# Lecture Scribe: Gaussian Simulation and Density Estimation

Name: Masoom Choksi
AUID: AU2440263

## Course Information

**Course:** CSE400 – Probability and Random Variables

## 1 In-Class Activity: Gaussian Simulation and Density Estimation

This lecture covers:

- Simulation of Gaussian random variables
- Visualization of probability distributions
- Gaussian density estimation from data
- Application of Gaussian density estimation to real-world problems

## 2 Why Do We Need Probability Models?

Real-world measurements are uncertain due to noise and randomness. Examples include:

- Sensor measurements
- Network packet delays
- Image sensor noise

Such uncertainty is modeled using **Gaussian random variables**, and their distributions are estimated from data.

## 3 Gaussian Random Variable Simulation (Known Parameters)

### 3.1 Assumptions

- The random variable is Gaussian.
- The mean and standard deviation are known.
- Parameters are not estimated at this stage.

### 3.2 Definition (Gaussian Random Variable)

A random variable $X$ is Gaussian with mean $\mu$ and standard deviation $\sigma$ if its probability density function (PDF) is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left( -\frac{(x-\mu)^2}{2\sigma^2} \right).$$

### 3.3 Simulation Parameters

$$N = 100000, \quad \mu_{\text{true}} = 0.0, \quad \sigma_{\text{true}} = 4.0.$$

### 3.4 Simulation

Samples are generated as

$$x_i = \sigma_{\text{true}} Z_i + \mu_{\text{true}},$$

where $Z_i \sim \mathcal{N}(0, 1)$.

## 4 Visualizing the Distribution

### 4.1 Histogram Construction

- Bin range: $[-15, 25]$
- Bin width: 1.0
- Histogram normalized to form a density

### 4.2 True PDF

For visualization, the true Gaussian PDF is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_{\text{true}}} \exp\left( -\frac{(x-\mu_{\text{true}})^2}{2\sigma_{\text{true}}^2} \right).$$

## 5 Reality Check: Unknown Parameters

### 5.1 Assumptions

- The distribution is Gaussian.
- The parameters (mean and variance) are unknown.
- Only data samples are observed.

Gaussian density estimation is used to estimate the underlying distribution from data.

## 6 Gaussian Density Estimation from Data

### 6.1 Simulation Parameters

$$N_e = 200, \quad \mu_{\text{true}} = 0.0, \quad \sigma_{\text{true}} = 4.0.$$

### 6.2 Observed Data Model

- True samples: $x_i \sim \mathcal{N}(\mu_{\text{true}}, \sigma^2_{\text{true}})$
- Measurement noise: $n_i \sim \mathcal{N}(0, \sigma_n^2)$, where $\sigma_n = 1.5$

Observed data:

$$x_i^{\text{obs}} = x_i + n_i.$$

### 6.3 Parameter Estimation

The Gaussian parameters are estimated from observed data as

$$\hat{\mu} = \frac{1}{N_e} \sum_{i=1}^{N_e} x_i^{\text{obs}}, \qquad \hat{\sigma} = \sqrt{\frac{1}{N_e} \sum_{i=1}^{N_e} \left(x_i^{\text{obs}} - \hat{\mu}\right)^2}.$$

### 6.4 Estimated PDF

$$\hat{f}_X(x) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp\left(-\frac{(x-\hat{\mu})^2}{2\hat{\sigma}^2}\right).$$

## 7 Application: Packet Delay in Networks

### 7.1 Assumptions

- Packet delay is modeled as a single Gaussian random variable.
- Delays are non-negative (physical constraint).

### 7.2 Simulation Parameters

$$N = 150, \quad \mu = 45.0 \text{ ms}, \quad \sigma = 8.0 \text{ ms}.$$

### 7.3 Data Generation

Packet delays are generated as

$$D_i \sim \mathcal{N}(\mu, \sigma^2), \quad D_i > 0.$$

### 7.4 Density Estimation

Estimated parameters:

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} D_i, \qquad \hat{\sigma} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (D_i - \hat{\mu})^2}.$$

Estimated Gaussian density:

$$\hat{f}_D(d) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} \exp\left(-\frac{(d-\hat{\mu})^2}{2\hat{\sigma}^2}\right).$$

## 8 Key Takeaways

- Gaussian simulation uses randomized algorithms.
- Density estimation learns distributions from data.
- The same method applies across different applications.
- The objective is to estimate **distributions**.