CSE400 — Fundamentals of Probability in Computing
Lecture 8 Scribe (Exam-Preparation Style)

Topic: Gaussian Random Variable (Definitions, Standard Forms, Q/Φ Relations, Worked Examples, and CDF-Based Moment Analysis)

1 Gaussian Random Variable — Definition and Properties

1.1 Definition (Gaussian Random Variable)

A random variable X is Gaussian if its PDF can be written in the general form:

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

and it is denoted as:

$$X \sim N(m, \sigma^2)$$

This identifies:

- m as the mean parameter
- σ^2 as the variance parameter

2 Gaussian Random Variable — Standard Forms

This lecture introduces standard integrals and special functions used to compute Gaussian probabilities.

2.1 Error Function and Complementary Error Function

Definition: Error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt$$

Definition: Complementary error function

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)\, dt$$

2.2 Φ-function and Q-function

Definition: Φ-function (CDF of Standard Normal)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right)\, dt$$

Definition: Q-function (Gaussian Tail Function)

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\left(-\frac{z^2}{2}\right)\, dz$$

Note (given in lecture)

$$Q(x) = 1 - \Phi(x)$$

3 Relation Between Φ-function and Q-function

This section derives how to compute Gaussian CDFs and tail probabilities using $\Phi(\cdot)$ and $Q(\cdot)$.

3.1 Evaluating the CDF $F_X(x)$

For a Gaussian random variable $X \sim N(m, \sigma^2)$, the CDF is:

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)\, dy$$

The lecture expresses the standardized form as evaluating the Φ-function at the standardized point.

3.2 Evaluating Tail Probabilities

For tail probabilities:

$$\Pr(X > x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)\, dy$$

3.3 Key Identities (Given)

The lecture states the following key identities:

$$Q(x) = 1 - \Phi(x)$$

$$F_X(x) = 1 - Q\left(\frac{x-m}{\sigma}\right)$$

3.4 Lecture Guidance on When to Use Which Function

To evaluate the CDF of a Gaussian RV, evaluate the Φ-function at the points.

The Q-function is more natural for evaluating probabilities of the form $\Pr(X > x)$.

4 Worked Example — Gaussian Probability Expressions Using Q-functions

4.1 Problem Statement (Given)

A random variable X has a PDF given by a Gaussian form. The lecture then asks:

Find each of the following probabilities and express in terms of Q-functions:

- $\Pr(X < 0)$
- $\Pr(X > 4)$
- $\Pr(|X + 3| < 2)$
- $\Pr(|X - 2| > 1)$

Hint (given): Use $Q(x)$ for right-tail and $\Phi(x)$ for left-tail probabilities.

4.2 Step 1 — Identify Parameters by Comparing with Gaussian PDF

The solution begins:

Comparing with the Gaussian PDF we identify Gaussian relations:

$$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right), \qquad \Pr(X > x) = Q\left(\frac{x - m}{\sigma}\right)$$

4.3 Step-by-Step Solution for Each Subpart

4.3.1 (3) Compute $\Pr(|X + 3| < 2)$

The lecture explicitly shows:

Step 1. Start with the inequality:

$$|X + 3| < 2$$

Step 2. Convert absolute value inequality to a double inequality:

$$-2 < X + 3 < 2$$

Step 3. Subtract 3 throughout:

$$-5 < X < -1$$

Step 4. Convert interval probability to CDF form:

$$\Pr(-5 < X < -1) = F_X(-1) - F_X(-5)$$

Step 5. Express in terms of Φ:

$$F_X(-1) - F_X(-5) = \Phi(1) - \Phi(-1)$$

Step 6. Use the identity given in the solution:

$$\Phi(-x) = 1 - \Phi(x)$$

Step 7. Apply it:

$$\Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1$$

Step 8. Convert to Q-form using $Q(x) = 1 - \Phi(x)$:

$$2\Phi(1) - 1 = 1 - 2Q(1)$$

Final:

