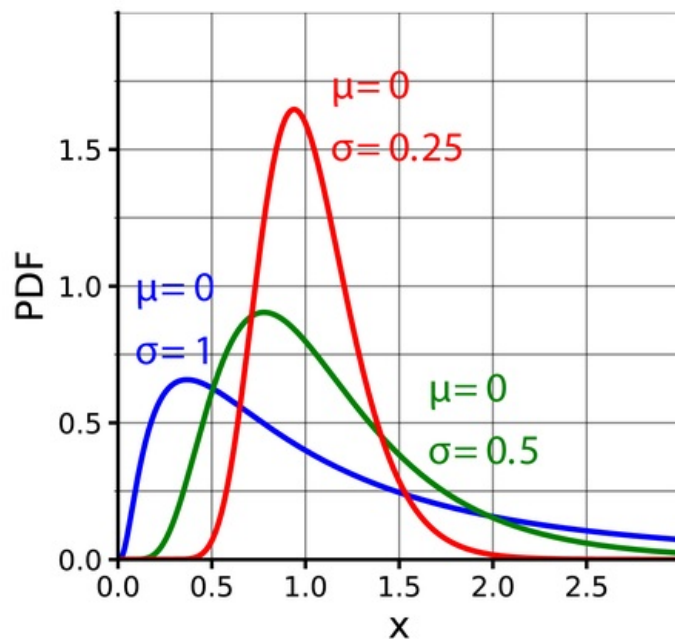


Log normal distribution: In probability theory, a log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then $Y = \ln(X)$ has a normal distribution.



We can see curve which more skewed towards right - Log normal distribution

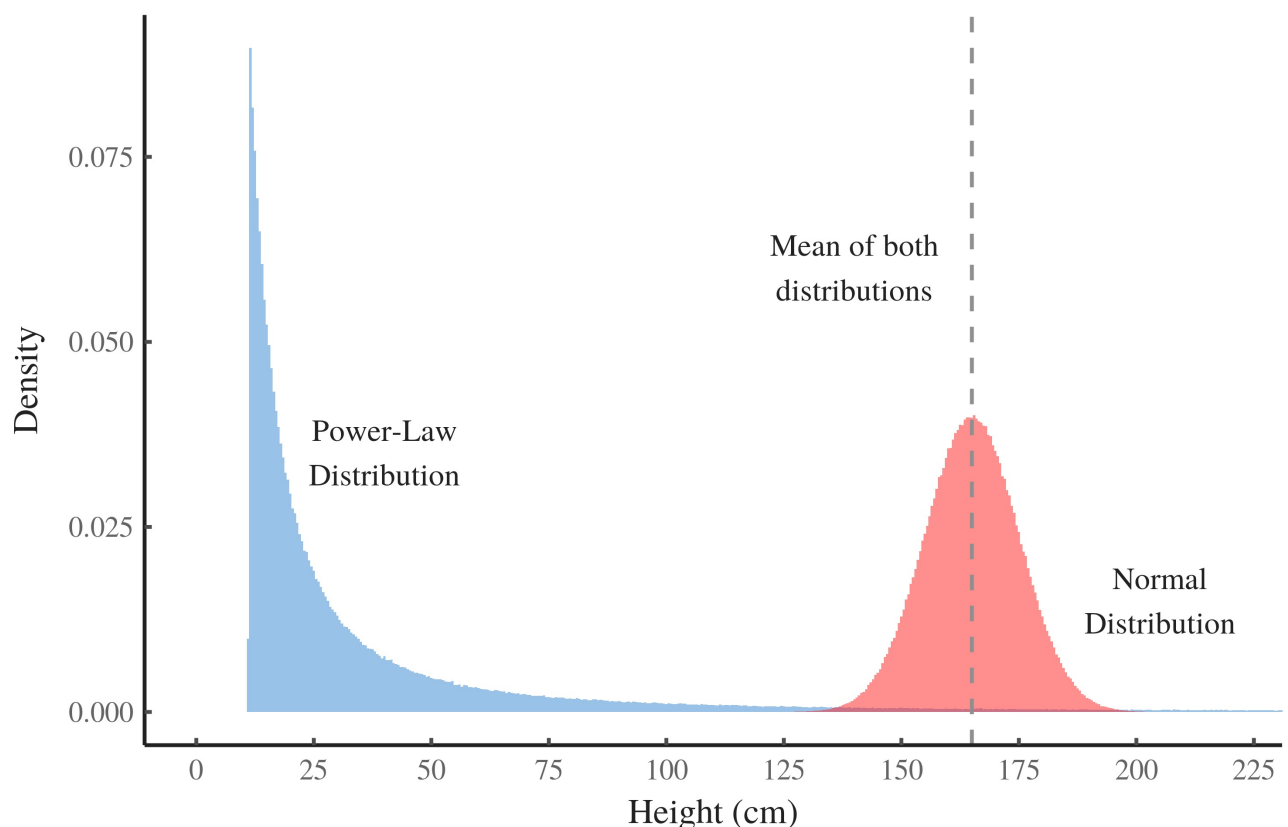
and, here, if we $\ln(x)$ - natural log, and value of x will be new then, if we plot again, it will be normal distribution - This is property of log normal distribution.

