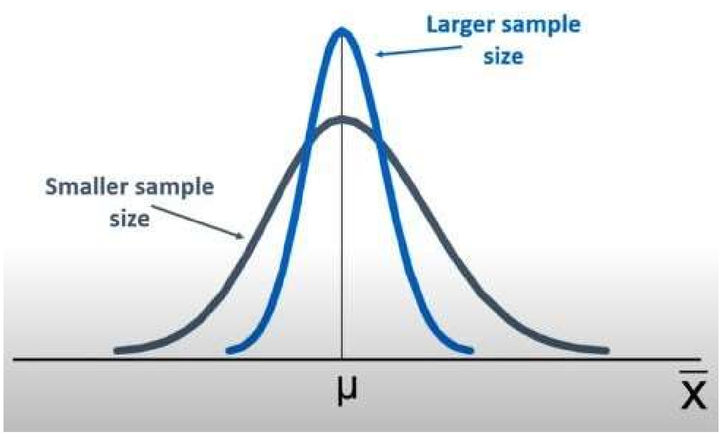
- *It states that the sampling distribution of the sample means approaches a normal distribution as the sample size gets larger.*
- *Holds True when sample size > 30.*

Essential component of the Central Limit Theorem is that the average of your sample means will be the population mean.

Relationship between shape of population distribution and shape of sampling distribution of mean is known as Central Limit theorem

As the sample size get large, sampling distribution becomes almost Normal regardless of shape of population



$$Z = \frac{X - \mu}{\sigma_x} \qquad \sigma_x = \frac{\sigma}{\sqrt{n}}$$

Where,
μ = Population mean
σ = Population standard deviation
$\mu_{\bar{x}}$ = Sample mean
$\sigma_{\bar{x}}$ = Sample standard deviation
n = Sample size

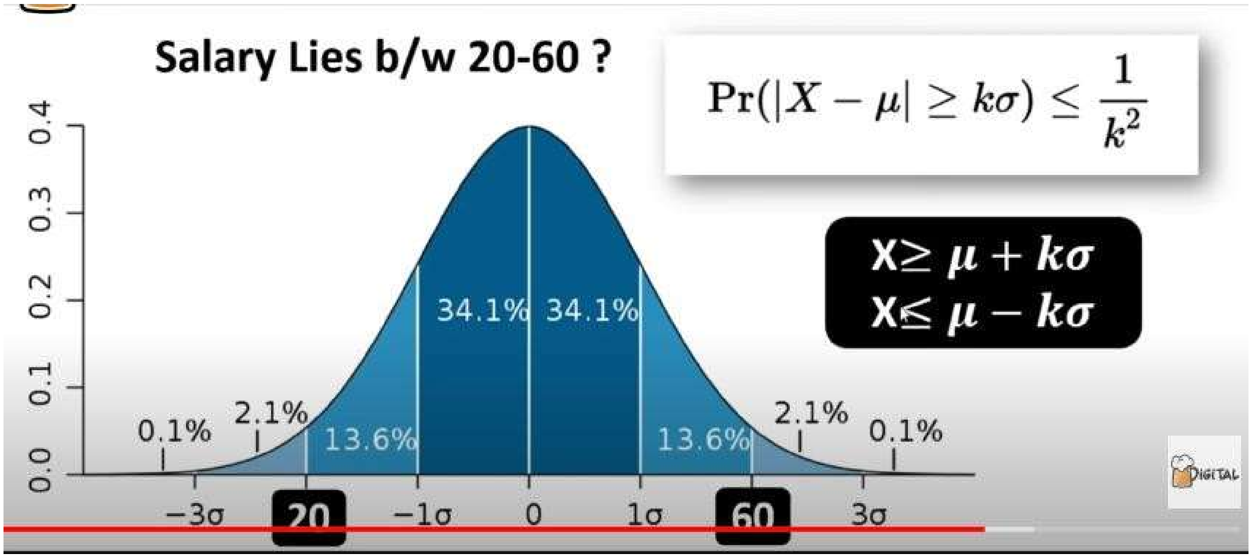
The Central Limit Theorem (CLT) is a fundamental theorem in probability theory and statistics. It states that, given a sufficiently large sample size, the distribution of the sample mean will approach a normal distribution (also known as a Gaussian distribution), regardless of the original distribution of the population from which the sample is drawn. This is true provided that the samples are independent and identically distributed (i.i.d.) and that the population has a finite mean and variance.

Key Points of the Central Limit Theorem

1. **Sample Size:** The sample size should be large enough, typically $n \geq 30$ is considered sufficient in many practical situations.
2. **Independence:** The samples must be independent of each other.
3. **Identically Distributed:** Each sample should be drawn from the same population with the same probability distribution.
4. **Finite Mean and Variance:** The population from which the samples are drawn must have a finite mean (μ) and variance (σ^2).

