Sampling types :

1-with replacement (huge data)

2-without replacement (sample data)

3-random (balanced data)

4-stratified sampling (unbalanced data)

When I can remove the outliers from dataset ?

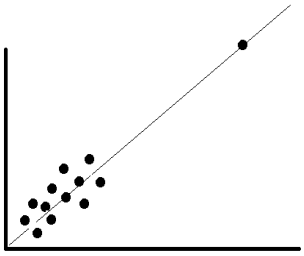
When the outlier is certainly part of the data and need a legitimate result to drop it

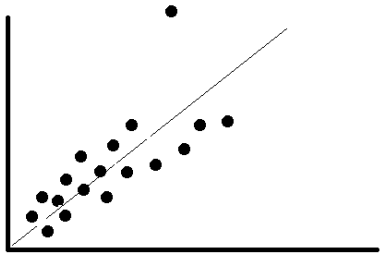
Ex: In this case, we cannot merely drop the outlier.

Try to run the analysis with or without the outlier and see how the result is.

Let's see in the example below;

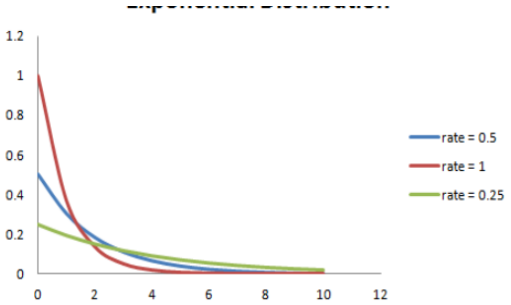
if we remove the outlier, the regression line will move



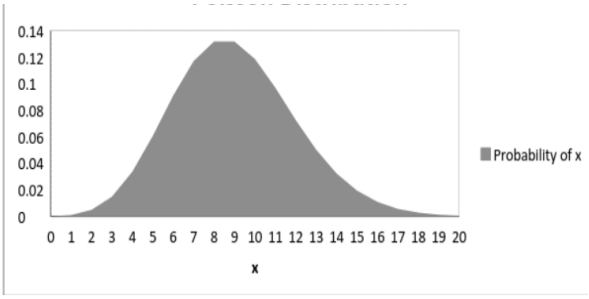


Types of Distribution :

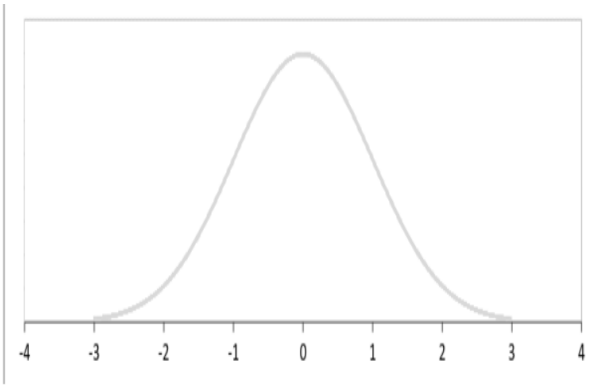
Exponential Distribution



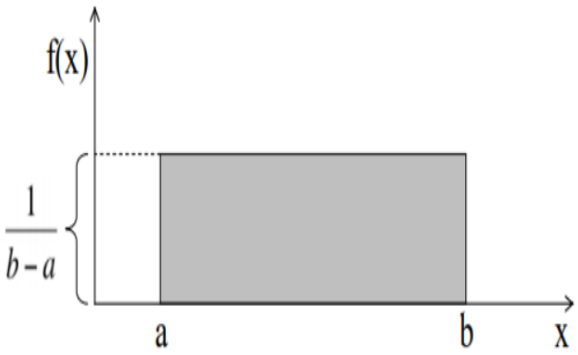
Poisson Distribution



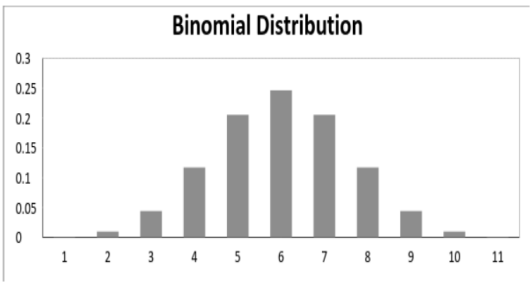
standard normal distribution



Uniform Distribution



Binomial Distribution



Convert To Normal Distribution :

- Log Transformation
- Square root Transformation
- Reciprocal Transformation
- Exponential Transformation
- Box-Cox Transformation

The Importance of Z_Score :

1_exam score

2_Newborn Weights

3_Blood Pressure