X}$ for every sample size *n*.
- Even when the population distribution is highly non-normal, averaging produces a distribution more bell-shaped than the one being sampled.
- A reasonable conjecture is that if *n* is large, a suitable normal curve will approximate the actual distribution of $\overline{X}$ The formal statement of this result is the most important theorem of probability.

The Central Limit Theorem

- The CLT provides insight into why many random variables have probability distributions that are approximately normal.
- A practical difficulty in applying the CLT is in knowing when n is sufficiently large.
- The problem is that the accuracy of the approximation for a particular n depends on the shape of the original underlying distribution being sampled.

The Central Limit Theorem

- If the underlying distribution is close to a normal density curve, then the approximation will be good even for a small *n*, whereas if it is far from being normal, then a large *n* will be required.

- **Rule of Thumb**
  If $n > 30$, the Central Limit Theorem can be used.

- There are population distributions for which even an *n* of 40 or 50 does not suffice, but such distributions are rarely encountered in practice.

Some rules about the sampling distribution of the mean…

1. For a random sample of size *n* from a population having mean μ and standard deviation σ, the sampling distribution of X-bar has mean μ and standard error $\sigma_{X\text{-bar}} = \sigma/\sqrt{n}$
2. The Central Limit Theorem says that for random sampling, as the sample size *n* grows, the sampling distribution of X-bar approaches a normal distribution.
3. The sampling distribution will be normal *no matter what* the population distribution's shape as long as $n > 30$.

Some rules about the sampling distribution of the mean…

4. If $n < 30$, the sampling distribution is likely normal only if the underlying population's distribution is normal.
5. As n increases, the standard error (remember that this word means standard deviation of the sampling distribution) gets smaller.
6. Precision provided by any given sample increases as sample size n increases.

A Sampling Distribution

So we know in advance of ever collecting a sample, that if sample size is sufficiently large:

- Repeated samples would pile up in a normal distribution
- The sample means will center on the true population mean
- The standard error will be a function of the population variability and sample size
- The larger the sample size, the more precise, or efficient, a particular sample is
- 95% of all sample means will fall between +/- 2 s.e. from the population mean

What is probability?

- Probability defines the likelihood of occurrence of an event.
- Many real-life situations in which we may have to predict the outcome of an event. We may be sure or not sure of the results of an event.
- In such cases, we say that there is a probability of this event to occur or not occur.
- Probability generally has great applications in games, in business to make probability-based predictions, and also probability has extensive applications in this new area of artificial intelligence.
- Probability can be defined as the ratio of the number of favorable outcomes to the total number of outcomes of an event. For an experiment having 'n' number of outcomes, the number of favorable outcomes can be denoted by x.

What is probability?

- The formula to calculate the probability of an event is as follows.
- Probability(Event) = Favorable Outcomes/Total Outcomes = x/n
- Example: Suppose we have to predict about the happening of rain or not.
- The answer to this question is either "Yes" or "No".
- There is a likelihood to rain or not rain. Here we can apply probability.
- Probability is used to predict the outcomes for the tossing of coins, rolling of dice, or drawing a card from a pack of playing cards.

Terms in Probability

- **Experiment:** A trial or an operation conducted to produce an outcome is called an experiment.
- **Sample Space:** All the possible outcomes of an experiment together constitute a sample space. For example, the sample space of tossing a coin is head and tail.
- **Favorable Outcome:** An event that has produced the desired result or expected event is called a favorable outcome. For example, when we roll two dice, the possible/favorable outcomes of getting the sum of numbers on the two dice as 4 are (1,3), (2,2), and (3,1).
- **Trial:** A trial denotes doing a random experiment.

Terms in Probability

- **Random Experiment:** An experiment that has a well-defined set of outcomes is called a random experiment. For example, when we toss a coin, we know that we would get ahead or tail, but we are not sure which one will appear.
- **Event:** The total number of outcomes of a random experiment is called an event.
- **Equally Likely Events:** Events that have the same chances or probability of occurring are called equally likely events. The outcome of one event is independent of the other. For example, when we toss a coin, there are equal chances of getting a head or a tail.
- **Exhaustive Events:** When the set of all outcomes of an experiment is equal to the sample space, we call it an exhaustive event.
- Mutually Exclusive Events: Events that cannot happen simultaneously are called mutually exclusive events. For example, the climate can be either hot or cold. We cannot experience the same weather simultaneously.

