Measure of Dispersion and Intro to Probability

Relevel
by Unacademy

Measures of Dispersion

In this section, we will cover Measures of Dispersion topics :

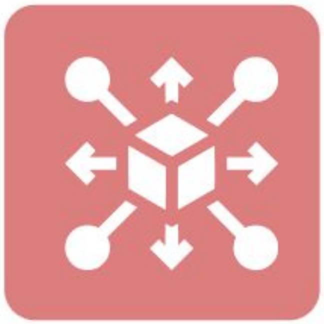
- Range
- Variance
- Standard Deviation

Measures of Dispersion

Being stretched out or in a state of dispersion. The degree to which numerical data is expected to deviate from an average value is known as Statistical Dispersion.
In other words, dispersion aids in the comprehension of data distribution.
Statistics measures of dispersion are used to assess data variability, or how homogeneous or heterogeneous the data is. Simply put, it demonstrates how constrained or dispersed the variable is.

A few measures of dispersion are:

- Range
- Variance
- Standard Deviation



Range

It is simply the difference between the maximum value and the minimum value given in a data set.

Example: In {4, 6, 9, 3, 7}, the lowest value is 3, and the highest is 9. Here, range is 6 (=9-3)

However, the range can sometimes be misleading when there are extremely high or low values.

Example: In {8, 11, 5, 9, 7, 6, 3616}:
the lowest value is 5,
and the highest is 3616,
So the range is 3616 − 5 = 3611.

Variance

Variance is the sum of the squares of the deviations from the mean. It is employed to gauge how dispersed a set of data is about the mean. The variance is calculated by comparing each result to the mean, unlike the preceding variability measurements. Whether you are calculating the variance for the complete population or using a sample to estimate the population variance, there are two formulas for the variance.

Population Diversity

$$\sigma^2 = 1/N\ \Sigma^N_{i=1} (X_i - \mu)^2$$

In the equation, σ2 is the population parameter for the variance,
μ is the parameter for the population mean, and
N is the number of data points, which should include the entire population.

Variance

$$\sigma^2 = 1/N\ \Sigma^N_{i=1} (X_i - \mu)^2$$

Let’s look at how this will be used in a practical case.
Raj wants to statistically understand the mileage he is getting from his new car - he has measured the mileage of his car and has the following readings -
13,15,14,13
Ans:
N = 4
μ = mean = (13+15+14+13)/4 = 13.75
Here, this formula will become,

$\sigma2 = ((13-13.75)2 + (15-13.75)2 + (14-13.75)2+ (13-13.75)2)/4$
$= ((-0.75)\ 2 + (1.25)\ 2 +(0.25)\ 2 + (-0.75)\ 2\ \ )/4$
$= (0.5625+1.5625+0.0625+0.5625)/4$
$= 2.75/4$
$= 0.6875$

Standard Deviation

The standard deviation is the average distance between each data point and the mean. Simply taking the variance's square root yields the standard deviation.

You have a reduced standard deviation when the values in a dataset are clustered more closely together. Conversely, when values are more dispersed, the standard deviation is higher because the standard deviation is higher.

The standard deviation conveniently uses the data's original units, simplifying interpretation. The standard deviation is, therefore, the most frequently employed measure of variability.

Standard Deviation (Contd.)

Take pizza delivery as an example; a standard deviation of 5 means that the delivery time will vary by 5 minutes from the mean. It's frequently stated alongside the average: 20 minutes (s.d. 5).

Simply taking the Variance's square root yields the standard deviation. Remember that the variance is expressed in square units. The square root thereby converts the value to natural units.

The standard deviation is denoted by the population parameter symbol and the letter s for a sample estimate.

After determining the variance, take its square root to determine the standard deviation. Eureka! You have the standard deviation!

In our earlier example with Raj and his car, the variance was 0.6875; our standard deviation is a Square root of 0.6875, which is 0.83.

