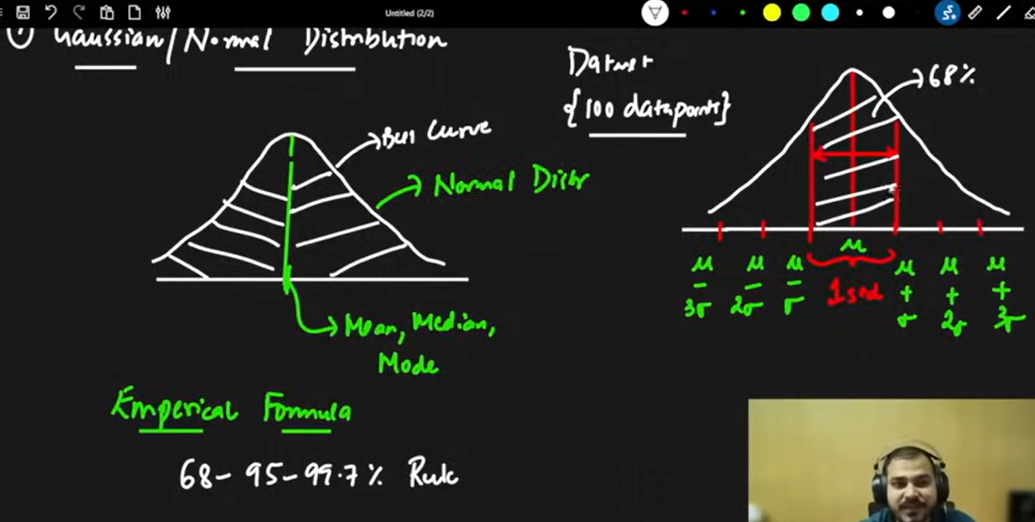
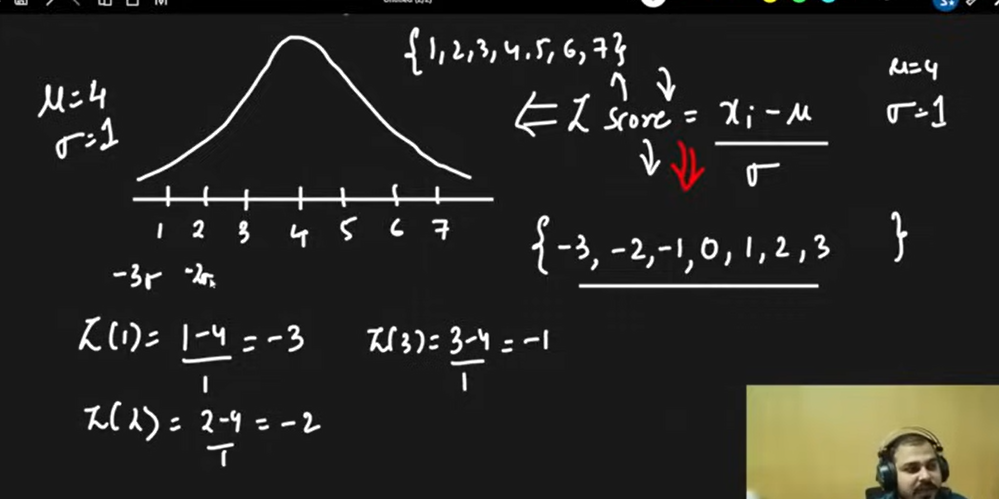
<https://www.freecodecamp.org/news/normal-distribution-explained/>

A diagram of normal distribution

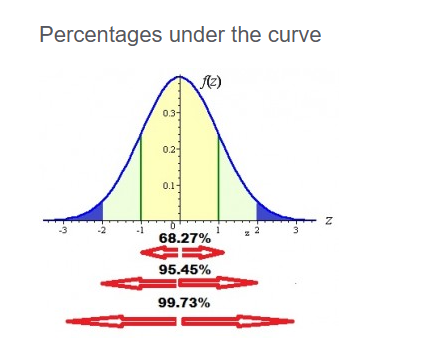
AI-generated content may be incorrect.