Probability Formula

cuemath THE MATH EXPERT

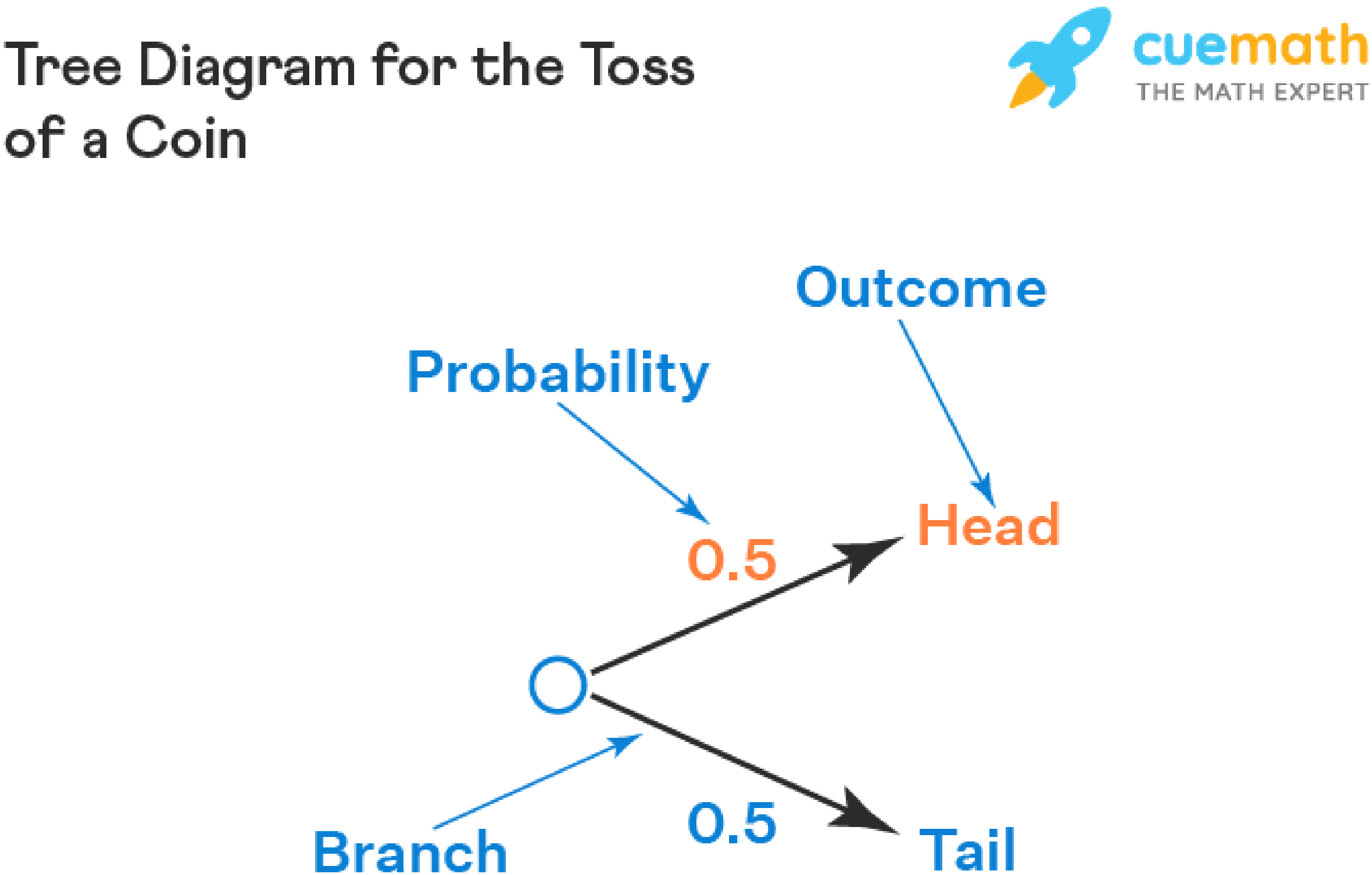
$$P(A) = \frac{\text{Number of favorable outcomes to A}}{\text{Total number of possible outcomes}}$$

