భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Biostatistics BT2023

Lecture 9

Himanshu Joshi
2 September 2022

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Start-up notes

History fact

Abraham de Moivre (1667-1754) in 18th century first derived the Normal distribution function which was later independently derived by Gauss and Laplace ~ 200 years later

Goal of biostatistics

To learn about a population of biological population using information in a sample.

Fun fact

The early development of statistics is intimately related to gambling.

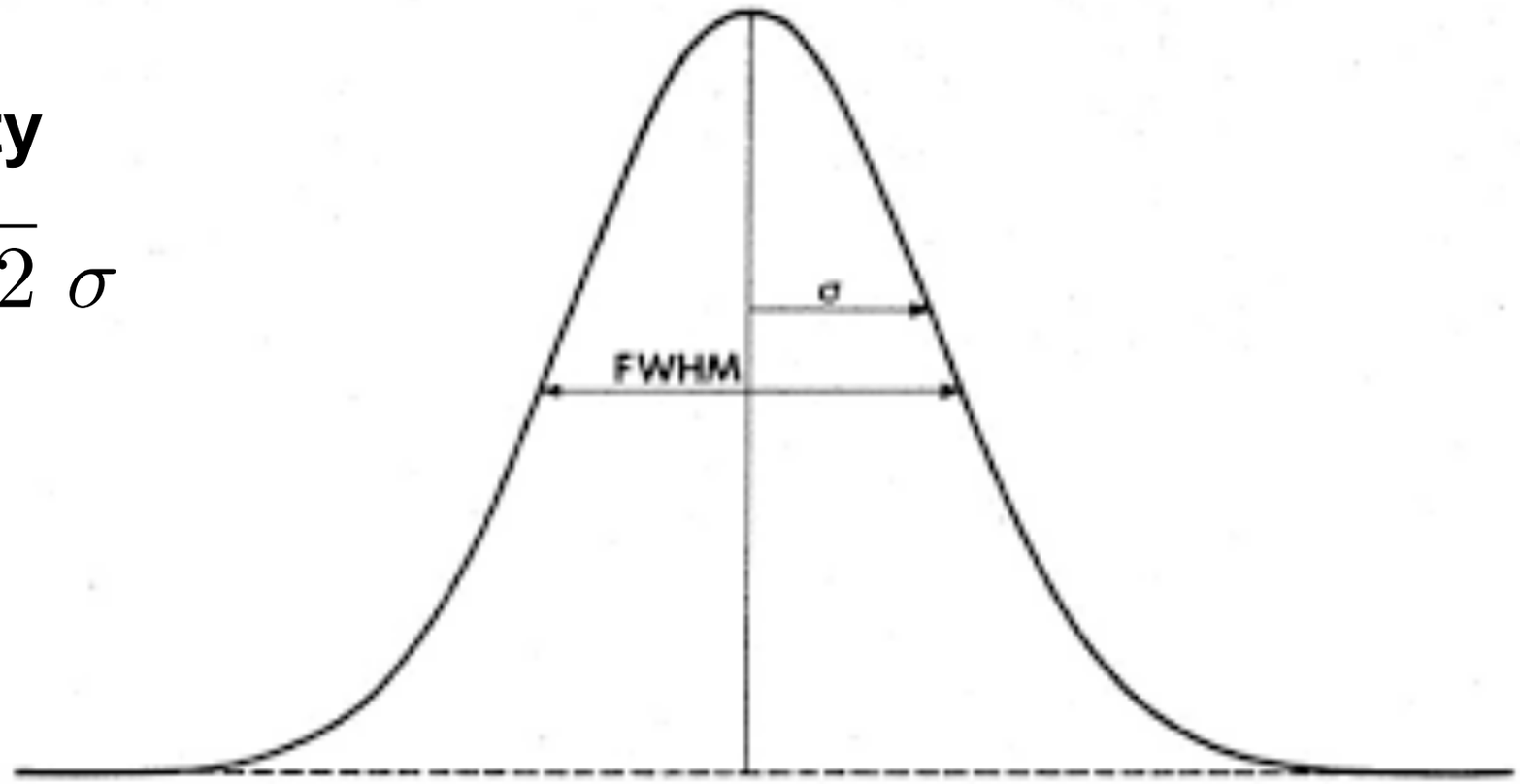
Now when are sitting in the class 0.7 population of the world population is drunk

