* Standard Normal Distribution

Let 'x' be a random variable which belongs to a Gaussian distribution having 'μ' mean and 'σ' standard deviation.

Then we can transform this 'x' to 'y' which will belong to Standard Normal Distribution having $\mu = 0$ and $\sigma = 1$, this process is called **standardization** and done by calculating **Z-score**.

$x \in$ Gaussian (μ, σ) → Transformation (standardization) using Z-score. → $y \in$ Normal standard

→ x to y called transformation.
→ y to x called inverse transformation.

$$\text{Z score} = \frac{x_i - \mu}{\sigma / \sqrt{n}}$$

here,

$\sigma/\sqrt{n}$ = Standard error

where n is sample size.

as in case of Gaussian to Normal standard transformation, we transform each value or features so 'n' tends to be 1, as area under probability distribution as 1.

So, Z-score formula for this specific transformation case is (n=1),

$$\text{Z-score} = \frac{x_i - \mu}{\sigma}$$

Ex→ $x = \{1, 2, 3, 4, 5\}$, $\mu = 3$, $\sigma = 1.414$

Zscores,
for 1 its -1.414 as $(1-3/1.414)$
Similarly for 2 its -0.707,
for 3 its 0, for 4 its 0.707, for 5 its 1.414

$\therefore y = \{-1.414, -0.707, 0, 0.707, 1.414\}$

Comparing x and y we can conclude that after standardization scope has decreased leading to faster calculation and easy visualization of data.

→ This standardization helps us build models efficiently as it decreases the scope for those datasets where each features range vary widely.

ex-

Age (yr)	Height (cm)	Weight (kg)
20	180	72
22	190	94
24	160	82

Here all these features have huge scope, so we should standardize it first.

→ Their is another way to scale random variables which is called **Normalization**.

→ In case of standardization range of scaled values determined by the formula outcome.

→ In case of Normalization range of scaled variable can be determined in advance.

One method to do Normalization is **min-max normalization** which gives the scaled variables withing range of 0 to 1.

And is used mostly in deep learning.

Another example is that pixels of image are kept within range of 0-255.

→ Normalization and standardization are part of procedure called **Feature scaling**.

→ Basically in normalization we are shifting our mean close to z-score, same with standardization.

→ Area under the curve can be calculated by Z-score and Z-table.