Implications of the CLT

- **Normal Distribution of Sample Means:** Regardless of the shape of the population distribution, the distribution of the sample means will tend to be normal if the sample size is large enough.
- **Standard Error:** The standard deviation of the sampling distribution of the sample mean (standard error) decreases as the sample size increases. It is given by $\sigma/\sqrt{n}$, where σ is the population standard deviation and n is the sample size.



The Chebyshev inequality is a fundamental result in probability theory and statistics that gives an upper bound on the probability that the value of a random variable deviates from its mean. Specifically, for a random variable X with finite mean μ and finite variance σ^2, the Chebyshev inequality states that for any $k > 0$:

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

This inequality can be interpreted as follows: the probability that the value of the random variable X differs from its mean μ by at least k standard deviations is at most $\frac{1}{k^2}$.

Applications and Implications

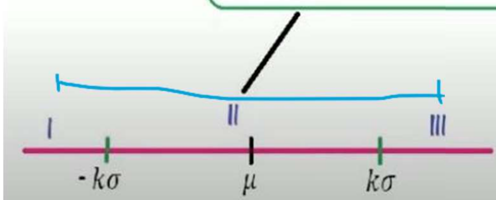
1. **General Probability Bounds:** Chebyshev's inequality provides a general bound that applies regardless of the underlying distribution of the random variable, as long as the mean and variance are finite. This makes it a very powerful and widely applicable tool.
2. **Concentration of Measure:** It shows that for a large class of distributions, the probability that a random variable deviates significantly from its mean is small. This principle is useful in many areas, such as quality control, risk management, and any field where understanding the variability around the mean is important.
3. **Law of Large Numbers:** Chebyshev's inequality can be used to provide a simple proof of the Weak Law of Large Numbers, which states that the sample average of a large number of independent and identically distributed random variables converges in probability to the expected value.

Limitations

While Chebyshev's inequality is very general, it is often not tight, meaning that the actual probability of deviation may be much smaller than the bound provided by the inequality. More specific distributions (e.g., normal distribution) have much sharper bounds (e.g., the empirical rule for normal distributions).

Let X be a random variable with mean μ and finite variance σ^2, then for any real number $k > 0$,

$$\boxed{P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}}$$



OR

$$\boxed{P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}}$$

Chebyshev's Inequality is a statistical theorem that provides a lower bound on the probability that the values of a random variable lie within a certain number of standard deviations from the mean. Specifically, it states that for any random variable with a finite mean μ and finite standard deviation σ, the probability that the variable lies within k standard deviations of the mean is at least $1 - \frac{1}{k^2}$ for any $k > 1$.

Example

Suppose we have a dataset with the following characteristics:

- Mean (μ): 50
- Standard Deviation (σ): 10

We want to apply Chebyshev's Inequality to determine the probability that a data point lies within k standard deviations from the mean.

Let's choose $k = 3$.

Applying Chebyshev's Inequality

Chebyshev's Inequality states:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

This can be restated as:

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

For $k = 3$:

$$P(|X - 50| < 3 \times 10) \geq 1 - \frac{1}{3^2} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0.89$$

So, Chebyshev's Inequality tells us that at least 89% of the data points lie within 3 standard deviations of the mean.

Calculating the Range

We now calculate the range within which 89% of the data points should lie:

$$\mu - 3\sigma \leq X \leq \mu + 3\sigma$$

Substituting the values:

$$50 - 3 \times 10 \leq X \leq 50 + 3 \times 10$$

$$20 \leq X \leq 80$$

Interpretation

According to Chebyshev's Inequality, at least 89% of the data points in our dataset will lie between 20 and 80. This gives us a way to understand the spread of the data around the mean, even if we do not know the exact distribution of the data.

Summary

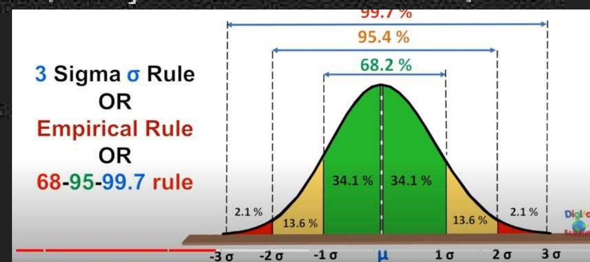
Chebyshev's Inequality is a powerful tool for understanding the dispersion of data. In this example, with a mean of 50 and a standard deviation of 10, the inequality guarantees that at least 89% of the data points lie within the range of 20 to 80, which is 3 standard deviations from the mean. This can be very useful when dealing with unknown or non-normal distributions.

