Continuous probability Distribution

v. Imp ⊙ Normal Distribution
⊙ Exponential "

Imp Central limit theorem

The central limit theorem states that the sampling distribution of the sample means approaches normal distribution as the sample size gets large no matter what the shape of the population distribution. This fact holds especially true for sample sizes over 30. All this is saying is that as you take more samples, especially large ones, your graph of the sample means will look more like a normal distribution.

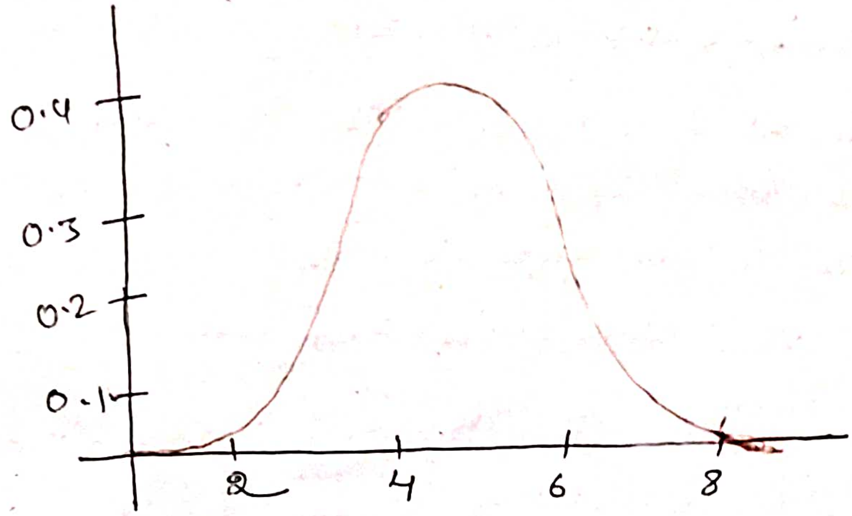
The Central Limit Theorem and Means

~~As~~ An essential component of the Central Limit theorem is that the average of your sample means will be population mean. In other words, add up the means from all of your samples, find the average and that average will be your actual population mean. Similarly, if you find the average of all of the standard deviations in your sample, you'll find the actual standard deviation for your population. It's ~~as~~ a pretty useful phenomenon that can help accurately predict characteristics of a population.

V.V. Imp

Normal Distribution

One of the most well-known distribution in the class of probability distributions is the normal distribution or Gaussian distribution.



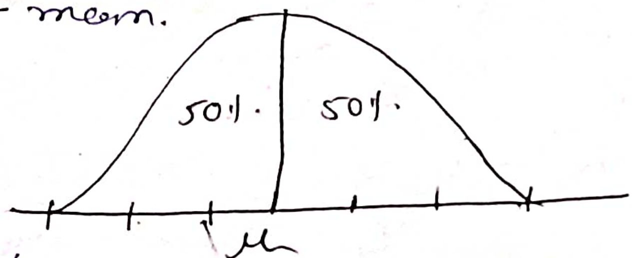
Why is the normal distribution so important?

- Many biological characteristics conform to a Normal distribution – for example,
- Heights of adult men and women.
- Blood pressure in a healthy population.
- Random errors in many types of laboratory measurements and biochemical data.
- Many physical quantities follow the normal distribution at least approximately.

Characteristics of the Normal Distribution

- The symmetric bell shaped curve

It is symmetric about mean.