Probability Examples: Tossing a Coin

- A single coin on tossing has two outcomes, a head, and a tail.
- Total number of possible outcomes = 2; Sample Space = {H, T}; H: Head, T: Tail
- P(H) = Number of heads/Total outcomes = 1/2
- P(T)= Number of Tails/ Total outcomes = ½
- While tossing two coins, we have a total of four outcomes
- Total number of outcomes = 4; Sample Space = {(H, H), (H, T), (T, H), (T, T)}
- P(2H) = P(0 T) = Number of outcome with two heads/Total Outcomes = 1/4
- P(1H) = P(1T) = Number of outcomes with only one head/Total Outcomes = 2/4 = 1/2
- P(0H) = (2T) = Number of outcome with two heads/Total Outcomes = 1/4

Probability Examples: Tossing 3 Coins

- The number of total outcomes on tossing three coins simultaneously is equal to $2^3 = 8$.
- For these outcomes, we can find the probability of getting one head, two heads, three heads, and no head.
- A similar probability can also be calculated for the number of tails.
- Total number of outcomes = $2^3 = 8$ Sample Space = {(H, H, H), (H, H, T), (H, T, H), (T, H, H), (T, T, H), (T, H, T), (H, T, T), (T, T, T)}
- P(0H) = P(3T) = Number of outcomes with no heads/Total Outcomes = 1/8
- P(1H) = P(2T) = Number of Outcomes with one head/Total Outcomes = 3/8
- P(2H) = P(1T) = Number of outcomes with two heads /Total Outcomes = 3/8
- P(3H) = P(0T) = Number of outcomes with three heads/Total Outcomes = 1/8



https://www.cuemath.com/data/probability/

Probability distributions: Permutations

What is the probability distribution of number of girls in two families with two children?

