Thesis 3

Anna Lucia Lamacchia
Statistics

January 11, 2024

The Gaussian Distribution

1 Meaning

The Gaussian distribution, commonly known as the normal distribution, holds a central and foundational position in probability theory and statistics. At its essence, the Gaussian distribution represents a continuous probability distribution that is symmetrically shaped, forming the well-known bell curve. This symmetry signifies that the probabilities of observing values to the left and right of the mean are equal. The meaning of the Gaussian distribution extends beyond its mathematical representation; it serves as an abstraction capturing the inherent variability and randomness present in countless natural and human phenomena. Its bell-shaped curve is a manifestation of the central limit theorem, indicating that, under certain conditions, the distribution of the sum (or average) of a large number of independent, identically distributed random variables tends to follow a Gaussian distribution. In practical terms, the Gaussian distribution often emerges as an idealized model for describing the variability in data, making it a fundamental tool in statistical analysis, hypothesis testing, and various scientific disciplines. The concept of standardization, with its reliance on the mean and standard deviation, further emphasizes the Gaussian distribution's meaning as a standardized framework for understanding and characterizing uncertainty and variability in diverse datasets.

2 Proof

Theorem Gaussian Distribution. The Gaussian distribution is characterized by its probability density function (PDF), and its standard form is given by:

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

where μ is the mean and σ is the standard deviation.

Proof.

2.1 Step 1: Derivation of Standard Normal Distribution PDF

The standard normal distribution PDF is given by:

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

To derive this, consider the cumulative distribution function (CDF) of the standard normal distribution:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t) dt$$

The derivative of this CDF with respect to x gives the standard normal distribution PDF:

$$\phi(x) = \frac{d}{dx}\Phi(x)$$

2.2 Step 2: Transformation to General Gaussian Distribution

For a general Gaussian distribution with mean μ and standard deviation σ, use the linear transformation:

$$Z = \frac{X - \mu}{\sigma}$$

where X is a random variable from the standard normal distribution. The probability density function of Z is derived by applying the transformation rule:

$$f_Z(z) = f_X(x) \left| \frac{dx}{dz} \right|$$

Substitute $X = \sigma Z + \mu$ into the expression to get the PDF of the general Gaussian distribution.

2.3 Step 3: Final Result

After simplifying the expression, the PDF of the general Gaussian distribution is obtained:

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

This completes the formal derivation of the Gaussian distribution PDF. $\square$

3 Simulation in Python

This Python code uses NumPy to generate random samples from a normal distribution and then plots a histogram of the samples alongside the theoretical Gaussian distribution. The visual comparison illustrates how the simulated samples align with the expected Gaussian distribution.

```
import numpy as np
import matplotlib.pyplot as plt

# Parameters
mu = 0
sigma = 1
num_samples = 1000

# Generate random samples from a normal distribution
samples = np.random.normal(mu, sigma, num_samples)

# Plot histogram of the samples
plt.hist(samples, bins=30, density=True, alpha=0.5, color='blue')

# Plot theoretical Gaussian distribution
xmin, xmax = plt.xlim()
x = np.linspace(xmin, xmax, 100)
pdf = (1 / (sigma * np.sqrt(2 * np.pi))) * np.exp(-0.5 * ((x - mu) / sigma)**2)
plt.plot(x, pdf, 'k', linewidth=2)

# Set plot labels and title
plt.xlabel('Value')
plt.ylabel('Probability Density')
plt.title('Gaussian Distribution Simulation')
plt.show()
```

Figure 1: Code

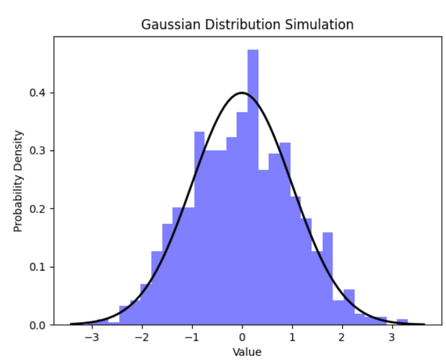

Figure 2: Chart