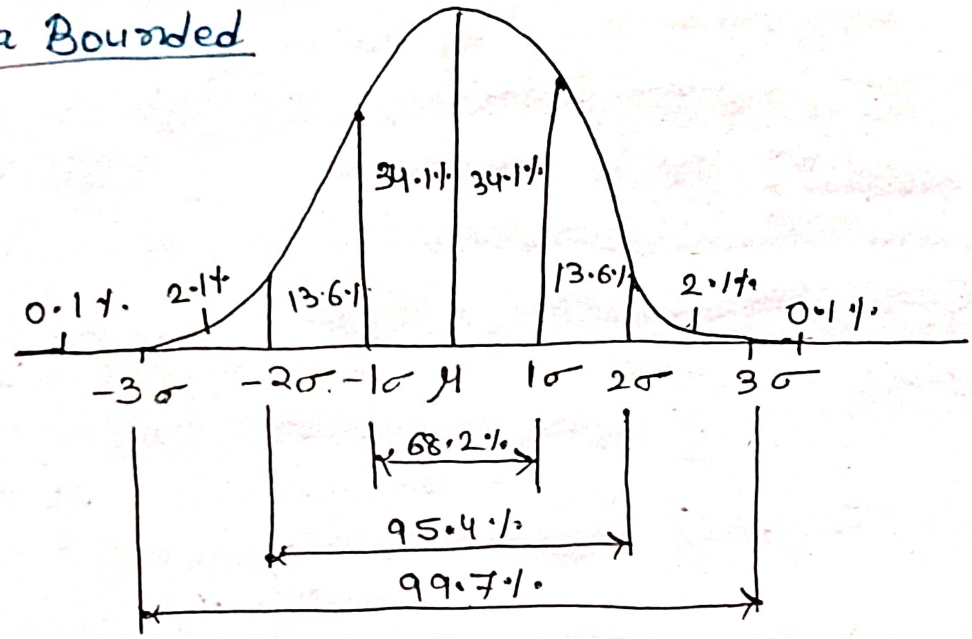


- The mean, median and the mode are equal.
- Total area under the curve, above the x axis is one, because of the symmetry, 50% to the area is to the right of a ⊥ erected at the

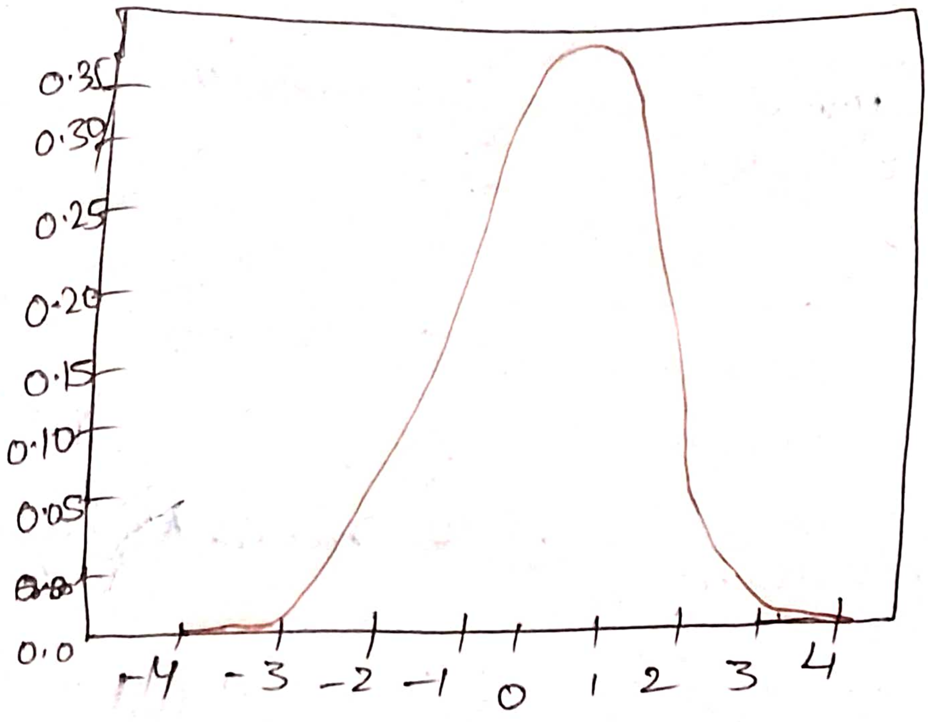
mean, and 50% is to the left.

⑤ Characteristics ---
Empirical rule - In normally distributed data, there is a constant proportion of ~~distan~~ distance lying under the curve between the mean and specific number of standard deviations from the mean. For example, 68.5% of all cases fall within +- one standard deviation from the mean. 95% of all cases fall within +/- one standard deviation from the mean. 95% of all cases fall within +/- two standard deviations from the mean, while 99% of all cases fall within +/- three standard deviations from the mean.

Area Bounded




```
→ from numpy import random
import matplotlib.pyplot as plt
import seaborn as sns
x = random.normal(loc=1, scale=2, size=(2,3))
print(x)
sns.displot(random.normal(size=1000), hist=False)
plt.show()
```



0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.0
-4
-3
-2
-1
0
1
2
3
4