This means his car has a mileage of 13.75, typically with a standard deviation of 0.83

Formula for Variance And Standard Deviation Ungrouped Data

The formula for the variance of an entire population is the following:

$$\sigma = \sqrt{\frac{\Sigma(X - \mu)^2}{n}}$$

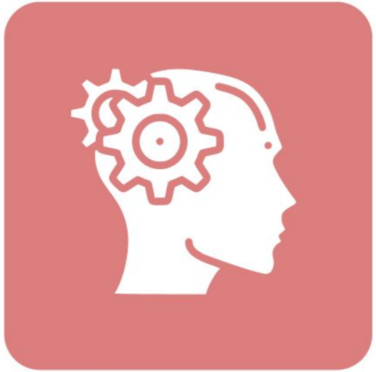
In the equation,

- σ2 is the population parameter for the variance
- σ is the population parameter for the standard deviation
- μ is the parameter for the population mean, and
- n is the number of data points, which should include the entire population

This is the same formula as before.

Variance and Standard Deviation – Practice Problem -1

Question-1: Calculate the variance and standard deviation of 4, 2, 5, 8, 6.



Variance and SD – Practice Problem:1 (Solution)

Step-1: Calculate the mean

$$\overline{x} = \frac{\Sigma x}{n} = \frac{x_1 + x_2 + x_3 + x_4 + \ldots + x_n}{n}$$

$$= (4+2+5+6+8) / 5 = 5$$

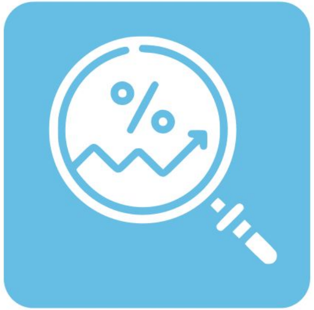
Step-2: Calculate the difference of each observation from mean and take square of that. Let's build this table on Excel

Observation (x)	x-mean	$(x\text{-mean})^2$
4	-1	1
2	-3	9
5	0	0
8	3	9
6	1	1

Variance and SD– Practice Problem:1 (Solution)

Step-3: Calculate Variance and Standard Deviation.

The variance for dataset is 4 and standard deviation is 2



Variance and Standard Deviation – Practice Problem -2

Question-1: Calculate the variance and standard deviation of **3, 8, 6, 10, 12, 9, 11, 10, 12, 7**.

Variance and SD– Practice Problem: 2 (Solution)

Step-1: Calculate the mean
Step-2: Calculate the difference of each observation from mean and take square of that.
Step-3: Calculate Variance and Standard Deviation.

The variance for dataset is 7.36 and standard deviation is 2.71

Formula for Variance And Standard Deviation Grouped Data

For Discrete frequency distribution of the type, where you have data points and their frequency
x: x1, x2, x3, ... xn and
f: f1, f2, f3, ... fn
The formula for standard deviation becomes:

$$(\sigma) = 1/N \sqrt{N\sum_{i=1}^{n} f_i x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2}$$

Here, N is given as:
$\mathbf{N} = {}^{n}\sum_{i=1} f_i$ (sum of number of readings (or sum of frequency))

Variance and Standard Deviation – Practice Problem -3

Question-1: Calculate the mean and standard deviation for the below data

x	f
60	2
61	1
62	12
63	29
64	25
65	12
66	10
67	4
68	5

Variance and Standard Deviation – Practice Problem: 3 (Solution)

x	f	x*f	f*x^2
60	2	120	7200
61	1	61	3721
62	12	744	46128
63	29	1827	115101
64	25	1600	102400
65	12	780	50700
66	10	660	43560
67	4	268	17956
68	5	340	23120
	100	6400	409886

Mean	64
Variance	2.86
Standard Deviation	1.691153453

Variance is 2.86 and standard deviation is 1.69

Variance and Standard Deviation – Practice Problem -4

Question-1: Calculate the mean and standard deviation for the below data

Class	f
0-10	27
10-20	10
20-30	7
30-40	5
40-50	4
50-60	2

Variance and Standard Deviation – Practice Problem: 4(Solution)

Class	f	Midpoint of ith class interval (Mi)	fiM	f*Mi^2
0-10	27	5	135	675
10-20	10	15	150	2250
20-30	7	25	175	4375
30-40	5	35	175	6125
40-50	4	45	180	8100
50-60	2	55	110	6050
	55	180	925	27575
			855625	