Continuous Distributions

Continuous distributions describe variables that can take on an infinite number of possible values within a given range. These distributions are characterized by a smooth probability density function (PDF). Here are some examples:

1. Normal Distribution

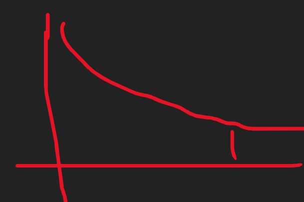
- **Description:** Also known as the Gaussian distribution, it is symmetric about the mean, with a bell-shaped curve.
- **Parameters:** Mean (μ) and standard deviation (σ).
- **PDF:** [redacted] $z = \left(\frac{x-\mu}{\sigma}\right)$
- **Example:** Heights of people, test scores.



1^{st} standard devialtion 2^{nd} and 3rd

2. Exponential Distribution

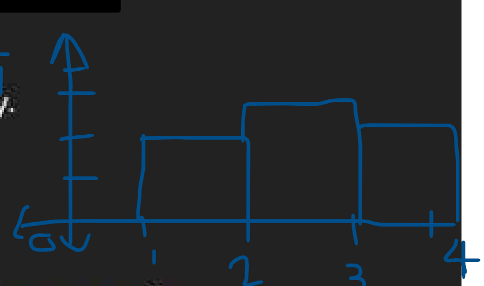
- **Description:** Describes the time between events in a Poisson process.
- **Parameter:** Rate parameter (λ). lambda
- **PDF:** $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$
- **Example:** Time between arrivals of buses.



Time per single event || the event must be constant and independent

3. Uniform Distribution

- **Description:** All outcomes in the range [a, b] are equally likely.
- **Parameters:** Lower bound (a) and upper bound (b).
- **PDF:** $f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
- **Example:** Rolling a fair die (continuous if we consider a continuous range).



Probability distributions with equally likely outcomes

Aspect	Discrete Distribution	Continuous Distribution
Variable Type	Takes on countable, distinct values (often integers)	Takes on an infinite number of values within a range
Probability of Specific Value	Non-zero for specific values (e.g., $P(X = k)$)	Zero for specific values; probability is defined over intervals
Common Distributions	Binomial, Poisson, Geometric, Negative Binomial	Normal, Exponential, Uniform, Gamma
Examples	Number of heads in coin flips, number of customers in a store	Time between events, heights of people

Discrete Distributions

Discrete distributions describe variables that can take on a finite or countably infinite number of possible values. These distributions are characterized by a probability mass function (PMF). Here are some examples:

1. Binomial Distribution

- **Description:** Describes the number of successes in a fixed number of independent Bernoulli trials.
- **Parameters:** Number of trials (n) and probability of success (p).
- **PMF:**
- **Example:** Number of heads in 10 coin flips.

The binomial distribution function is calculated as:

$$P_{(x:n,p)} = {}^{n}C_{x}\, p^{x} (1-p)^{n-x}$$

Where:

n is the number of trials (occurrences)

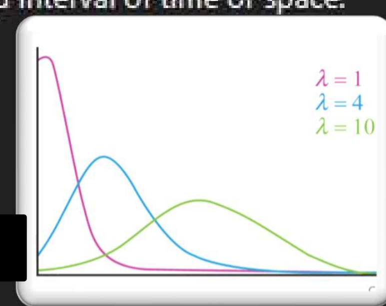
x is the number of successful trials

p is the probability of success in a single trial

${}^{n}C_{x}$ is the combination of n and x.

2. Poisson Distribution

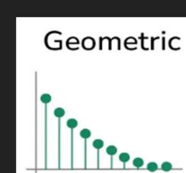
- **Description:** Describes the number of events occurring in a fixed interval of time or space.
- **Parameter:** Rate parameter (λ).
- **PMF:** $P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$
- **Example:** Number of emails received in an hour.



• **Lam: number of occurrences** • **Size: The shape of returned ary (it is count distributi)**

3. Geometric Distribution

- **Description:** Describes the number of trials until the first success in a series of Bernoulli trials.
- **Parameter:** Probability of success (p).
- **PMF:** $P(X=k) = (1-p)^{k-1}p$
- **Example:** Number of coin flips until the first heads.



<u>**Bernoulli distribution**:</u> Models a single trial with two possible outcomes (success or failure).

