AI 110 MEC Academy

Normal Distribution Task

Session 10

Prepared by:

Kamel hassan kamel Soliman

Prepared for:

Eng Anas Emad

What is Normal Distribution?

The normal distribution, also known as the Gaussian distribution or bell curve, is a continuous probability distribution characterized by its symmetric, bell-shaped curve, it is defined by two parameters: the mean (μ) and the standard deviation (σ).

properties of the normal distribution

- **Symmetry:** The curve is symmetric around the mean, meaning that the left and right halves of the distribution are mirror images.
- **Mean, Median, and Mode**: For a normal distribution, these three measures of central tendency are all equal and located at the center of the distribution.
- **Asymptotic:** The tails of the distribution approach the horizontal axis but never actually touch it.
- **Unimodal:** There is only one peak in the distribution, representing the mean, median, and mode.

The probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where:

- x is a value from the dataset,
- μ is the mean,
- σ is the standard deviation,
- π is Pi (approximately 3.14159),
- e is the base of the natural logarithm (approximately 2.71828).

Types of normal distribution

1- Standard Normal Distribution:

- Mean (μ\muμ) = 0
- Standard Deviation (σ\sigmaσ) = 1
- The standard normal distribution is a special case of the normal distribution and is often used in statistics for standardizing data.

2- General Normal Distribution:

- Any normal distribution with mean (μ\muμ) and standard deviation (σ\sigmaσ).
- It can be expressed as N(μ,σ2)N(\mu, \sigma^2)N(μ,σ2).

3- Multivariate Normal Distribution:

- Generalization of the one-dimensional normal distribution to higher dimensions.
- Characterized by a mean vector and a covariance matrix.
- Used to describe the joint distribution of a set of random variables

4- Mixture of Normals:

- A distribution formed by combining multiple normal distributions.
- Useful in modeling data with multiple modes.

AI 110 MEC Academy

Converting non-normal distribution data into normally distributed data

1. Log Transformation: used when the data is positive and the distribution tends to be right-skewed. You can use the natural logarithm or the common logarithm.

```
import numpy as np

# Original data
data = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9])

# Log transformation
log_data = np.log(data)

print("Original Data:")
print(data)
print("\nLog Transformed Data:")
print(log_data)
```

```
Original Data:
[1 2 3 4 5 6 7 8 9]

Log Transformed Data:
[0.         0.69314718 1.09861229 1.38629436 1.60943791 1.79175947
 1.94591015 2.07944154 2.19722458]
```

2. Square Root Transformation: used for positive data that tends to be right skewed.

```
import numpy as np

# Original data
data = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9])

# Log transformation
sqrt_data = np.sqrt(data)

print("Original Data:")
print(data)
print("\n sqrt_data Transformed Data:")
print(sqrt_data)
```

```
Original Data:
[1 2 3 4 5 6 7 8 9]

 sqrt_data Transformed Data:
[1.         1.41421356 1.73205081 2.         2.23606798 2.44948974
 2.64575131 2.82842712 3.        ]
```