Link: [Complete Statistics For Data Science In 6 hours By Krish Naik](https://www.youtube.com/watch?v=LZzq1zSL1bs)

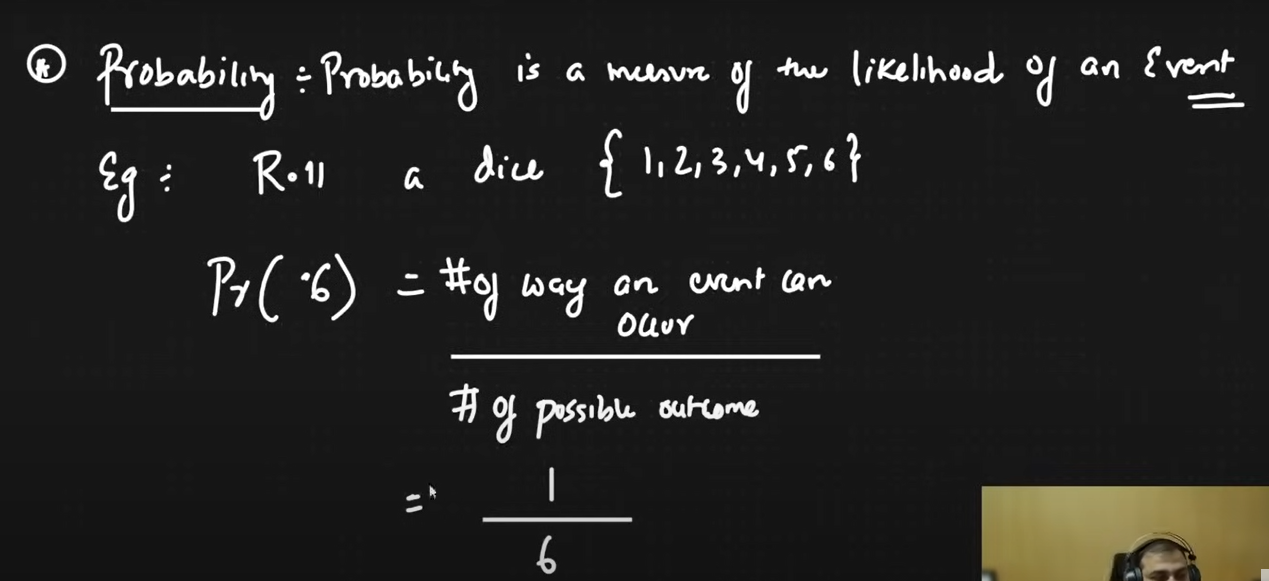


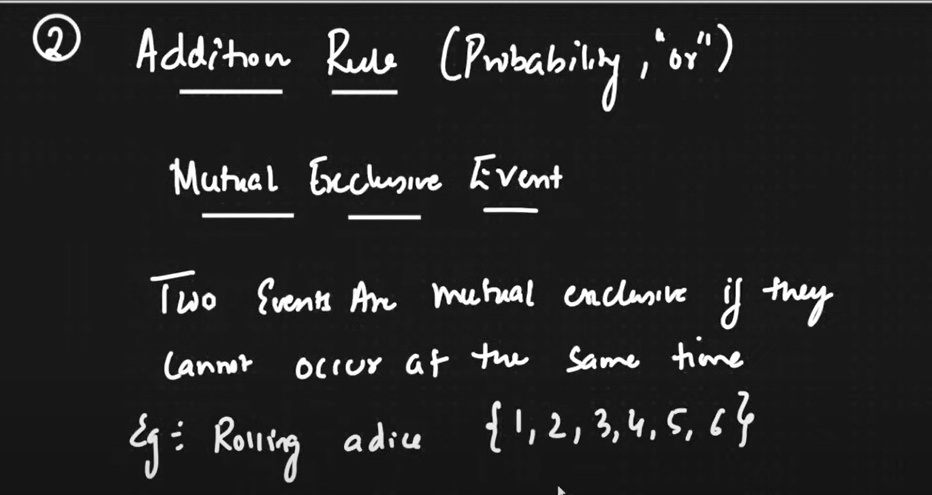


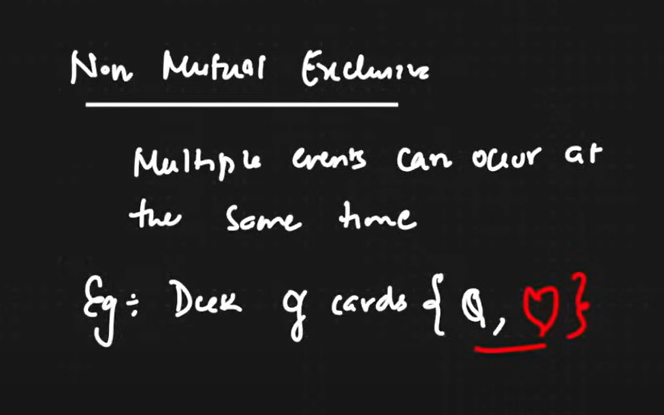
Here {-3,-2,-1,0,1,2,3} are SD derived via Z Score for {1,2,3,4,5,6,7}



<https://www.statisticshowto.com/tables/z-table/>







* **Mutually Exclusive Events:** Two events are mutually exclusive if they **cannot both happen at the same time**. Think of it as "either/or, but not both".
  + **Example:** Flipping a coin, the result can be either heads or tails, but not both at the same time.

