ECE 302: Lecture 4.7 Gaussian Random Variable

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University

PURDUE
UNIVERSITY

Outline

Overall schedule:

- Continuous random variables, PDF
- CDF
- Expectation
- Mean, mode, median
- Common random variables
  - Uniform
  - Exponential
  - **Gaussian**
- Transformation of random variables
- How to generate random numbers

Today's lecture:

- Definition of Gaussian
- Mean and variance
- Skewness and kurtosis
- Origin of Gaussian

Definition

Definition

Let X be an Gaussian random variable. The PDF of X is

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

where (μ, σ^2) are parameters of the distribution. We write

$$X \sim \text{Gaussian}(\mu, \sigma^2) \qquad \text{or} \qquad X \sim \mathcal{N}(\mu, \sigma^2)$$

to say that X is drawn from a Gaussian distribution of parameter (μ, σ^2).

Interpreting the mean and variance

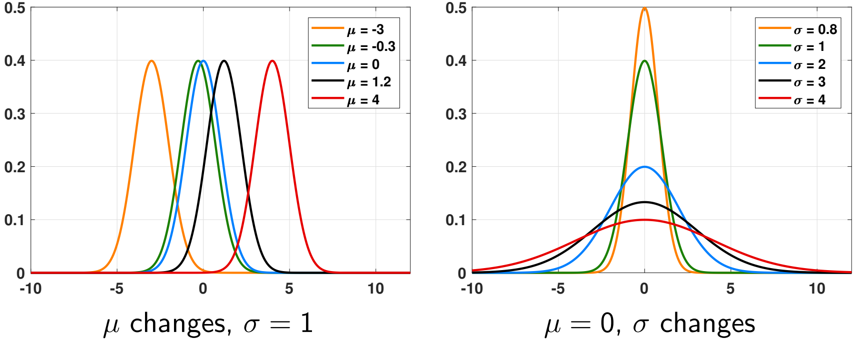


Figure: A Gaussian random variable with different μ and σ

Proving the mean

Theorem

If $X \sim \mathcal{N}(\mu, \sigma^2)$*, then*

$$\mathbb{E}[X] = \mu, \quad \textit{and} \quad \mathrm{Var}[X] = \sigma^2. \tag{2}$$

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&\overset{(a)}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y + \mu) e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \\
&\overset{(b)}{=} \\
&\overset{(c)}{=} \mu.
\end{aligned}$$

Proving the variance

Theorem

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$\mathbb{E}[X] = \mu, \quad \textit{and} \quad \mathrm{Var}[X] = \sigma^2. \tag{3}$$

$$\begin{aligned}
\mathrm{Var}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&\overset{(a)}{=} \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy, \quad \text{by letting} \quad y = \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left(-y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} \right) + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \\
&= \\
&= \sigma^2
\end{aligned}$$

Standard Gaussian PDF

Definition

A **standard Gaussian** (or standard Normal) random variable X has a PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{4}$$

That is, $X \sim \mathcal{N}(0, 1)$ is a Gaussian with $\mu = 0$ and $\sigma^2 = 1$.

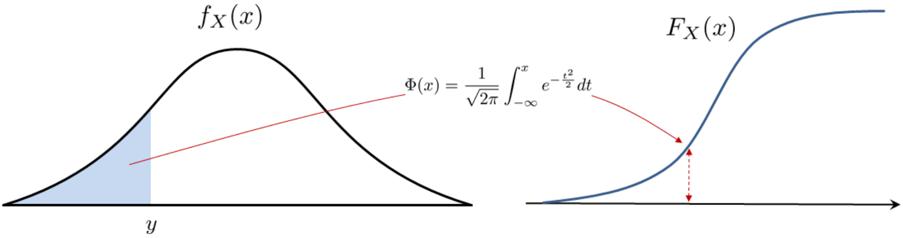


Figure: Definition of the CDF of the standard Gaussian $\Phi(x)$.