Normal or symmetrical distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp - \left[\frac{(x - x_0)^2}{2\sigma^2}\right]$$

Limits of Variability

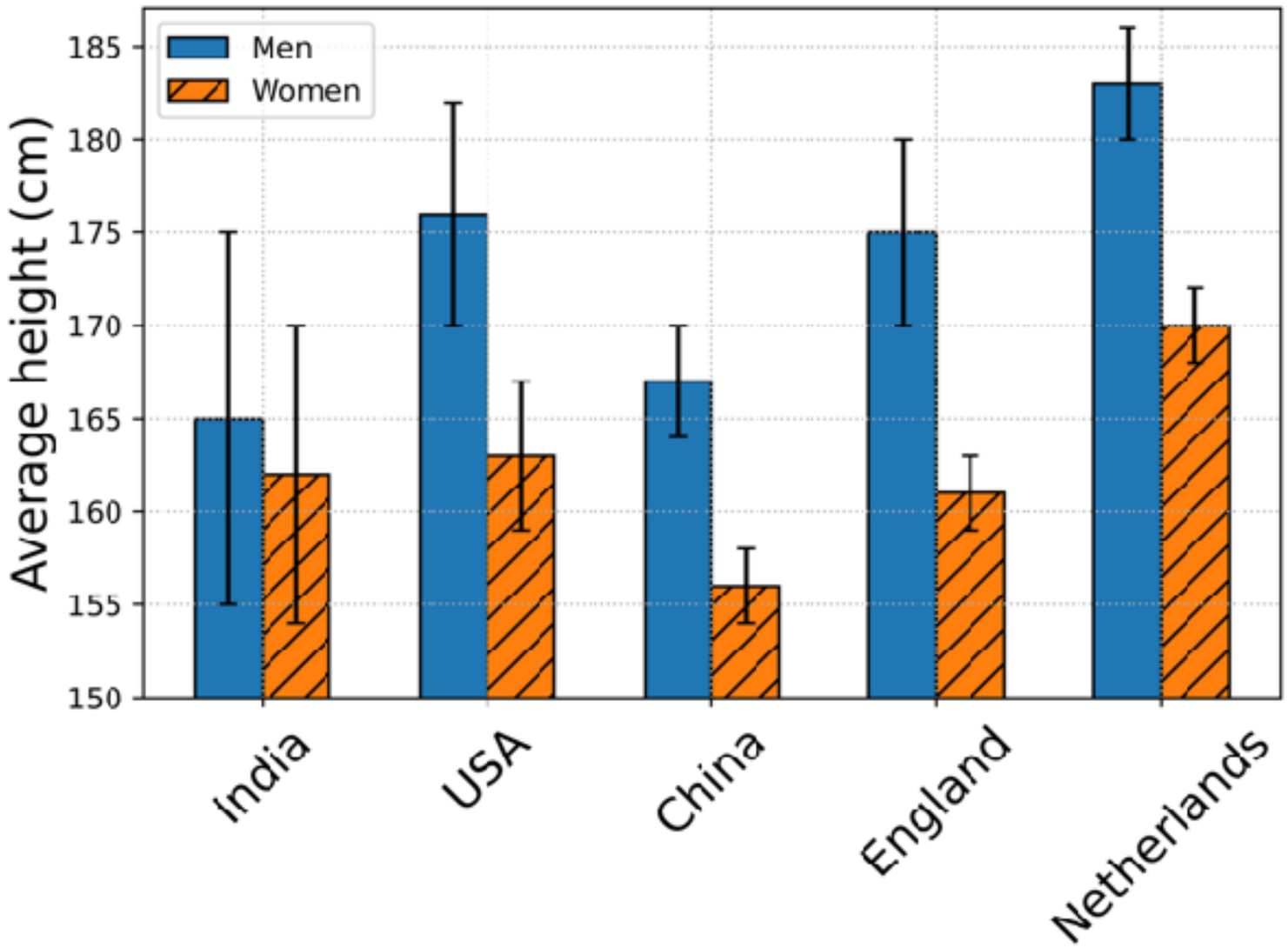
$$FWHM = 2\sqrt{2ln2}\ \sigma$$



Plotting this function

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Plot of the day



Question

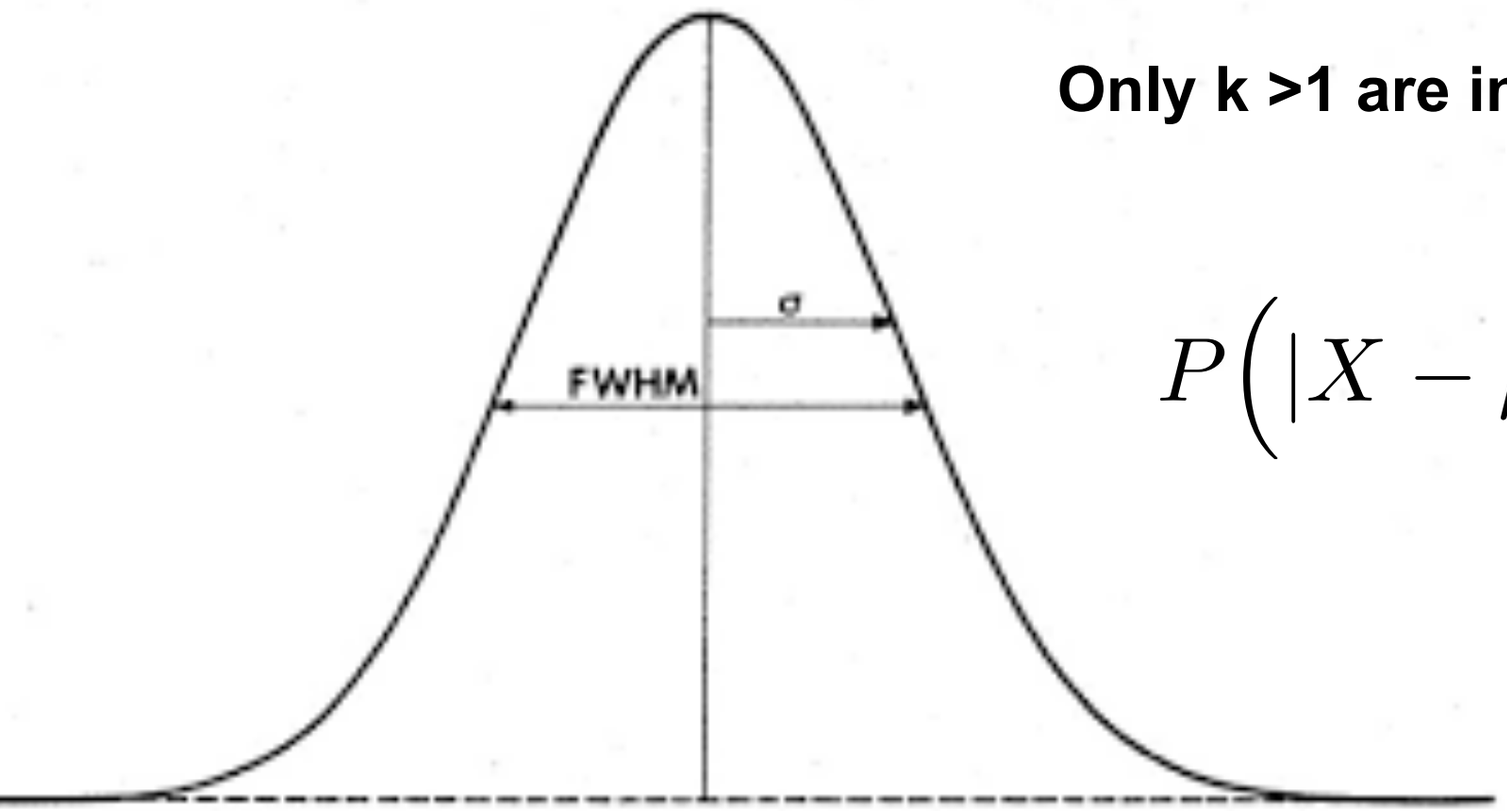
What is difference b/w a bar plot and histogram

Measure of dispersion

Chebyshev's inequality

The rule is often known as Chebyshev's theorem, tells about the range of standard deviations around the mean, in statistics.In a probability distribution, no more than a certain fraction of values can be more than a certain distance from the mean.