**Non-Mutually Exclusive Events: The Focus Here**

* **Non-Mutually Exclusive Events:** Two events are non-mutually exclusive if they **can happen at the same time**. There's some overlap or shared outcomes between the events.
  + **Example:** Consider drawing a card from a standard deck of 52 playing cards.
    - Event A: Drawing a **heart**.
    - Event B: Drawing a **queen**.
    - These are non-mutually exclusive because you can draw a card that is both a heart **and** a queen (the queen of hearts).

The key to understanding non-mutually exclusive events is recognizing that **both events can occur simultaneously**. The addition rule, with the subtraction of the intersection (P(A and B)), adjusts for this overlap to accurately calculate the probability of either event happening.

The key is to understand that "non-mutually exclusive" doesn't mean both events *happen simultaneously in a single draw*. Instead, it means that *the outcomes of the events can overlap*.

Let's clarify with an example:

Imagine you have a bag with these items:

* Red ball
* Blue ball
* Red cube
* Green cube

Now, consider these events:

* Event A: Drawing a red item.
* Event B: Drawing a cube.

These events are non-mutually exclusive. Why? Because *it's possible* to draw an item that satisfies *both* conditions – the red cube. Even though you draw only *one* item at a time, the *possibility* of an outcome fulfilling both events makes them non-mutually exclusive.

**Relating back to the cards:**

When we say drawing a heart and drawing a queen are non-mutually exclusive, we're not saying you can draw the queen of hearts *and* the queen of spades at the same time in one draw. We're saying that *if you were to list all the possible outcomes of drawing a single card*, *some outcomes would belong to both the "hearts" group and the "queens" group*. That overlapping outcome is the queen of hearts.

**The Probability Perspective:**

The reason we care about whether events are mutually exclusive or not is for calculating probabilities.

* **Mutually Exclusive:** If events are mutually exclusive, the probability of either event happening is simply the sum of their individual probabilities.
* **Non-Mutually Exclusive:** If events are non-mutually exclusive, we have to account for the overlap (the queen of hearts, in our example) to avoid double-counting when calculating the probability of either event happening.

**In short:** **Non-mutually exclusive means the possibility of outcomes belonging to both events, not that both events occur at the exact same instant in a single trial*. It's about the overlap in the sample space of possible outcomes.***

Let's break down independent and dependent events with examples:

**1. Independent Events**

Two events are independent if the outcome of one event **does not affect** the outcome of the other event. In probability terms, knowing that one event has occurred gives you no information about whether the other event will occur.

**Examples:**

* **Example 1: Coin Flip and Dice Roll:**
  + Event A: Flipping a fair coin and getting heads.
  + Event B: Rolling a fair six-sided die and getting a 3.
  + These events are independent. Whether you get heads on the coin flip has absolutely no impact on what number you roll on the die.
* **Example 2: Drawing with Replacement:**
  + Imagine a bag with 5 red marbles and 3 blue marbles.
  + Event A: Drawing a red marble, replacing it, and then drawing another marble.
  + Event B: The second marble drawn is blue.
  + Because you replace the first marble, the probabilities for the second draw are the same as the first. The events are independent.
* **Example 3: Two Separate Coin Flips:**
  + Event A: Flipping a coin and getting heads.
  + Event B: Flipping the same coin again and getting tails.
  + The outcome of the first flip does not influence the outcome of the second flip. Each flip is a fresh start.
* **Example 4: Choosing two cards from deck with replacement**
  + Event A: Choosing a heart from the deck of cards, putting it back in the deck.
  + Event B: Choosing a spade from the deck of cards.
  + Since we are putting the card back in the deck, the probability of choosing a spade remains the same, irrespective of what we chose first.

**Probability of Independent Events:**

If events A and B are independent, the probability of both A and B occurring is:

P(A and B) = P(A) \* P(B)

**2. Dependent Events**

Two events are dependent if the outcome of one event **does affect** the outcome of the other event. Knowing that one event has occurred *does* change the probability of the other event.

