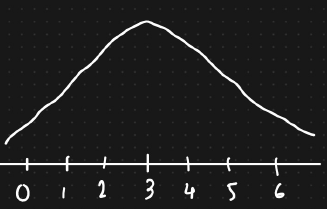
Standard Normal Distribution and Z Score:

Standard Normal distribution:

X = [1,2,3,4,5] follows a normal distribution

Mean = μ = 3

Standard deviation = σ = 1.414 = 1 approx



We will get this normal distribution

However if we want to convert the Gaussian normal distribution such that the mean = 0 and the standard deviation = 1 by applying a transformation technique. Then it will be transformed into a standard normal distribution from a normal distribution.

To convert, Z Score formula is used

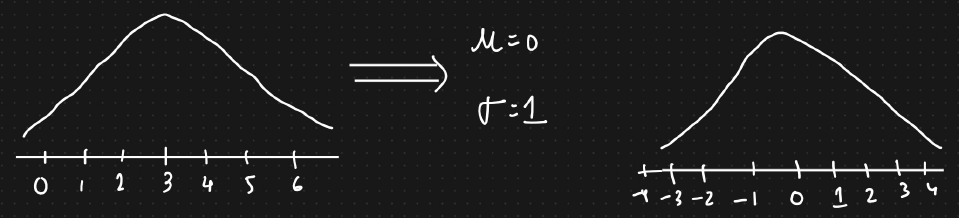
Z score =

Applying this, we will get another random variable y = […]

Applying Z transform on X = [1,2,3,4,5]

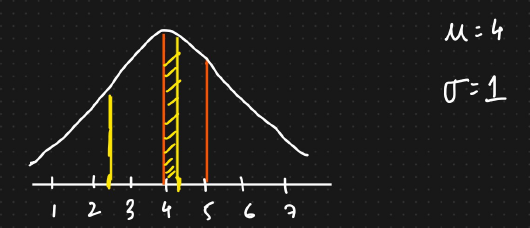
We get Y = [-2,-1,0,1,2]

Thus the mean of the new distribution gets zero



Notation:   
X SND (μ = 0, σ=1)

Meaning any random variable x belongs to a standard normal distribution if it’s mean is 0 and it’s standard deviation is 1



Q. How many standard deviation is 4.25 from the mean

Applying Z score formula = = 0.25

Example: Dataset

Features: Age, Weight, Height, Salary

These will all be in different units: years, kg, cm, rupees

Whenever we are solving a machine learning problem statement and whenever we create a machine learning model, the first rule is that we try to bring all the features in a same unit scale.

Many algorithms like linear regression and clustering algorithms require them to be in the same units so as to optimize the model, leading to better performance as well. Otherwise the datapoints will be scattered.

Thus the features are converted using standardization, i.e, applying Z score over every feature.



Z score = = =

Once we apply Z score, we get the values based on the distribution, bringing them down to the scale units.

Application: ML Model

Clustering algorithm calculates the distance, so if they are near, the calculation would be fast.