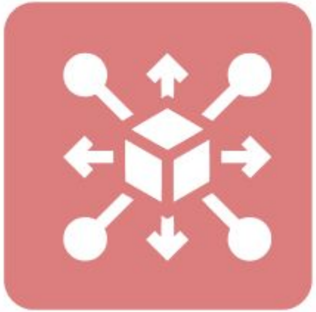
Mean	16.81818182
Variance	218.5123967
Standard Deviation	14.78216482

Variance is 218.5 and standard deviation is 14.78

Intro to Probability

Provide an overview on:

- Introduction to probability
- Key Terminologies
- Practice Problems



Intro to Probability

Probability is an intuitive concept. We use it daily without realising that we are speaking and applying probability to work.
There are numerous unknowns in life. Until something happens, we cannot predict how a situation will turn out. Today, will it rain? Will I succeed on my next math exam? Which team will win the coin toss? Will I be promoted in the upcoming six months? These queries are all illustrations of the unpredictability of our world. Let's translate them into a few terms we will use frequently moving forward.

Why do we need probability?

Knowing and comprehending the probabilities of different outcomes can be quite helpful in a world full of uncertainty. You can arrange your plans accordingly. I would bring my umbrella if it looked like it could rain. I will have myself tested if my eating habits indicate that I might have diabetes. I will notify my customer of a renewal premium if he is unlikely to do so without a reminder.
So knowing the likelihood might be very beneficial.

Definition of probability

Probability is the measurement of an event's likelihood. Probability aids in determining the chance of an event occurring because many events cannot be predicted with 100% accuracy. It is the proportion of positive events to all of the events in an experiment.

Probability(Event)=Favorable Outcomes/Total

Numerically the probability value always lies between 0 and 1.

0≤P(E)≤1

It is expressed in percentage, decimal, or fraction.

Key Terminologies

- **An experiment** is a task for which the results are unknown. Every experiment has a mix of successful and unsuccessful results. Before Thomas Alva Edison successfully attempted to create the light bulb, he made more than a thousand unsuccessful attempts during his historical experiments.

- **Random Experiment:** A random experiment is one for which the range of potential outcomes is known, but it is impossible to predict which specific consequence will occur on a given execution of the experiment in advance. Random experiments include tossing a coin, rolling a dice, and picking a card randomly from a deck.

- **Trial:** The term "trials" refers to all attempts made throughout an experiment. In other terms, a trial is any specific outcome of a random experiment. Tossing a coin, for instance, is a trial.

- **Event:** A trial with a certain conclusion qualifies as an event. An event might be something like throwing a coin and getting a tail.

Key Terminologies

Random Event: A random event is something unforeseen. You can never assign an exact value or probability because it is unexpected. For instance, because falling down a flight of stairs is purely random, it is impossible to predict the likelihood that you will do so in the next ten years.

Outcome: This is the outcome of the trial. Two obvious consequences exist when a sportsperson kicks a ball towards the goal post. He has the potential to score or fail to do so.

Possible Outcomes: A conceivable outcome is just a list of every possible result of an experiment. There are two possible results when tossing a coin: heads or tails.

Equally likely outcomes: Equally likely outcomes result from an experiment when each possible result has the same chance of happening. Any number can be rolled on a six-sided die with an equal chance of coming up.

P(any number)=1/6

Key Terminologies

- **Sample Space:** This collects results from every experiment's trials. There are six possible results when tossing a die: 1, 2, 3, 4, 5, and 6. The sample space is comprised of these results. S = {1, 2, 3, 4, 5, 6}
- **Probable Event:** A probable event can be foreseen. We can determine the likelihood of such occurrences. One can assess the likelihood that a certain child will advance to the following grade. As a result, we can describe this as a likely event.
- **Impossible Event:** An event that does not occur during the experiment or does not fit within the results' sample space is an impossible event. In a region with a temperate climate, there is no snowfall. Since there is zero chance that it would snow in this situation, it can be said that the event is impossible.