Parameter estimation is a statistical technique used to determine the values of parameters in a statistical model based on observed data. The goal is to find the best-fit values for the parameters so that the model accurately represents the underlying data distribution or relationship. Here's a basic outline of the process:

1. **Model Specification:** Define the statistical model that describes the relationship between the observed data and the parameters. This model could be linear, nonlinear, probabilistic, etc.
2. **Collect Data:** Gather the observed data that will be used to estimate the parameters.
3. **Choose an Estimation Method:** Select a method for estimating the parameters. Common methods include:
   - **Maximum Likelihood Estimation (MLE):** Finds the parameter values that maximize the likelihood function, which measures how likely the observed data is given the parameter values.
   - **Least Squares Estimation:** Minimizes the sum of the squared differences between observed values and the values predicted by the model.
   - **Bayesian Estimation:** Incorporates prior beliefs about the parameters and updates them with the observed data to produce a posterior distribution.
4. **Compute Estimates:** Apply the chosen method to the data to compute the parameter estimates.
5. **Evaluate Estimates:** Assess the quality of the estimates using techniques like confidence intervals, hypothesis testing, or goodness-of-fit measures.
6. **Refine Model:** If necessary, refine the model or estimation method based on the evaluation results to improve accuracy.

In essence, parameter estimation helps to fit a model to data so that predictions or inferences can be made about the system or phenomenon being studied.

Descriptive Statistics are used to summarize and describe the main features of a data set in a quantitative manner. They provide a way to present data in a clear and concise format, making it easier to understand the overall patterns and characteristics of the data. Here's a closer look at the key components:

1. **Measures of Central Tendency:**
   - **Mean:** The average of all data points. It is calculated by summing all values and dividing by the number of values.
     - Formula: $\text{Mean} = \frac{\sum x_i}{N}$
   - **Median:** The middle value in a data set when the values are ordered from smallest to largest. If there is an even number of observations, the median is the average of the two middle numbers.
   - **Mode:** The value that appears most frequently in the data set. A data set can have one mode, more than one mode, or no mode if all values occur with the same frequency.
2. **Measures of Dispersion:**
   - **Range:** The difference between the highest and lowest values in the data set.
     - Formula: $\text{Range} = \text{Max} - \text{Min}$
   - **Variance:** A measure of how spread out the data points are around the mean. It is the average of the squared differences from the mean.
     - Formula: $\text{Variance} = \frac{\sum (x_i - \bar{x})^2}{N}$
   - **Standard Deviation:** The square root of the variance. It provides a measure of the average distance of each data point from the mean.
     - Formula: $\text{Standard Deviation} = \sqrt{\text{Variance}}$
3. **Measures of Distribution Shape:**
   - **Skewness:** Indicates the asymmetry of the data distribution. Positive skewness means the distribution tail is longer on the right, while negative skewness means the tail is longer on the left.
   - **Kurtosis:** Describes the "tailedness" of the distribution. High kurtosis means more data is in the tails, while low kurtosis means the data is more evenly distributed.
4. **Percentiles and Quartiles:**
   - **Percentiles:** Values below which a certain percentage of observations fall. For example, the 25th percentile is the value below which 25% of the data falls.
   - **Quartiles:** Specific percentiles that divide the data set into four equal parts:
     - **First Quartile (Q1):** 25th percentile
     - **Second Quartile (Q2):** 50th percentile (median)
     - **Third Quartile (Q3):** 75th percentile

Descriptive statistics provide a snapshot of the data set, helping to summarize and communicate its essential features without making predictions or inferences about a larger population.

Graphical Statistics involve the use of visual tools to represent and interpret data. These visual representations help in understanding patterns, trends, and relationships within the data more intuitively than numerical summaries alone. Here's an overview of common graphical statistics methods:

1. **Histograms:**
   - **Purpose**: To show the distribution of a continuous variable.
   - **Description:** A histogram displays data by dividing it into intervals (bins) and plotting the frequency (count) of data points within each interval. The height of each bar represents the number of observations in that bin.
   - **Use:** Useful for understanding the shape of the distribution, including skewness and modality (number of peaks).
2. **Bar Charts:**
   - **Purpose**: To compare the frequencies or counts of different categories.
   - **Description:** Bar charts use rectangular bars to represent the count or frequency of each category. The length or height of each bar corresponds to the value for that category.
   - **Use:** Ideal for categorical data or discrete data to compare different groups.
3. **Box Plots (Box-and-Whisker Plots):**
   - **Purpose**: To summarize the distribution of a dataset and identify outliers.
   - **Description**: A box plot shows the median, quartiles, and potential outliers. The central box represents the interquartile range (IQR), with a line inside indicating the median. "Whiskers" extend from the box to the minimum and maximum values within 1.5 times the IQR from the quartiles.
   - **Use:** Helpful for visualizing the spread and symmetry of data, and for comparing distributions across groups.
4. **Scatter Plots:**
   - **Purpose**: To explore relationships between two continuous variables.
   - **Description**: A scatter plot displays data points on a two-dimensional axis. Each point represents a pair of values for the two variables.
   - **Use:** Useful for identifying correlations, trends, and patterns between variables.