**Examples:**

* **Example 1: Drawing without Replacement:**
  + Imagine the same bag with 5 red marbles and 3 blue marbles.
  + Event A: Drawing a red marble.
  + Event B: Drawing a blue marble without replacing the first marble.
  + These events are dependent. If you draw a red marble first, there are now only 7 marbles left in the bag, and the probability of drawing a blue marble on the second draw is different than if you had replaced the red marble.
* **Example 2: Weather and Outdoor Activity:**
  + Event A: It rains tomorrow.
  + Event B: You go on a picnic tomorrow.
  + These events are dependent (though not strictly deterministically). Rain makes it less likely that you'll go on a picnic. The occurrence of rain influences the probability of the picnic.
* **Example 3: Card Drawing without Replacement:**
  + Event A: Drawing a king from a standard deck of cards.
  + Event B: Drawing a second king from the deck without replacing the first card.
  + These are dependent because after you draw one king, there are only three kings left in the deck, changing the probability of drawing a king on the second draw.
* **Example 4: Picking two cards from deck without replacement**
  + Event A: Choosing a heart from the deck of cards, without putting it back in the deck.
  + Event B: Choosing a spade from the deck of cards.
  + Since we are not putting the card back in the deck, the probability of choosing a spade changes, since one card has already been picked.

**Probability of Dependent Events:**

If events A and B are dependent, the probability of both A and B occurring is:

P(A and B) = P(A) \* P(B|A)

Where P(B|A) (read as "the probability of B given A") is the probability of event B happening

*given that* event A has already occurred.

Imagine you have a group of friends, and you want to choose some of them for different activities.

* **Permutation:** Order matters! Think of it like arranging your friends in a line for a photo. If you choose Alice, then Bob, that's different from choosing Bob, then Alice. Permutations answer the question: "How many different ways can I arrange a certain number of items from a group where the order is important?"
* **Combination:** Order doesn't matter! Think of it like picking a team of friends to go to the movies. If you choose Alice and then Bob, it's the same team as choosing Bob and then Alice. Combinations answer the question: "How many different groups can I choose from a larger group where the order of selection doesn't matter?"

**In short:**

* **Permutation:** Arrangement matters (like a password or a race).
* **Combination:** Grouping matters (like a committee or a hand of cards).

Let's explore some practical use cases for permutations and combinations:

**Permutations (Order Matters)**

1. **Password Creation:** A password is a permutation because the order of characters is crucial. "password123" is different from "123password". The number of possible passwords of a certain length with a given character set is a permutation problem.
2. **Scheduling Tasks:** Suppose you have to complete five tasks, and the order in which you do them matters (e.g., because some tasks depend on others). The number of different schedules you can create is a permutation.
3. **Arranging Books on a Shelf:** The way you arrange books on a shelf is a permutation. A different order creates a different arrangement.
4. **Assigning Roles:** If you have 10 people and need to assign them distinct roles (president, vice president, treasurer), the number of ways to assign these roles is a permutation.
5. **License Plates/Serial Numbers:** The order of letters and numbers on license plates or serial numbers matters, making them permutations.

**Combinations (Order Doesn't Matter)**

1. **Lottery:** In most lotteries, the order in which you pick the numbers doesn't matter; what matters is the set of numbers you choose. The number of possible lottery tickets is a combination problem.
2. **Card Games (Hands):** In card games like poker, the order in which you receive your cards doesn't matter; it's the combination of cards in your hand that determines its value.
3. **Choosing a Committee/Team:** When selecting a committee or a team from a larger group, the order in which you select the members usually doesn't matter. The composition of the committee is what counts.
4. **Selecting a Group for a Project:** If you have a group of students and need to select a subset to work on a project, the order in which you choose the students is irrelevant (unless roles within the group are later assigned).
5. **Pizza Toppings:** When choosing pizza toppings, the order you tell the pizza maker doesn't change the final pizza. It's the combination of toppings that matters.

**Key Difference in a Nutshell**

The core difference is this: **If rearranging the items creates a different outcome, you're dealing with a permutation. If rearranging the items creates the same outcome, you're dealing with a combination.**

**Example: Consider the letters A, B, and C. If you want to choose 2 letters and the order matters, the possible permutations are:**

* **AB**
* **BA**
* **AC**
* **CA**
* **BC**
* **CB There are 6 permutations.**

**Example:** Using the same letters A, B, and C, if you want to choose 2 letters and the order *doesn't* matter, the possible combinations are:

* AB (which is the same as BA)
* AC (which is the same as CA)
* BC (which is the same as CB) There are 3 combinations.