Key Terminologies

- **Complementary Events:** When there are only two possible outcomes, and one of them is completely the opposite of the other, complementary events take place. The complement of an occurrence with probability P(A) is P(A'). P(A)+P(A`)=1 The success and failure events in an examination are mutually supportive. P(Success)+P(Failure)=1 In a coin toss, getting heads and tails are complementary outcomes.

- **Mutually Exclusive Events:** Two events are said to be mutually exclusive if one event prohibits the occurrence of the other. In other words, if two occurrences cannot happen simultaneously, they are said to be mutually exclusive. For instance, flipping a coin has a chance of producing either heads or tails. Both cannot be seen simultaneously.

Probability Example – 1

A football team plays 120 matches and wins 80 matches. What is the probability of the team winning the next match?

Probability Example – 1 (Solution)

Total number of matches played = 120
Number of matches won by the team = 80
Probability of winning = Probability(Winning) = Favorable Outcomes/Total
= Number of matches won/Total number of matches
=80/120
=2/3
= 66.67%
This means that when the team plays a new match, we can say there is a 66.67% chance that they will win

Probability Example - 2

Tom's teacher wrote each English alphabet on a different piece of paper and jumbled all those in a box. He asked Tom to randomly pick a piece of paper. What is the probability of having a vowel written on that piece of paper?

Probability Example - 2

Total number of English alphabet = 26

Number of vowels = 5

Probability of getting a vowel = Probability(getting a vowel)=Favorable Outcomes/Total

= Number of vowels/Total number of alphabets

=5/26

Probability Example - 3

Sam takes two coins and flips them both at once. What is the probability of getting heads on both the coins?

Probability Example – 3 (Solution)

Sample space on flipping two coins = {(H, H), (H, T), (T, H), (T, T)}
Total number of outcomes = 4
Favorable outcome of two heads = 1
Probability of getting two heads =Number of outcomes with two heads/Total number of outcomes
=1/4

Probability Example - 4

Mary had a jar containing 8 red balls, 5 blue balls, and 7 green balls. She called one of her friends and asked them to pick a ball from the jar. What is the probability that the ball which is picked is either a red or a blue ball?

Probability Example – 4 (Solution)

Let us take a count of the number of balls in the jar.

Number of red balls = 8

Number of blue balls = 5

Number of green balls = 7

Total number of balls = 20

P(red or a blue ball) = (No of red balls + No of blue balls)/Total number of balls

=(8+5)/20

=13/20

Permutations and Combinations

Provide an overview on:

- Principle of counting
- Permutation
- Combination
- Probability on PnC

Principle of Counting– Example:1

Suppose there are 14 boys and 9 girls. If a boy or a girl has to be selected to be the monitor of the class, in how many ways can the teacher select a monitor?

Solution:

The teacher can select 1 out of 14 boys or 1 out of 9 girls. She can do it in 14 + 9 = 23 ways (using the sum rule of counting).