6. **Line Graphs**:
    - **Purpose**: To track changes over time or the relationship between two continuous variables.
    - **Description**: Line graphs plot data points on a Cartesian plane and connect them with lines. This helps to illustrate trends and changes over a period.
    - **Use**: Ideal for time series data or continuous data to show trends and patterns.
7. **Heat Maps**:
    - **Purpose**: To show the magnitude of values in a matrix format.
    - **Description**: A heat map uses color to represent data values in a matrix. Different colors represent different ranges of values.
    - **Use**: Useful for visualizing data intensity and spotting patterns in large datasets.

Graphical statistics complement descriptive statistics by providing a visual summary that can reveal insights not always apparent in numerical summaries alone. They help in exploring data, detecting outliers, and presenting findings in a more accessible format.

Descriptive Statistics

Descriptive statistics are used to summarize and describe the features of a dataset. They provide simple summaries and measures that give insight into the data's central tendency, dispersion, and overall shape.

help to provide a quick snapshot of the data and do not make any inferences beyond the data itself.

involves summarizing, organizing, and presenting data meaningfully and concisely.

Key Measures of Descriptive Statistics:

1. **Measures of Central Tendency:**
   - **Mean:** The average of all data points.
   - **Median:** The middle value when the data is sorted.
   - **Mode:** The most frequently occurring value.
2. **Measures of Dispersion:**
   - **Range:** The difference between the maximum and minimum values.
   - **Variance:** The average of the squared differences from the mean.
   - **Standard Deviation:** The square root of the variance, indicating how much the values deviate from the mean on average.
   - **Interquartile Range (IQR):** The range of the middle 50% of the data (the difference between the first quartile (Q1) and the third quartile (Q3)).
3. **Shape of the Data Distribution:**
   - **Skewness:** A measure of the asymmetry of the data distribution.(left,right,no)
   - **Kurtosis:** A measure of the "tailedness" of the distribution, indicating the presence of outliers.

Graphical Statistics

Graphical statistics refer to the use of visual representations to analyze and interpret data. These visual tools help in identifying patterns, trends, distributions, and relationships within a dataset. Graphical statistics are an essential part of exploratory data analysis (EDA).

Common Graphical Techniques:

1. **Histograms:** Display the distribution of a dataset by dividing the data into bins and showing the frequency of data points in each bin.
2. **Box Plots (Box-and-Whisker Plots):** Visualize the distribution of data based on five-number summary (minimum, Q1, median, Q3, and maximum) and highlight potential outliers.
3. **Bar Charts:** Used to represent categorical data with rectangular bars whose lengths are proportional to the frequencies or relative frequencies of the categories.
4. **Pie Charts:** Show proportions of a whole for different categories.
5. **Scatter Plots:** Visualize relationships between two continuous variables by plotting points corresponding to pairs of data values.
6. **Line Charts:** Useful for showing trends over time by connecting data points with lines.
7. **Heat Map**: Color represent the correlation between the attributes

Estimation Methods

Estimation methods are statistical techniques used to infer or estimate unknown population parameters based on sample data. The two main types of estimation are **point estimation** and **interval estimation**.

1. Point Estimation:

- A **point estimator** provides a **single value as an estimate** of a population **parameter**. Common point estimators include:
  - **Sample Mean ($\bar{x}$):** Used to estimate the population mean.
  - **Sample Proportion (p):** Used to estimate the population proportion.
  - **Sample Variance (s^2):** Used to estimate the population variance.

A good point estimator should be unbiased, consistent, and efficient. However, point estimates do not provide information about the variability or reliability of the estimate.

2. Interval Estimation:

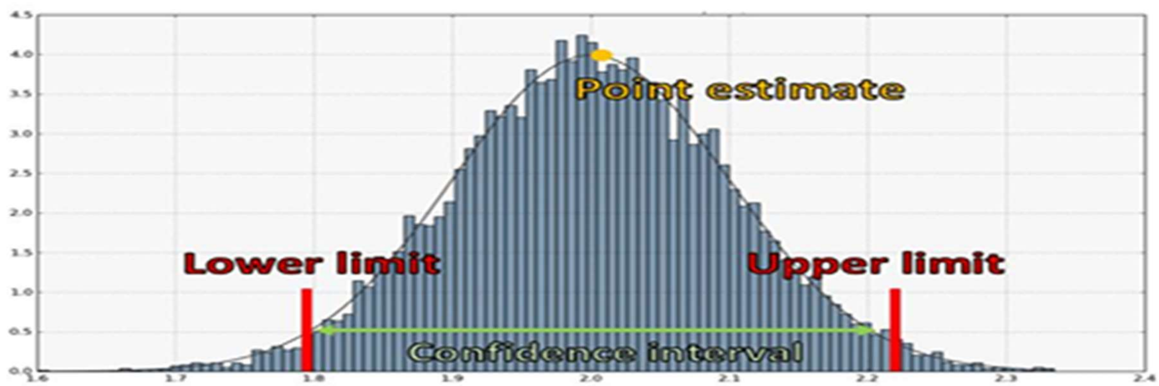
- An **interval estimator** provides a range of values within which the population parameter is expected to lie, with a certain level of confidence. The most common interval estimation method is the **confidence interval (CI)**.
- The **confidence level** (e.g., 95%) indicates the proportion of times the interval would capture the true population parameter if the study were repeated multiple times.

