**RANDOMNESS**

Randomness is a concept that refers to the lack of predictability or pattern in events, outcomes, or data. In a random process, the next outcome cannot be reliably determined based on previous outcomes or any other available information. It is often associated with a sense of unpredictability, chance, or uncertainty.

Example:

* When making a choice between two equally appealing options, a person might make a random or arbitrary decision
* The timing and location of lightning strikes.

**Statistical randomness**

The term “random” is often used colloquially to refer to things that are unexpected, but in statistics the term has a very specific meaning - A numeric sequence is said to be statistically random when it contains no recognizable patterns or regularities

Statistical randomness, also known as stochastic randomness, refers to the property of a sequence of data or a set of events where each element in the sequence or each event appears to be chosen or occurring at random, without any discernible pattern or predictability. In statistical terms, a sequence or process is considered statistically random when its elements exhibit the following characteristics:

* Independence: Each data point or event in a random sequence is independent of the others, meaning that the occurrence or value of one element does not influence the occurrence or value of subsequent elements.
* Uniformity of Distribution: In a statistically random sequence, each possible outcome or value is equally likely. This means that over a sufficiently long series of observations, each element occurs with roughly the same frequency.
* Unpredictability: In a statistically random sequence, it is difficult or impossible to predict the next element or event based on the previous ones. There is no underlying pattern or deterministic relationship between the elements

For example, if you flip a fair coin 10 times, the value of the outcome on one flip does not provide any information that lets me predict the outcome on the next flip. It’s important to note that the fact that something is unpredictable doesn’t necessarily mean that it is not deterministic. For example, when we flip a coin, the outcome of the flip is determined by the laws of physics; if we knew all of the conditions in enough detail, we should be able to predict the outcome of the flip. However, many factors combine to make the outcome of the coin flip unpredictable in practice.

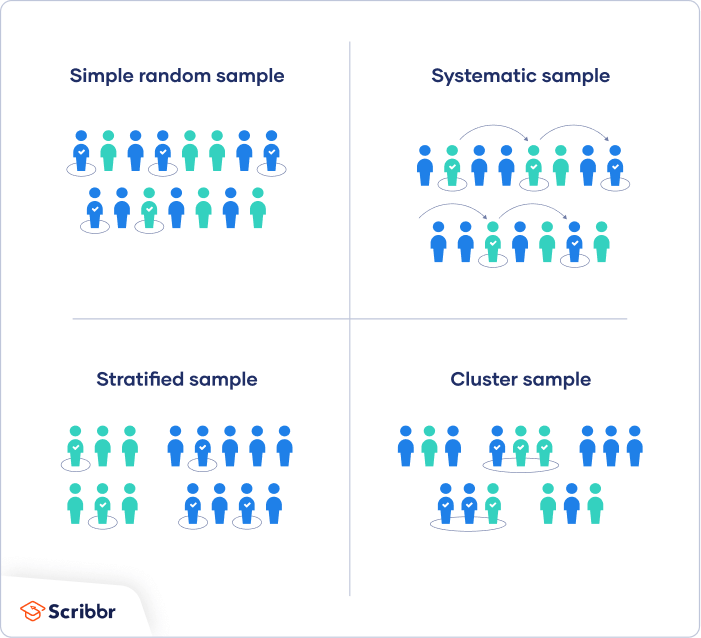
**The Role of Randomness in Statistics**

**1. Data Collection:**

Random Sampling: Randomness is essential in the process of data collection, particularly when selecting samples from a larger population.Random sampling is a part of the sampling technique in which each sample has an equal probability of being chosen. A sample chosen randomly is meant to be an unbiased representation of the total population. If for some reasons, the sample does not represent the population, the variation is called a sampling error.

There are four major types of this sampling method, they are;

1. Simple Random Sampling
2. Systematic Sampling
3. Stratified Sampling
4. Clustered Sampling



Example : In a classroom of 100 students, the teacher decided to use random sampling to select 10 students for a Group discussion, ensuring that every student had an equal chance of being chosen

**2. Reducing Bias:**

Bias: Bias refers to systematic errors in data collection or analysis that can lead to inaccurate or unfair conclusions. Randomness in sample selection helps reduce both selection bias and measurement bias. Selection bias occurs when certain subgroups are overrepresented or underrepresented in the sample due to non-random sampling. Measurement bias can arise from systematic errors in data collection instruments or methods. Random sampling reduces the risk of these biases, promoting fairness and accuracy in statistical analyses.

**3. Probability and Inference:**

Probabilistic Reasoning: Randomness is the cornerstone of probabilistic reasoning in statistics. It allows statisticians to quantify uncertainty, assess risk, and make informed decisions. Through the use of probability distributions, statistical inference becomes possible. This means that we can make educated guesses and predictions based on data.

Statistical Tests: Many statistical tests, such as hypothesis testing and regression analysis, rely on the assumption that data are collected randomly. These tests have been developed based on the properties of random samples. Without randomness, the results of these tests may lack validity and could lead to erroneous conclusions.

**4. Generalizability:**

Sample to Population: Randomness is crucial for the generalizability of findings from a sample to a larger population. When a sample is randomly selected and representative, the statistical conclusions drawn from the sample can be extended to the entire population with a quantifiable level of confidence. This is the basis for making population-wide predictions and decisions based on sample data.

**Pseudo-Random Number Generators (PRNGs)**

Pseudo-Random Number Generators are algorithms designed to produce sequences of numbers that mimic true randomness. These sequences are known as pseudo-random sequences and are widely used due to their computational efficiency and practical utility. However, it is essential to understand that PRNGs are not truly random; they are deterministic algorithms that produce sequences that pass many statistical tests for randomness.

The Importance of Randomness Testing

While PRNGs can generate sequences that appear to be random, it is crucial to subject these sequences to randomness testing to ensure their quality and suitability for various applications

Randomness test examples:

* Kolmogorov-Smirnov Test:
* Runs Test
* Chi-square test

[What is Random?](https://youtu.be/9rIy0xY99a0?si=WDQJ0nQLaCmk4TuJ)

<https://www.youtube.com/watch?v=sMb00lz-IfE&t=45s>