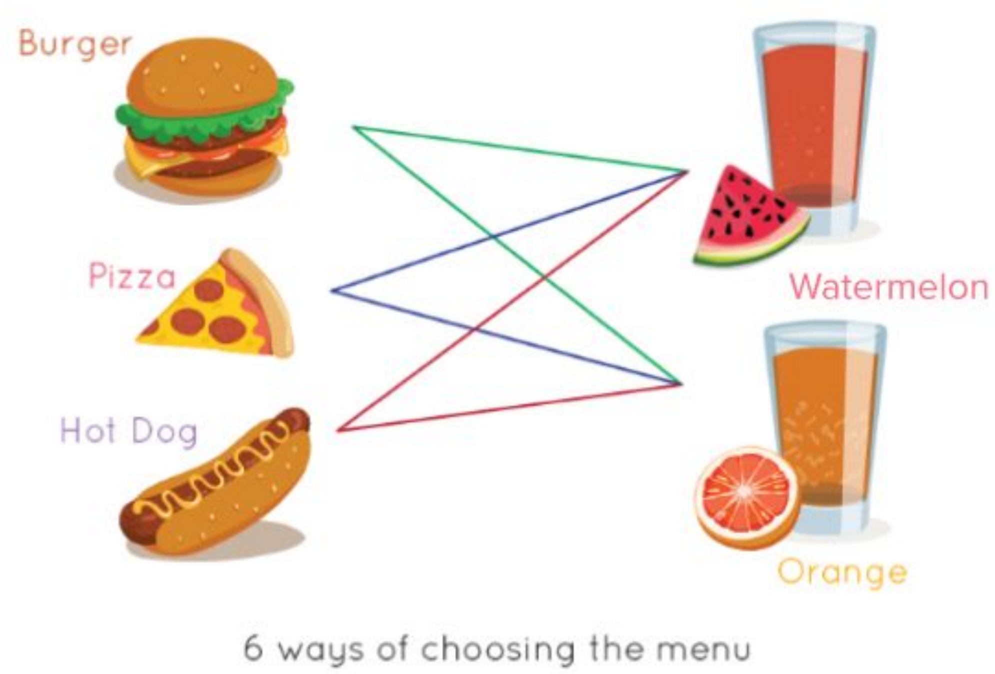
Principle of Counting – Example: 2

Let us look at another example. Let's say Sam typically orders a drink and one main entrée. He has the option of a watermelon juice, orange juice, hot dog, pizza, or burger today. What are the various combinations he may try? There are three food options as well as two drink options. To discover the combinations, we multiply.

Solution:

$3 \times 2 = 6$. Thus Sam can try six combinations using the product rule of counting. This can be shown using the figure.

Principle of Counting – Example: 2



Permutation And Combinations

Permutation and combination form the fundamentals of counting and are used in various analytics situations.

A permutation is a list of the various configurations that can be created from the given set of items.
Detailing is vital in permutation because the order or sequence matters. Writing the names of three countries {India, China, Israel} or {China, India, Israel) or { Israel, China, India } is different or unique to each other (as the order in which they are mentioned as different), and this sequence in which the names of the countries are written is important.

Let's explore this example further - There are the following ways to write these country names in groups of 3 -
{ Israel, China, India }, { Israel, India, China }, { India, China, Israel }, { India, Israel, China }, { China, India, Israel }, { China, Israel, India}

A combination here would be just the group of countries regardless of order. Hence we have 6 permutations and 1 combination

The sequence or order of the names is irrelevant when three countries are combined because their names from only one group. **Combinations and permutations** are referred to as **selections and arrangements**, respectively. The sum and product rules, based on the basic concept of counting, make counting simple to use.

Permutation

Permutation refers to arranging a number of objects in different orders taken some or all at one time.

$$^{n}P_{r} = n!/(n-r)!$$

n is the number of objects from where selection is done

r is the number of objects required in each selection

Does everybody remember factorials?

Permutation – Example: 1

Let us take 10 numbers: 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. What is the number of 4-digit-passcode which can be formed using these 10 numbers?

Permutation – Example: 1 (Solution)

These are the easiest to calculate.
When a thing has n different types … we have n choices each time!
Here we have 10 different types. Hence, we have 10 choices

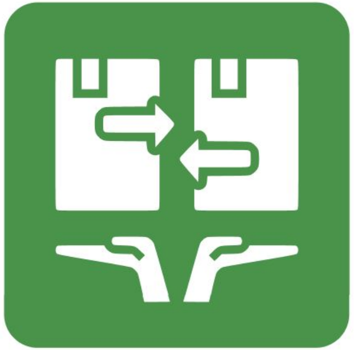
Choosing ***r*** of something that has ***n*** different types, the permutations are:
n × n × … (r times)

Here, we are choosing 4 numbers, so we have 10 x 10 x 10 x10 = 10^4 times

Here, the reason we have not used the formula is because the numbers can be repeated - 0000, 1111, 1122, are all possible

Permutation – Example: 2

Let us take 10 numbers: 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. What is the number of different 4-digit-passcode which can be formed using these 10 numbers without repeating the numbers?



Permutation – Example: 2 (Solution)

This can be calculated using permutations -
P(10,4) = 5040.
Hence, you can form 5040 4 digit passcodes using the numbers 0-9.
The passcodes formed will all be unique i.e they cannot be repeated.