2 GG

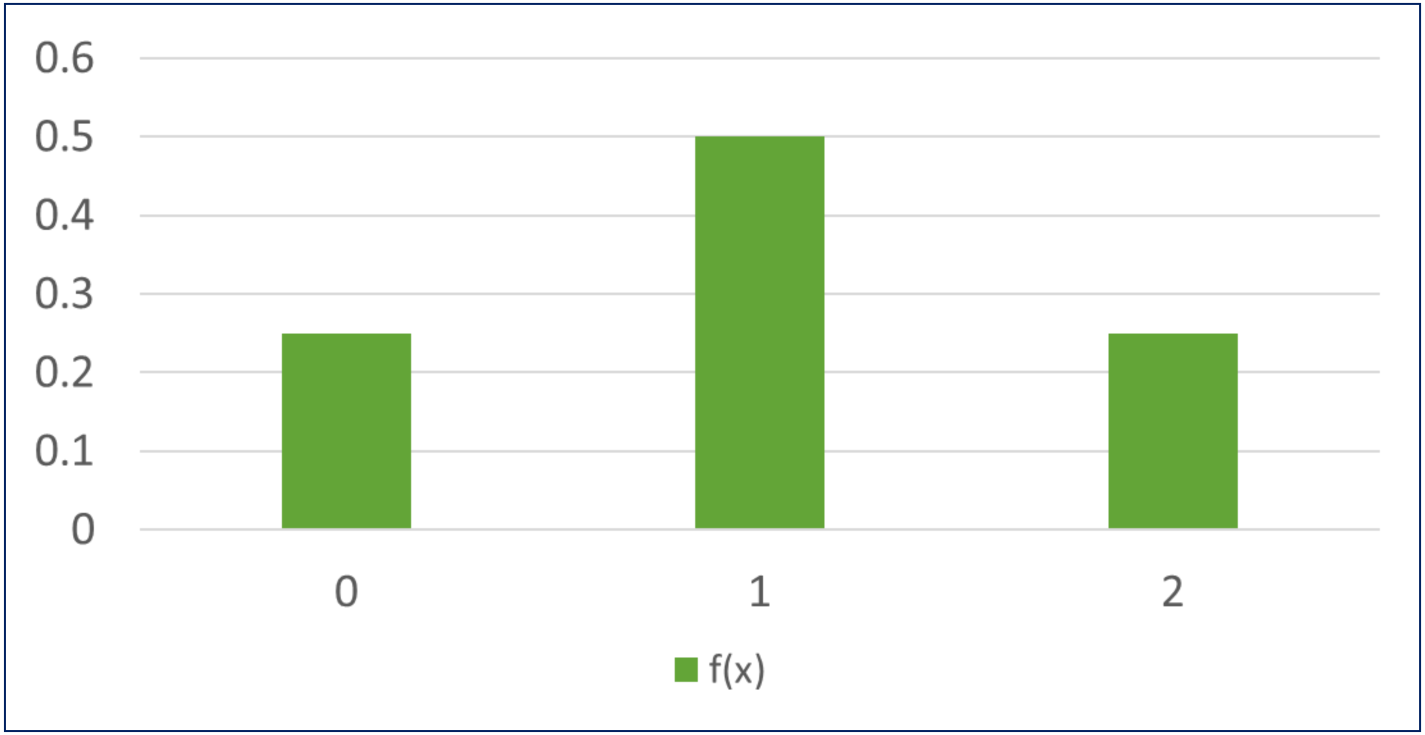
1 BG

1 GB

0 BB

	x	f(x)
GG	0	0.25
BG, GB	1	0.5
BB	2	0.25

Probability Distribution of Number of Girls

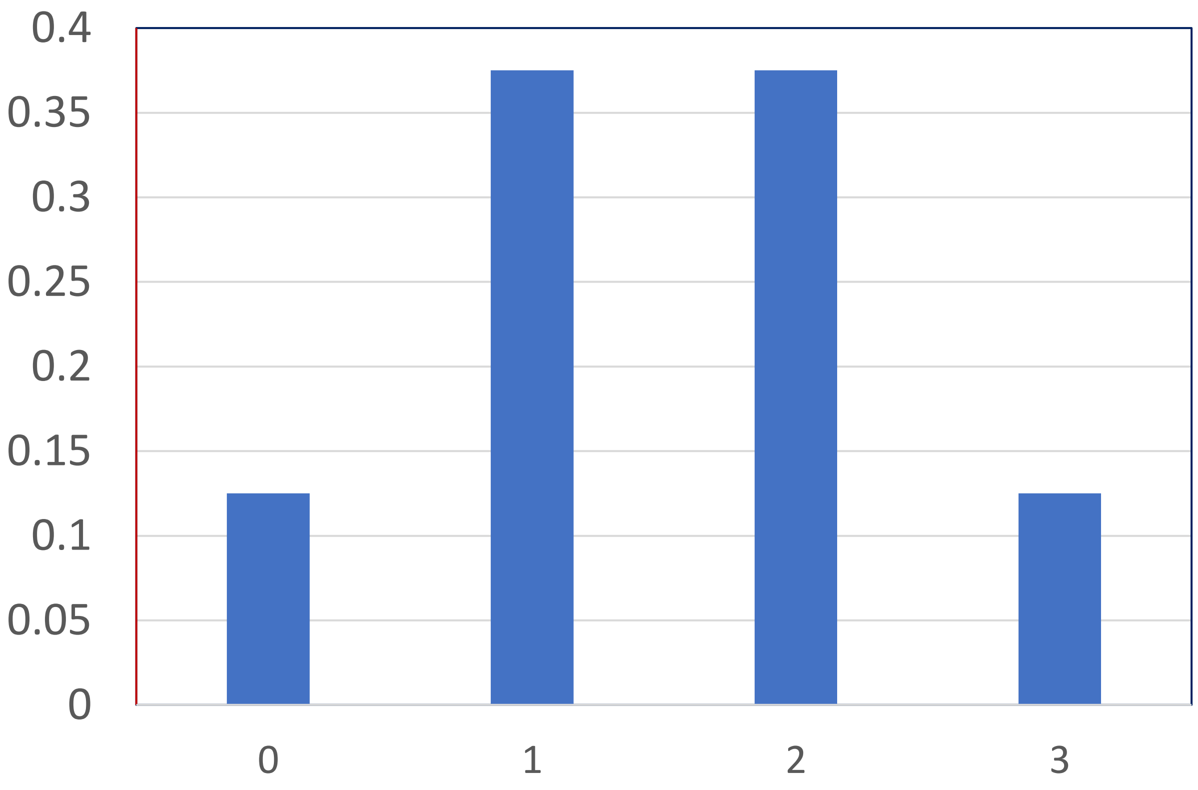


How about family of three?