3. Methods of Estimation:

Several methods are commonly used for estimation in statistics, including:

- **Maximum Likelihood Estimation (MLE):** MLE finds the parameter values that maximize the likelihood of observing the given sample data. It is widely used in estimating parameters of probability distributions (explained previously).
- **Method of Moments (MoM):** The method of moments estimates parameters by equating sample moments (like sample mean, sample variance) with the corresponding population moments. For example, to estimate the mean and variance of a distribution, the method of moments solves equations based on the relationships between sample moments and theoretical moments.
- **Bayesian Estimation:** Bayesian estimation combines prior information about a parameter (in the form of a prior distribution) with the likelihood of the observed data to produce a posterior distribution. The posterior distribution represents the updated beliefs about the parameter after considering the data. The **Bayes estimator** is the value of the parameter that maximizes the posterior distribution.
- **Least Squares Estimation (LSE):** Least squares estimation is commonly used in regression analysis. It minimizes the sum of the squared differences between observed values and the values predicted by a model (e.g., a linear regression model).

- The primary goal of descriptive statistics is to provide a clear and concise summary of the data, enabling researchers or analysts to gain insights and understand patterns, trends, and distributions within the dataset.
- This summary typically includes measures such as central tendency (e.g., mean, median, mode), dispersion (e.g., range, variance, standard deviation), and shape of the distribution (e.g., skewness, kurtosis).



Maximum Likelihood Estimation (MLE):

1. MLE finds the parameter values that maximize the likelihood of observing the given sample data.
2. It is widely used in estimating parameters of probability distributions
3. The point in which the parameter value that maximizes the likelihood function is called the maximum likelihood estimate.
4. the parameter vector is considered which maximizes the likelihood function
5. The **goal** of maximum likelihood **estimation is to make inference about the population**

Steps for Maximum Likelihood Estimation:

1. **Specify the Likelihood Function:** Define the likelihood function for the given data and parameter θ, based on the assumed probability distribution.

2. **Compute the Log-Likelihood**: Take the natural logarithm of the likelihood function to obtain the log-likelihood.

3. **Differentiate the Log-Likelihood**: Take the derivative of the log-likelihood function with respect to the parameter θ.

4. **Find the Maximum**: Set the derivative equal to zero and solve for θ. This value θ^ is the maximum likelihood estimate of the parameter.

Examples:

Toss a Coin – To find the probabilities of head and tail

Throw a Dart – To find your PDF of distance to the bull eye

Maximum Likelihood Estimation (MLE) is a method used in statistics to estimate the parameters of a probability distribution by maximizing the likelihood function. The likelihood function measures how likely it is to observe the given data under different parameter values. The goal of MLE is to find the parameter values that maximize this likelihood, meaning they make the observed data most probable.

1)

model parameters. If θ represents the parameters and X represents the data, the likelihood function $L(\theta|X)$ is given by:

$$L(\theta|X) = P(X|\theta)$$

2)

$$\ell(\theta|X) = \log L(\theta|X)$$

4)

$$\hat{\theta} = \arg\max_{\theta} \ell(\theta|X)$$

The Method of Moments is a statistical technique used for parameter estimation. It involves matching the moments of a sample distribution to the moments of a theoretical distribution. Here's a step-by-step explanation:

1. **Define Moments:** Moments are quantitative measures related to the shape of a distribution. For a random variable X, the k-th moment is defined as $E[X^k]$, where E denotes the expected value. For example, the first moment is the mean, and the second central moment is the variance.
2. **Obtain Sample Moments:** Calculate the sample moments from your data. For a sample $x_1, x_2, \ldots, x_n$, the k-th sample moment is given by:

$$m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

   This estimates $E[X^k]$ using the sample data.
3. **Set Up Theoretical Moments:** Determine the theoretical moments of the distribution that you assume describes the data. For a distribution with parameters θ, the theoretical k-th moment might be a function of these parameters.
4. **Match Moments:** Equate the sample moments to the theoretical moments to form a system of equations. For example, if the theoretical k-th moment is $\phi_k(\theta)$, set:

$$m_k = \phi_k(\theta)$$

   for the relevant values of k.
5. **Solve for Parameters:** Solve the system of equations obtained in the previous step to estimate the parameters θ. These parameter estimates are the Method of Moments estimators.