**In essence:**

* **Permutation: Think "arrangement" or "sequence."**
* **Combination: Think "selection" or "grouping."**

Skewness can be defined as a statistical measure that describes the lack of symmetry or asymmetry in the probability distribution of a dataset. It quantifies the degree to which the data deviates from a perfectly symmetrical distribution, such as a [normal (bell-shaped) distribution](https://www.geeksforgeeks.org/normal-distribution-in-business-statistics/). Skewness is a valuable statistical term because it provides insight into the shape and nature of a dataset's distribution.**For example,** understanding whether a dataset is positively or negatively skewed can be important in various fields, including [finance](https://www.geeksforgeeks.org/finance/), [economics](https://www.geeksforgeeks.org/economics/), and data analysis, as it can impact the interpretation of data and the choice of statistical techniques.

**1. Positive Skewness (Right Skew)**

In a positively skewed distribution, the tail on the right side (the larger values) is longer than the tail on the left side (the smaller values). This means that the majority of data points are concentrated on the left side of the distribution, and there are some extreme values on the right side. In the case of a positively skewed dataset,

*Mean > Median > Mode*

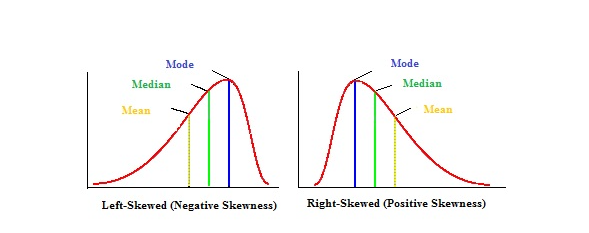
**Examples**of positively skewed data include income distribution (where most people earn a moderate income, but a few earn extremely high incomes), exam scores (where most students score in a certain range, but a few score exceptionally high), and stock market returns (where most days have modest returns, but a few days may have very high returns).

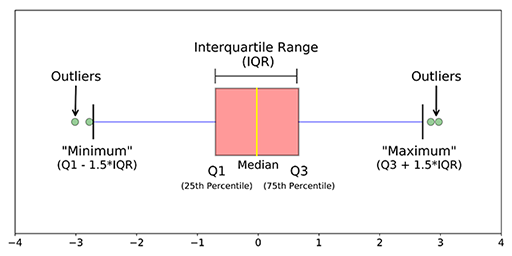
**2. Negative Skewness (Left Skew)**

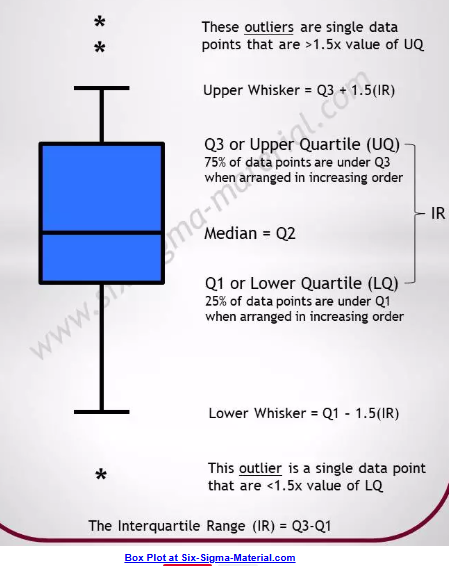
In a negatively skewed distribution, the tail on the left side (the smaller values) is longer than the tail on the right side (the larger values). This implies that most of the data points are concentrated on the right side of the distribution, with a few extreme values on the left side. In the case of a negatively skewed dataset,

*Mean < Median < Mode*

**Examples**of negatively skewed data include test scores on an easy test (where most students score well, but a few score very low), the age at retirement (where most people retire at a certain age, but a few retire exceptionally early), and the gestational age at birth (where most babies are born full-term, but a few are born prematurely).







|  |
| --- |
| Key Terms: |
| **Covariance:**   * Definition: Covariance measures the degree to which two variables change together. A positive covariance indicates that the variables tend to increase or decrease together, while a negative covariance indicates that they tend to move in opposite directions. A covariance close to zero suggests a weak linear relationship |
| **Correlation:**   * Definition: Correlation measures the strength and direction of the *linear* relationship between two variables. It's a standardized version of covariance, so it's easier to interpret and compare. |
| * Units: Correlation is a dimensionless quantity, meaning it has no units. It ranges from -1 to +1. * Interpretation:   + +1: Perfect positive linear relationship. As one variable increases, the other increases proportionally.   + -1: Perfect negative linear relationship. As one variable increases, the other decreases proportionally.   + 0: No linear relationship.   + Values close to +1 or -1 indicate a strong linear relationship.   + Values close to 0 indicate a weak or no linear relationship. |
|  |
|  |
|  |
|  |