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



Only k >1 are interesting because for k<1 it trivial

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

Mean, Mode, Median

Line plot,

Histogram,

Bar płot,

Pie chart,

Scatter plot,

double axis plot,

Log log plot

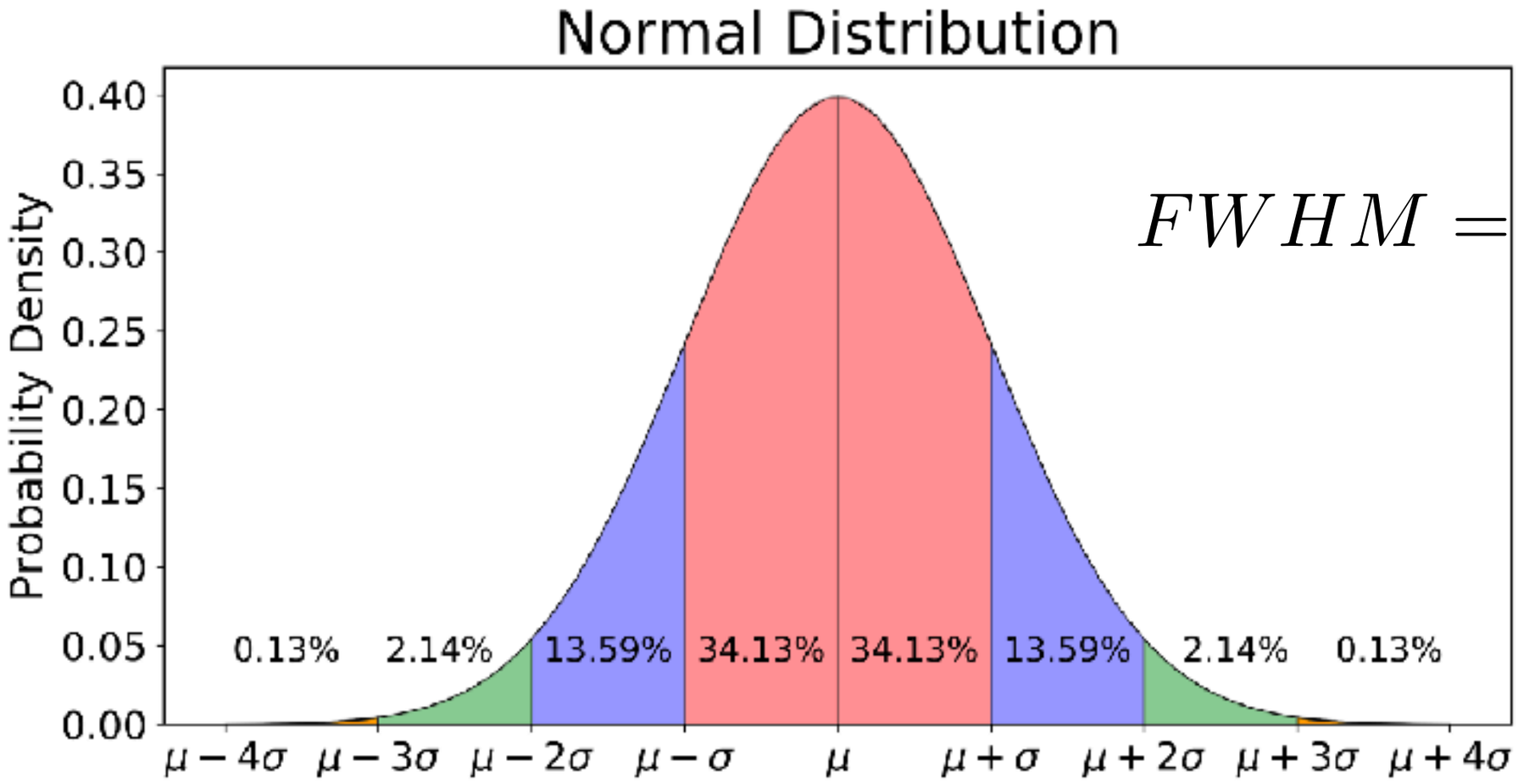
Plot the data with standard deviation of error,

Skewness, Kurtosis

भारतीय सांकेतिक विज्ञान संस्थ हैदराबाद
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Limits of Variability