Where it will be used in data science?

- Suppose we are doing linear regression and we have dataset which follow normal curve then I will get effective result but, if my data set follow log normal distribution and if I provide it natural log then it will get converted into normal distribution and I can give that data to train my machine KNOWN AS TRANSFORMATION DATA TECHNIQUE.

So, now let's look at power law distribution - We know normal curve has a perfect bell shape, log normal curve is more skewed towards right. Then Power law distribution looks like:

Power Law vs. Normal Distribution of Human Height



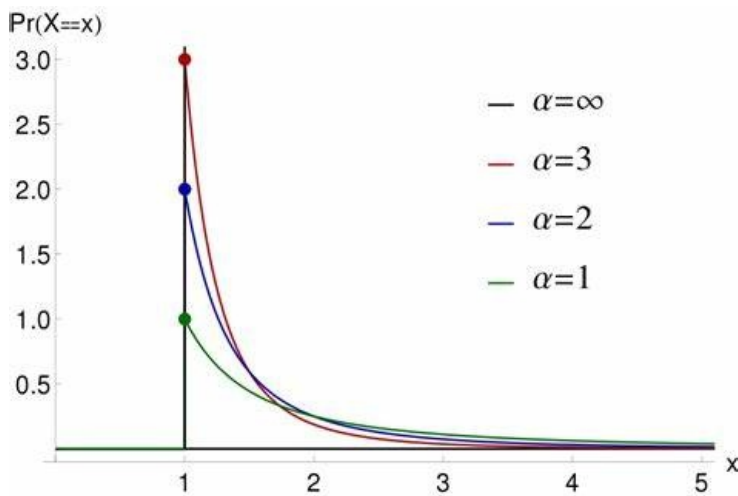
Example of log normal can be wealth distribution like there are very few people who are wealthiest. That's why there is long tail towards

Example of log normal can be wealth distribution like there are very few people who are wealthiest, that's why there is long tail towards right but curve has not that much height on that side. Similarly, Power law example can be Cricket or football match example. Always remember power law follows 80-20% rule like: In some Football team 20% of players from that team are responsible for winning 80% of matches.

Q. Can we convert log normal and Power law distribution to normal curve?

- Yes, to convert log normal to natural normal curve we can use natural log, for power law: Use Box Cox transform

Example: Pareto distribution is a non-gaussian distribution which follows power law distribution.



Hypothesis:

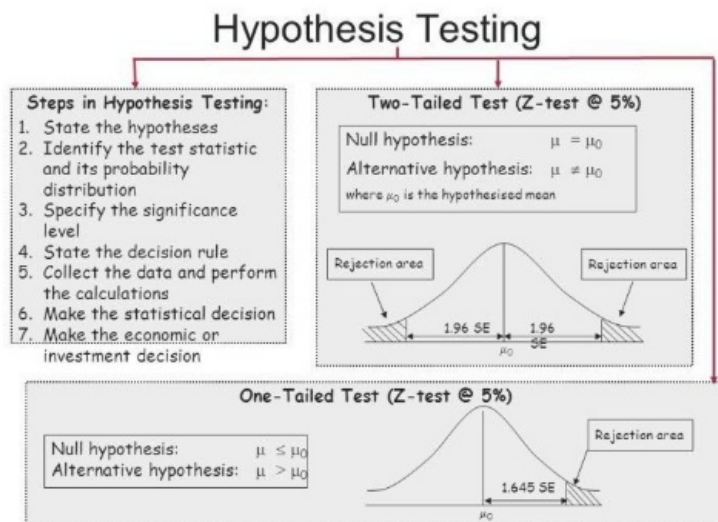
-Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter.

Example and step:

Statement: Tossing Coin is fair or not.

H_0 : Tossing coin is fair (Null hypothesis) H_1 : Tossing coin is not fair (alternative hypothesis)

Experiment: I will toss coin 100 times to see result.



In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js