What is the probability distribution of number of girls in three families with two children?

Num. Girls	child #1	child #2	child #3
0	B	B	B
1	B	B	G
1	B	G	B
1	G	B	B
2	B	G	G
2	G	B	G
2	G	G	B
3	G	G	G

Probability distribution of number of girls in three families



x	f(x)
0	0.125
1	0.375
2	0.375
3	0.125

Combinations: nCr

The **factorial function** (symbol: **!**) means to multiply a series of descending natural numbers.

Examples:

- 4! = 4 × 3 × 2 × 1 = 24
- 7! = 7 × 6 × 5 × 4 × 3 × 2 × 1 = 5,040
- 1! = 1

Note: it is generally agreed that **0! = 1**. It may seem funny that multiplying no numbers together gets 1, but it helps simplify a lot of equations.

Combinations: nCr

Dennis is ordering pizza. He is planning to order two toppings on his pizza. If the pizza place he is ordering from has 12 toppings to choose from, how many different combinations of 2 toppings can he order?

- In this example n = 12, r =2

$$\frac{n!}{r!(n-r)!} = \frac{12!}{2!(12-2)!} = \frac{12!}{2!(10!)} =$$

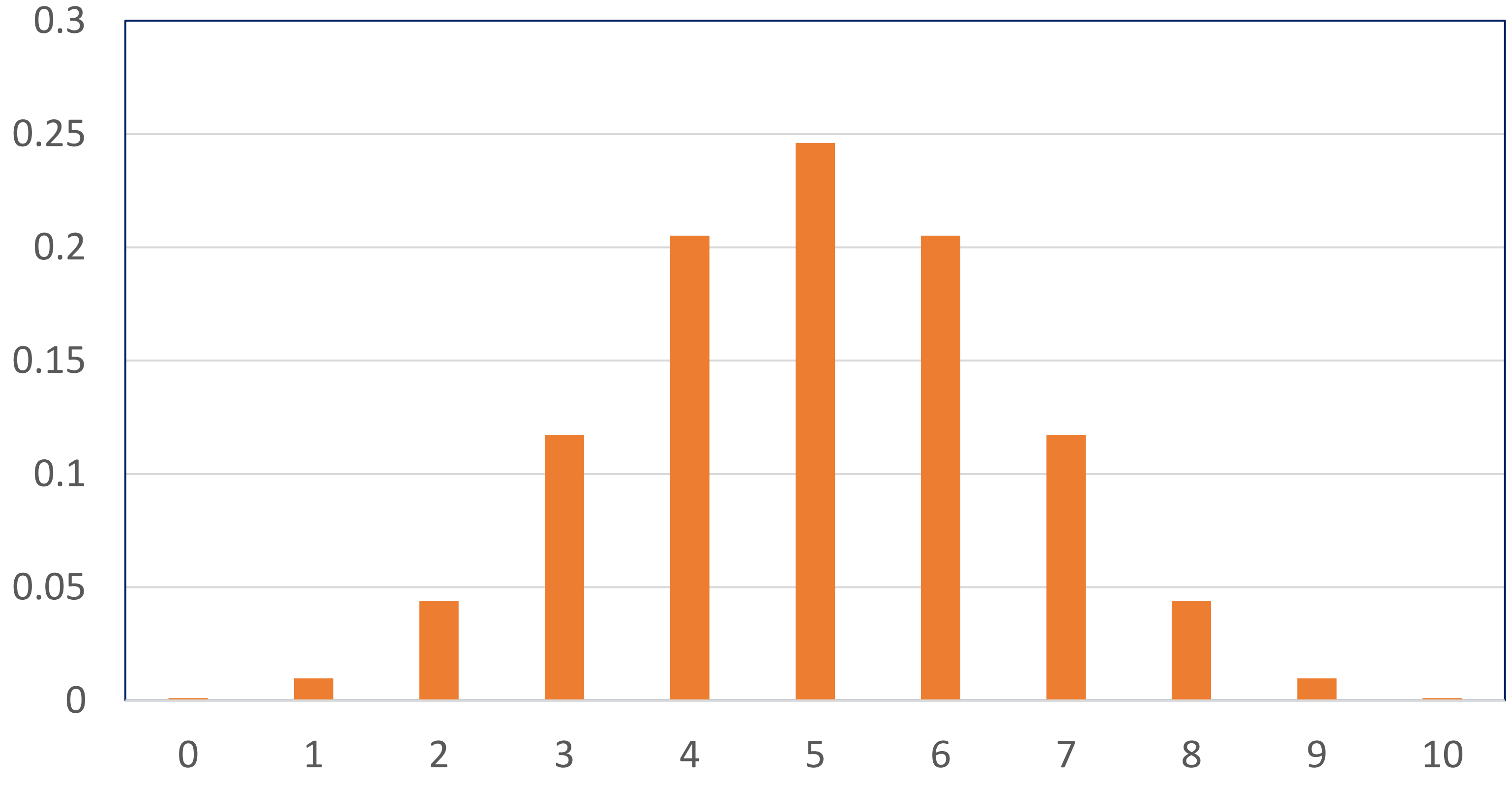
$$\frac{12*11*10*9*8*7*6*5*4*3*2*1}{2*1*10*9*8*7*6*5*4*3*2*1} = \frac{12*11}{2*1} = \frac{132}{2} = 66$$