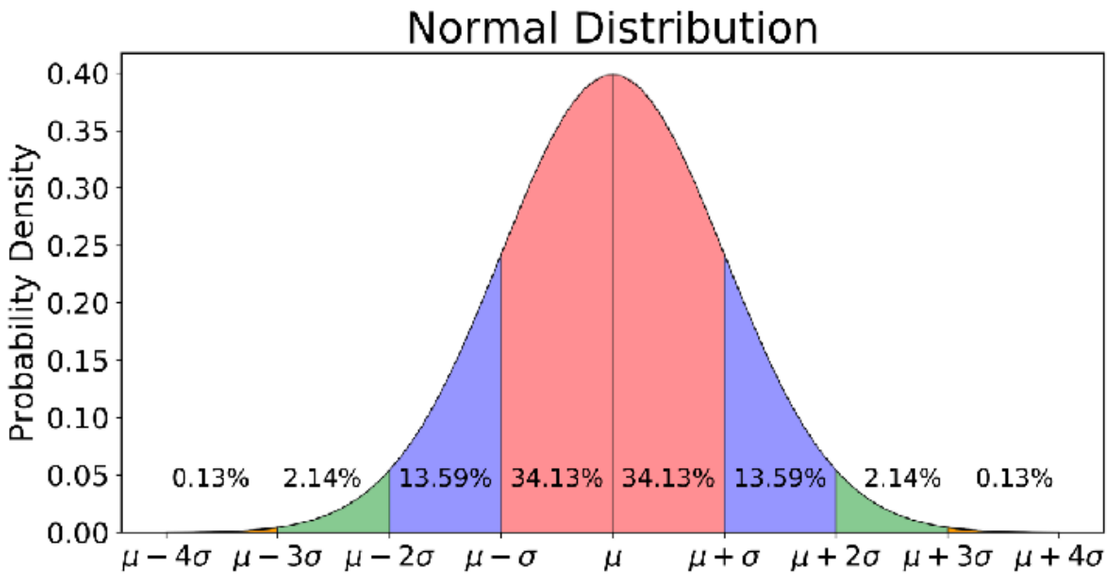
68–95–99.7 rule
For a normal distribution