$$\Pr(|X + 3| < 2) = 1 - 2Q(1)$$

4.3.2 (4) Compute $\Pr(|X - 2| > 1)$

The lecture explicitly shows:
Step 1. Start:

$$|X - 2| > 1$$

Step 2. Convert to a union of inequalities:

$$X < 1 \quad \text{or} \quad X > 3$$

Step 3. Convert to probability:

$$\Pr(|X - 2| > 1) = F_X(1) + \Pr(X > 3)$$

Step 4. Express using Q-function and CDF relation:

$$F_X(1) = 1 - Q(2)$$

and

$$\Pr(X > 3) = Q(3)$$

Step 5. Combine:

$$\Pr(|X - 2| > 1) = 1 - Q(2) + Q(3)$$

Final:

$$\Pr(|X - 2| > 1) = 1 - Q(2) + Q(3)$$

5 Gaussian Random Variable — Application Examples

The lecture lists the following application examples:

- Thermal noise voltage in an electronic circuit
- Measurement error in a sensor reading:

$$\text{Measured Value} = \text{True Value} + \text{Gaussian noise}$$

- Packet delay variation (jitter) in a communication network, where delays fluctuate around a mean value due to random congestion

6 Exercise: Problem Solving 2 (CDF Analysis)

6.1 Problem Statement (Given CDF)

The CDF of a random variable is:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 < x < 1 \\ 1, & x > 1 \end{cases}$$

Find:

- Mean (μ_X)
- Variance (σ^2)
- Skewness (c_s)
- Kurtosis (c_k)

6.2 Step-by-Step Solution (Exactly as Given)

Step 1: Compute the PDF from the CDF

The lecture states:

$$f_X(x) = \frac{d}{dx} F_X(x) = 2x, \qquad 0 < x < 1$$

Step 2: Mean

The lecture indicates "Mean:" but the explicit algebra is not fully visible in the provided context snippet.

Step 3: Variance

The lecture provides the variance computation line in a condensed form:

$$\sigma_X^2 = \int_0^1 x^2(2x)\,dx - \mu_X^2$$

and gives the final simplified value:

$$\sigma_X^2 = \frac{1}{18}$$

Step 4: Skewness

The lecture provides:

$$c_s = \frac{E(X-\mu)^3}{\sigma^3} = -\frac{2}{5}$$

Step 5: Kurtosis

The lecture provides:

$$c_k = \frac{E(X-\mu)^4}{\sigma^4} = \frac{12}{5}$$

6.3 Conclusion (Given)

The lecture concludes:

- Distribution is left-skewed
- Platykurtic (lighter tails than Gaussian)

7 Gaussian Modeling: From Noise to Math (Conceptual Notes)

7.1 Motivation (Given)

The lecture states:

- The world is noisy.
- Sensors do not give "flat" lines (thermal noise, interference).
- Repeated measurements form a "cloud" of uncertainty.

7.2 Gaussian Simulation Model (Given)

The lecture models noise using:

$$X \sim N(1, \sigma^2)$$

and provides the generative formula:

$$X = \sigma Z + 1$$

where Z is "standard randomness."

8 Density Estimation: Learning from Data (Conceptual Notes)

The lecture states:

- In real settings, we do not know μ or σ.
- We observe raw noisy samples (histogram).
- We use sample statistics to estimate the distribution.
- Key concept: We estimate probability distributions, not deterministic functions; this smooths noise into a predictive tool.

9 Applications: Network Systems (Packet Delay Modeling)

The lecture describes:

- Case study: 500 ICMP pings (latency analysis).
- Delay varies due to queuing, congestion, and routing hops.
- Engineering goal: typical delay μ and jitter σ.

10 Applications: Broader Computing Contexts (Given)

The lecture lists examples:

- Image processing denoising (Gaussian priors)
- Tail latency in distributed systems (99th percentile)
- Sensor fusion (Kalman filters; GPS uncertainty)