How about a family of 10?

What is the probability distribution of number of girls in ten families with two children?

x	ncr	f (x)
0	1	0.00098
1	10	0.00977
2	45	0.04395
3	120	0.11719
4	210	0.20508
5	252	0.24609
6	210	0.20508
7	120	0.11719
8	45	0.04395
9	10	0.00977
10	1	0.00098
	1024	1.00000

What is the probability distribution of number of girls in two families with two children?



Probability Density Functions

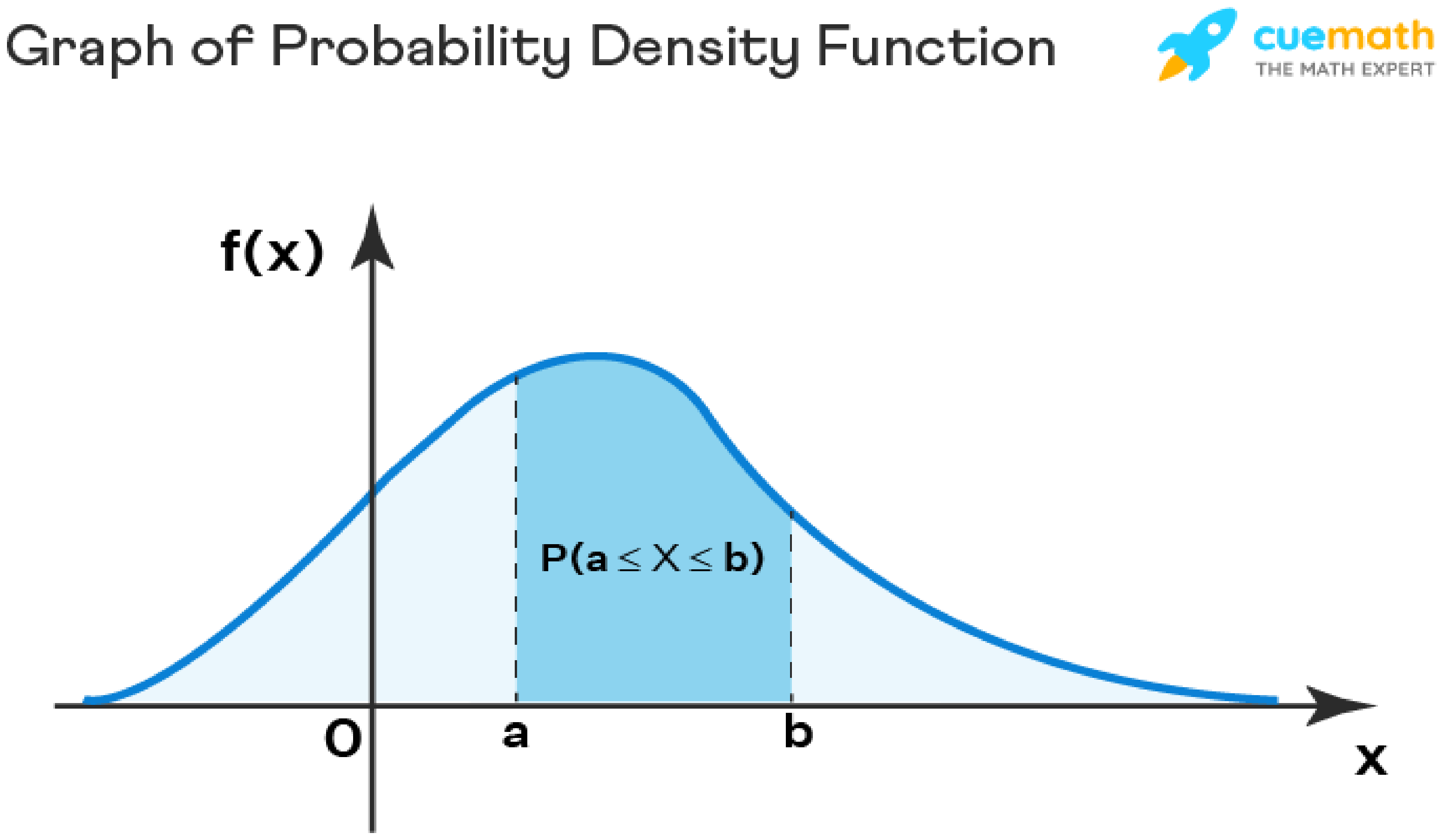
- Probability density function is a function that provides the likelihood that the value of a random variable will fall between a certain range of values.
- We use the probability density function in the case of continuous random variables. For discrete random variables, we use the probability mass function which is analogous to the probability density function.
- The graph of a probability density function is in the form of a bell curve. The area that lies between any two specified values gives the probability of the outcome of the designated observation.
- We solve the integral of this function to determine the probabilities associated with a continuous random variable.