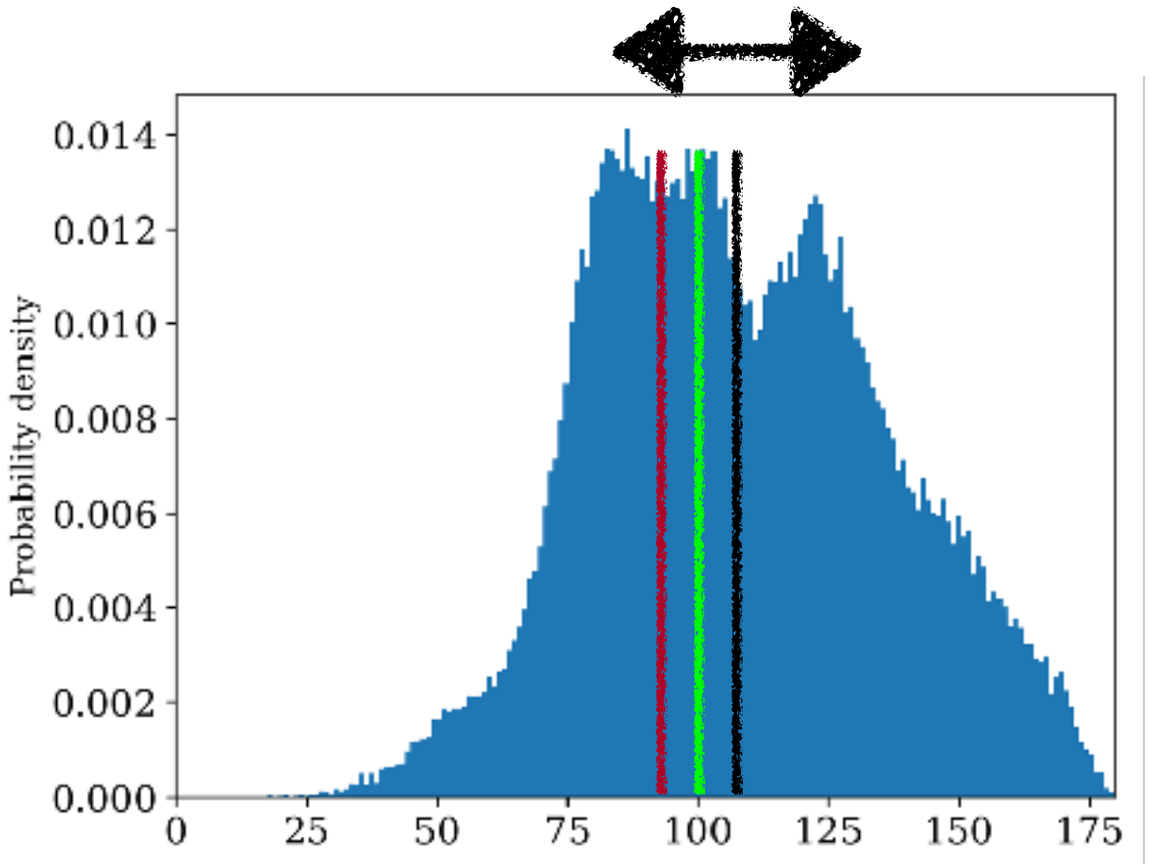


$$FWHM = 2\sqrt{2ln2}\ \sigma$$

भारतीय सांकेतिक विज्ञान संस्थ हैदराबाद
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Chebyshev inequality
Comparing the normal distribution with the any other distribution