Combinations

A combination gives entire grouping information. Combinations can be used to determine how many different groups can be created from the available items.

$$^{n}C_{r} = n!/r!(n-r)!$$

Combination – Example: 1

A team of 2 is formed from 5 students (Raj, Jagdish, Naveen, Lokesh, and Ojas). Find the possible combination of teams.

Combination – Example: 1 (Solution)

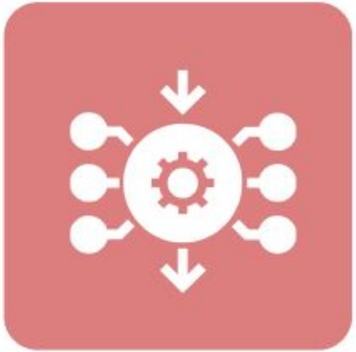
The combinations can happen in the following 10 ways by which the team of 2 could be formed.

- Raj Jagdish
- Raj Naveen
- Raj Lokesh
- Raj Ojas
- Jagdish Naveen
- Jagdish Lokesh
- Jagdish Ojas
- Lokesh Naveen
- Lokesh Ojas
- Ojas Naveen

This is a simple example of combinations. C(5,2) = 10.

Probability Question on PnC: Question-1

Example 1: Praveen has to choose 5 marbles from 12 marbles. In how many ways can she choose them?



Probability Question on PnC: Question- 1 (Solution)

Praveen has to choose 5 out of 12 marbles. The order doesn't matter here. Thus combinations used here are: she can choose it in 12C5 ways. Therefore there are 792 ways.

$$
\begin{aligned}
{}^{12}C_5 &= 12!/5!(12-5)! \\
&= 12!/5!7! \\
&= (12 \times 11 \times 10 \times 9 \times 8 \times 7!)/(5! \times 7!) \\
&= 12 \times 11 \times 10 \times 9 \times 8 \,/\, 5! \\
&= 792
\end{aligned}
$$

Probability Question on PnC: Question-2

A committee of 3 members will be formed with two males and one female. Find the number of ways this committee can be formed from 5 male and four female members.

Probability Question on PnC: Question- 2 (Solution)

The aim is to form a committee of 3 members, with 2 male members and 1 female member.

Number of male members = 5

Number of female members = 4

We can form this committee by taking 2 male members from 5 male members, and 1 female member from 4 female members.

We apply the combinations formula, to find the solution.

The number of ways of forming this committee =

$^5C_2 \times {}^4C_1$

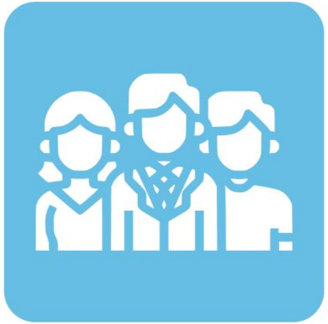
- As you can see, when we split a case into multiple permutations or combinations, the final valu [illegible] them

$=5!/[2!(3)!] \times 4!/[1!(3)!]$

$= [120/12] \times [24/6]$

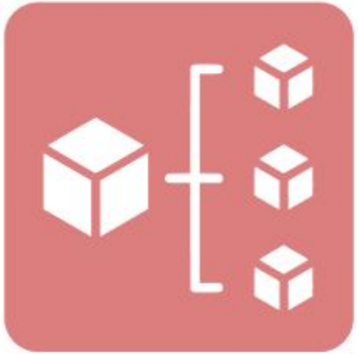
$= 10 \times 4 = 40$

Answer: Therefore, the committee can be formed in 40 ways.



Probability Question on PnC: Question- 3

There are 10 marbles in a rucksack, numbered from 0 to 9. How many ways of 3 different digits could be formed by picking them up from the rucksack, without replacement?



Probability Question on PnC: Question- 3 (Solution)

The number of permutations of 3 digits chosen from 10 marbles is 10P3

Using Permutations formula, we know:

$nPr=n!(n-r)!$

$10P3= 10\,!/\,7!$

$10\times 9 \times 8 =720$

Answer: Thus in 720 ways, 3 digits can be formed from 10 marbles

Conclusion

THANK YOU