What is Probability Density Function?

- Probability density function and cumulative distribution function are used to define the distribution of continuous random variables.
- If we differentiate the cumulative distribution function of a continuous random variable it results in the probability density function.
- Conversely, on integrating the probability density function we get the cumulative distribution function.

Definition of Probability Density Function

- Probability density function defines the density of the probability that a continuous random variable will lie within a particular range of values.
- To determine this probability, we integrate the probability density function between two specified points.
- Normal distribution is the most widely used type of continuous probability distribution.
- The notation for normal distribution is given as $X \sim N(\mu, \sigma 2)$



https://www.cuemath.com/data/probability-density-function/

Important Notes on Probability Density Function

- Probability density function determines the probability that a continuous random variable will fall between a range of specified values.
- On differentiating the cumulative distribution function, we obtain the probability density function.
- The mean of the probability density function can be give as $E[X] = \mu = \int \infty - \infty x f(x) dx$
- The variance of a probability density function is given by $Var(X) = \sigma 2 = \int \infty - \infty (x - \mu) 2 f(x) dx$

Animated Video on Understanding CLT

- By Dr. Nick under Statistics Learning Centre
- Where the coverage includes four aspects
- 1. Mean of the Sampling distribution
- 2. Sampling distribution tends to follow normal distribution
- 3. spread of the sampling distribution is related to the spread of the population
- 4. Bigger the size of n, smaller the spread or variance in the distribution
- https://www.youtube.com/watch?v=_YOr_yYPytM from Statistics learning Centre

Reference

Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Central Limit Theorem available at *https://pdp.sjsu.edu › gerstman › Chapt11_BPS*

https://www.youtube.com/watch?v=_YOr_yYPytM from Statistics learning Centre

https://www.cuemath.com/data/probability-density-function/

https://medium.com/@alb.arganese/the-most-important-families-of-theoretical-distribution-bd6974c5d323

https://personal.utdallas.edu/~scniu/OPRE-6301/documents/Important_Probability_Distributions.pdf

https://www.cuemath.com/data/probability/

What next?

- Some Important Theoretical Probability Distributions viz., Normal distribution, t, F and **Chi-squared distribution**
- Hypothesis testing