mean= 108.16
mode= 79.86
median= 105.78
standard deviation= 28.14

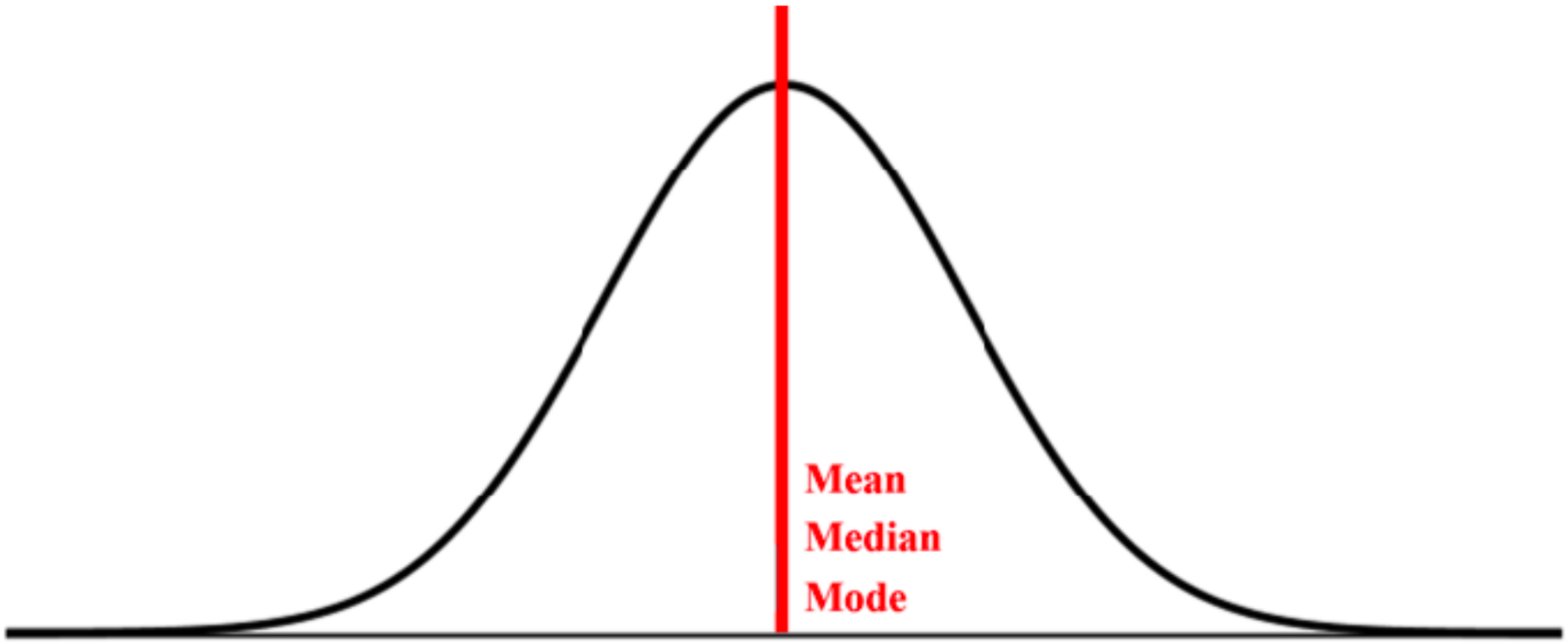
Absolute skewness

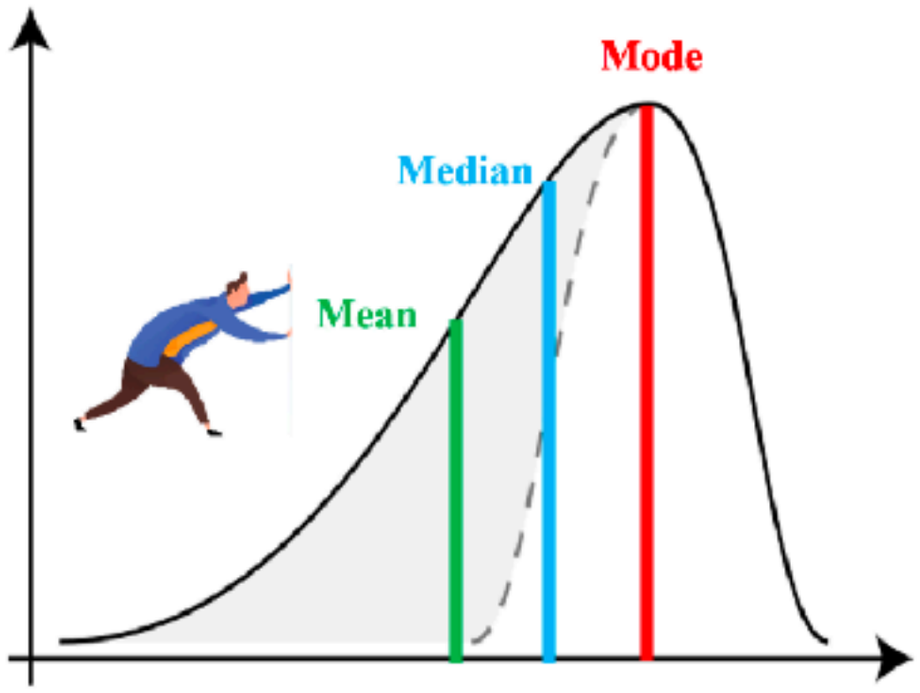
Mean - Mode

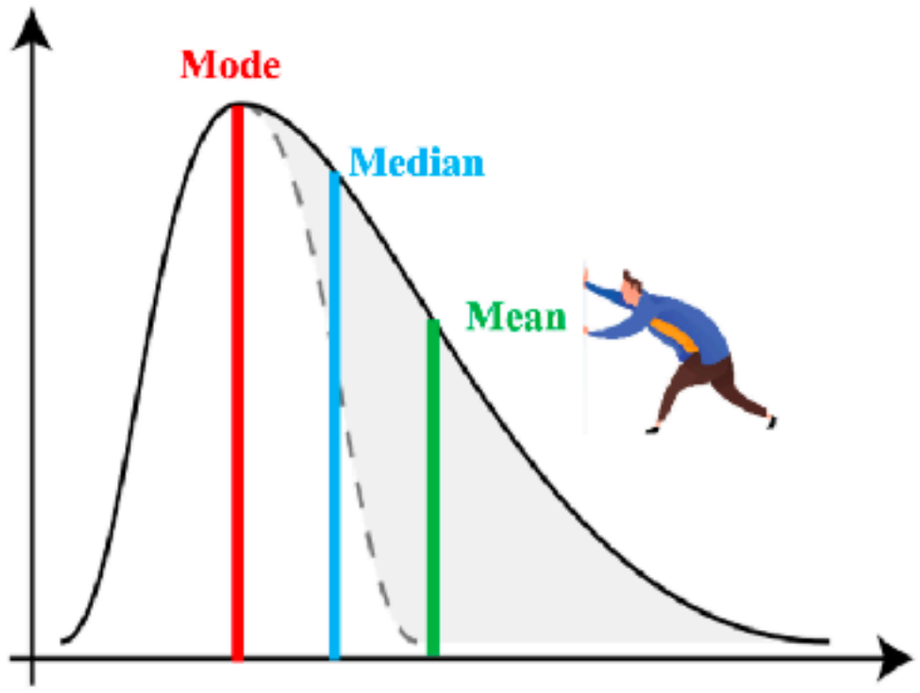
Karl Pearson coefficient of skewness

$$S_k = \frac{Mean - Mode}{Stanadrd\ deviation} = \frac{3(Mean - Median)}{Standard\ deviation}$$