Example

Suppose we have a sample and assume it follows a normal distribution with unknown mean μ and variance σ^2. The theoretical moments for this distribution are:

- First moment (mean): μ
- Second central moment (variance): σ^2

Using the sample data, calculate the sample mean and sample variance. Set these equal to the theoretical moments:

$$\hat{\mu} = \text{sample mean}$$

$$\hat{\sigma}^2 = \text{sample variance}$$

Here, $\hat{\mu}$ and $\hat{\sigma}^2$ are the Method of Moments estimates for the mean and variance of the normal distribution.

Fitting a distribution in statistics involves finding the best-fitting probability distribution for a given set of data. This process helps in understanding the underlying patterns and making predictions based on that data. Here's a step-by-step overview:

1. **Select Candidate Distributions:** Based on the nature of your data (e.g., continuous, discrete, skewed), you choose a set of potential probability distributions. Common distributions include normal, exponential, binomial, and Poisson.
2. **Estimate Parameters:** Each distribution has specific parameters (like mean and variance for the normal distribution). You estimate these parameters using statistical methods. For instance, the mean and variance of a normal distribution can be estimated using sample mean and sample variance.
3. **Fit the Distribution:** Use statistical techniques to fit the selected distributions to your data. This often involves using methods like Maximum Likelihood Estimation (MLE) or Method of Moments to find the parameter values that make the chosen distribution best represent your data.
4. **Evaluate the Fit**: Assess how well the distribution fits the data. This can be done using various goodness-of-fit tests, such as the Chi-Square test, Kolmogorov-Smirnov test, or graphical methods like Q-Q plots.
5. **Select the Best Distribution**: Compare the fit of different distributions using statistical criteria or goodness-of-fit measures. Choose the distribution that best represents the data, balancing complexity and goodness-of-fit.
6. **Validate the Model:** Ensure that the chosen distribution is appropriate for your data and the assumptions underlying the distribution are met.

Fitting a distribution helps in summarizing data, making predictions, and conducting further statistical analysis.

Random variable :-

Probability:- how probable something is to occure.
(A measure of how likely the event can happen is called Probability)

Random Variable

1) It can be any name that you are basically using
Suppose it as 'X' we can store something in that variable
2) A random variable X takes on defined set of values with diff probabilities.
3) value is unknown.
4) usually used in regression analysis. (to define stat relationship)

ex:- if you poll a die, the outcome is random (not fixed) & there are 6 possible outcomes, each of which occures with probability one-sixth (1/6)

Types of Random Variables :-

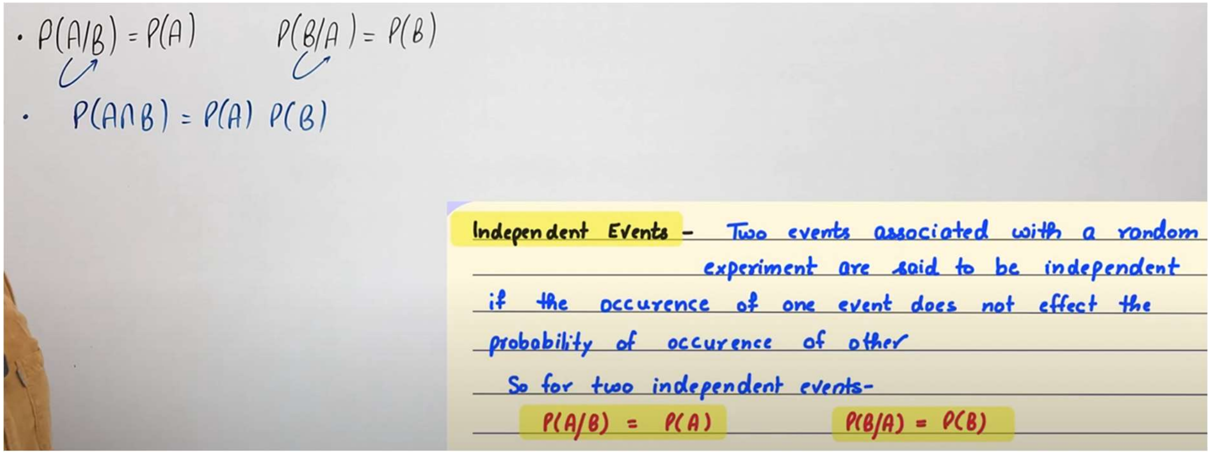
1) Discrete Random variables.
2) Continuous Random variables.

Aspect	Discrete Random Variable	Continuous Random Variable
Definition.	A discrete random variable can take on countable no. of distinct values. (	A continuous random variable can take on any ~~number~~ value within given range
Nature of values	finite or Countably infinite	infinite & uncountable within an interval
Types of Data	Usually integer	Real number or measurements.
Probability Distribution	Probability Mass function (PMF)	Probability density funn (PDF)
Graph Representation	Discreate points on graph	Continuous curve without break
ex :-	0,1,2,3,.... or no. of heads in coin tosses no. of students	height of students, 160.5cm, 172.8cm. weights, temperature
Summation vs integration	Probabilities are summed for individual outcomes	Probabilities are ~~integered~~ integrated over a range of values.

Discrete Random Variable	Continuous Variable
① They can take the specific or discrete values.	① They can take any value within a specific range.
② DRV are typically measured on a nominal or ordinal scale.	② CRV are typically measured on an interval or ration scale.
③ DRV are often represented by bar graph or histograms	③ CRV are often represented by line graphs or smooth curve.
④ Example:- Number of students in class, Outcomes of rolling die, etc	④ Example:- Measurments such a length, time or temp.
⑤ DRV have probability mass function (PMF)	⑤ CRV have probability density function (PDF)
⑥ Application They are employed in various mathematical contexts and applications where quantities are counted	⑥ They are often employed in various branches of mathematics, including calculus, differential equations, etc.
⑦ Probabilities are calculated for specific values (e.g. P(x = 5 students))	⑦ Probabilities are calculated for intervals (e.g. 160 cm < height < 170 cm)

Aspect	Discrete Distribution	Continuous Distribution
Variable Type	Takes on countable, distinct values (often integers)	Takes on an infinite number of values within a range
Probability of Specific Value	Non-zero for specific values (e.g., $P(X = k)$)	Zero for specific values; probability is defined over intervals
Common Distributions	Binomial, Poisson, Geometric, Negative Binomial	Normal, Exponential, Uniform, Gamma
Examples	Number of heads in coin flips, number of customers in a store	Time between events, heights of people

<u>**Independent event:**</u>

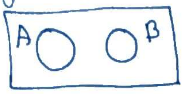


Independence :-

1) Happening of one event does not affect the happening of another event.

2) Probability of occurance of two events is independent

ex:- " B and A are independent if event A occuring confers no ~~ch~~ change to the probability of B, when ~~ω~~



Theorem 1 :-

Two events A & B are said to be independent if probability of their intersection is the product of individual probabilities.

$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

Theorem 2 :- if $A_1, A_2, \ldots A_n$ are independent events associated with a random experiment then $P(A_1 \cap A_2 \cap A_3 \ldots \cap A_n) = P(A_1)$ $\cdot P(A_2)\, P(A_3) \ldots P(A_n)$

Covariance :-

1) measures ~~of~~ the relationship between two random variables

2) Covariance is denoted by $Cov(x, y)$

3) Measured in units & calculated by multiplying the units of two variables.

Types of Covariance :-

Population $Cov(x, y)$

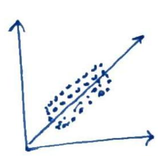
$$= \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{N}$$

Sample $Cov(x, y)$

$$= \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

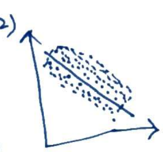
x_i = datavalue of x
y_i = —— 11 —— of y.
$\bar{x}$ = mean of x
$\bar{y}$ = —— 11 —— y
N = No. of data values.

1)



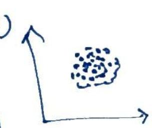
+ve co-relation
+ve covariance
$Cov(x, y) > 0$
moves in same dir^n
Both variables inc. together

2)



-ve corelation
-ve covariance
$Cov(x, y) < 0$
moves in opposite dir^n
~~Both~~ one incr. while other decr.

3)



zero co-relation
$Cov(x, y) \simeq 0$
No relationship btw 2 variables

CLT :- $\bar{x} \approx GD\left(\mu, \frac{\sigma^2}{n}\right)$

Bernouli: $p \rightarrow$ success $1-p \rightarrow$ failure mean $= p$, $var(x) = p(1-p)$

Binomial :- $p(x=k) = {}^nC_k \cdot p^k (1-p)^{n-k}$, mean $= np$, var $= np(1-p)$

Possion :- $p(x=k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!}$, λ , λ

Geometric :- $p(x=k) = (1-p)^{k-1} \cdot p$, mean $= \frac{1}{p}$, Var $= \frac{1-p}{p^2}$