Skewness



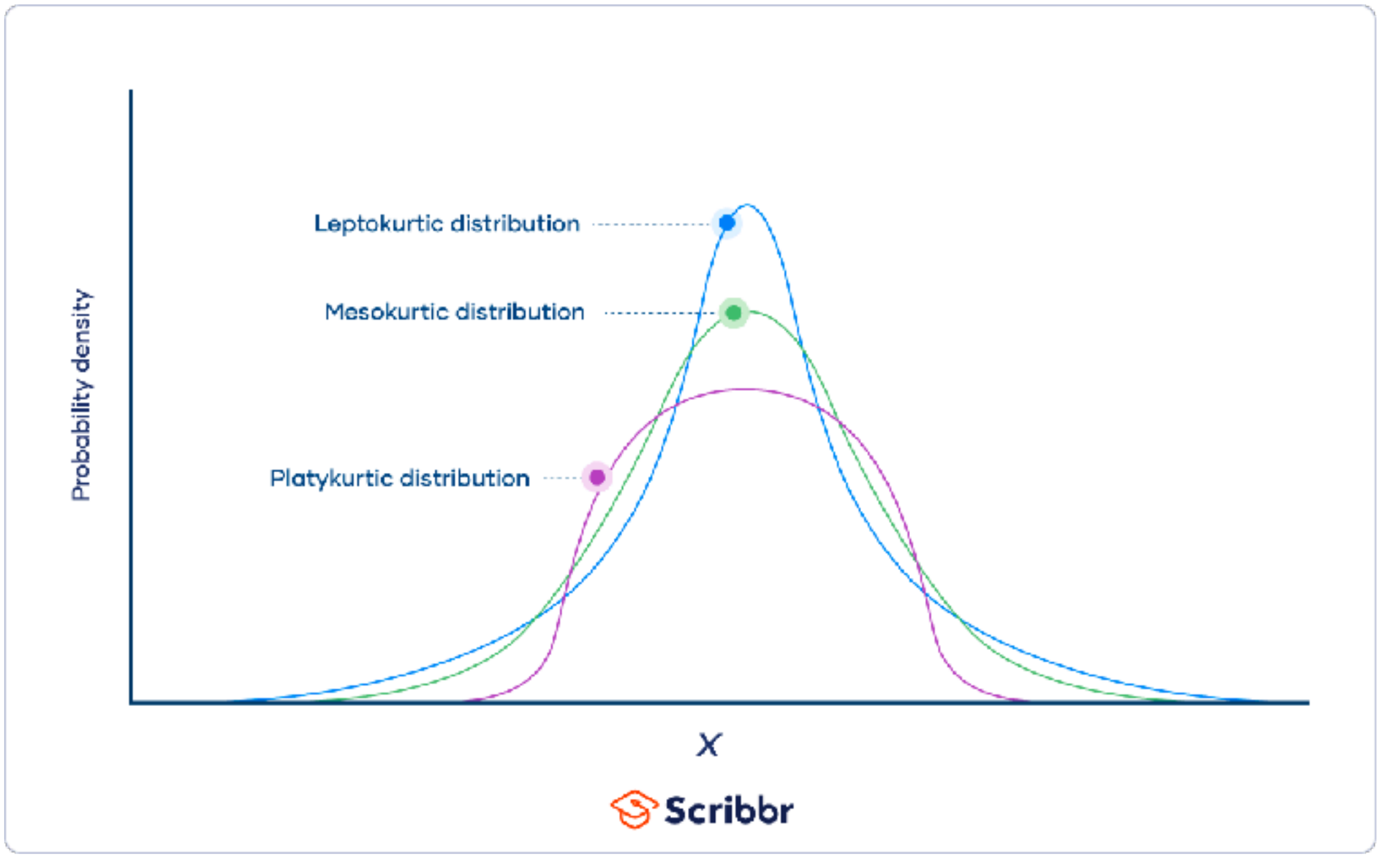




Reference
Towards Data Science

Kurtosis

Kurtosis refers to the degree of the peakedness of the hump in the distribution



भारतीय साङ्केतिक विज्ञान संस्थ हैदराबाद
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Next Class

2:30 PM Friday, 7 September 2022