Standard Gaussian CDF

Definition

The **CDF** of the standard Gaussian is defined as the $\Phi(\cdot)$ function

$$\Phi(x) \stackrel{\text{def}}{=} F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt. \tag{5}$$

The standard Gaussian's CDF is related to a so-called **error function** which is defined as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2}\, dt. \tag{6}$$

It is quite easy to link $\Phi(x)$ with $\text{erf}(x)$:

$$\Phi(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right], \qquad \text{and} \qquad \text{erf}(x) = 2\Phi(x\sqrt{2}) - 1.$$

CDF of arbitrary Gaussian

Theorem (CDF of an arbitrary Gaussian)

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then,

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right). \tag{7}$$

We start by expressing $F_X(x)$:

$$F_X(x) = \qquad\qquad\qquad\qquad .$$

Substituting $y = \frac{t-\mu}{\sigma}$, and using the definition of standard Gaussian, we have

$$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy =$$

Other results

$$\mathbb{P}[a < X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \tag{8}$$

To see this, note that

$$\mathbb{P}[a < X \leq b] = \mathbb{P}[X \leq b] - \mathbb{P}[X \leq a] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Corollary

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, the following results hold:

- $\Phi(y) = 1 - \Phi(-y)$.
- $\mathbb{P}[X \geq b] = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$.
- $\mathbb{P}[|X| \geq b] = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right) + \Phi\left(\frac{-b-\mu}{\sigma}\right)$

Skewness and Kurtosis

Definition

For a random variable X with PDF $f_X(x)$, define the following **central moments** as

$$
\begin{aligned}
\text{mean} &= \mathbb{E}[X] \stackrel{\text{def}}{=} \mu, \\
\text{variance} &= \mathbb{E}\left[(X-\mu)^2\right] \stackrel{\text{def}}{=} \sigma^2, \\
\text{skewness} &= \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \stackrel{\text{def}}{=} \gamma, \\
\text{kurtosis} &= \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] \stackrel{\text{def}}{=} \kappa.
\end{aligned}
$$

Skewness

What is skewness?

- $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$.
- Measures how **asymmetrical** the distribution is.
- Gaussian has skewness 0.

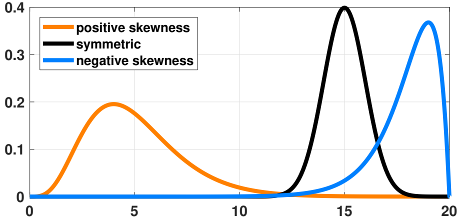


Figure: Skewness of a distribution measures how asymmetric the distribution is. In this example, the skewness are: orange = 0.8943, black = 0, blue = -1.414.

What is kurtosis?

- $\kappa = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$.
- Measures how **heavy tail** is. Gaussian has kurtosis 3.
- Some people prefer *excess kurtosis* $\kappa - 3$. Gaussian has excess kurtosis 0.

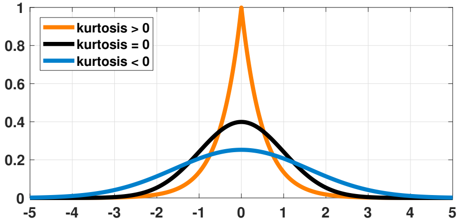


Figure: Kurtosis of a distribution measures how heavy tail the distribution is. In this example, the (excess) kurtosis are: orange = 2.8567, black = 0, blue = -0.1242.

Skewness and Kurtosis

Random variable	Mean μ	Variance σ^2	Skewness γ	Excess kurtosis $\kappa - 3$
Bernoulli	p	$p(1-p)$	$\frac{1-2p}{\sqrt{p(1-p)}}$	$\frac{1}{1-p} + \frac{1}{p} - 6$
Binomial	np	$np(1-p)$	$\frac{1-2p}{\sqrt{np(1-p)}}$	$\frac{6p^2-6p+1}{np(1-p)}$
Geometric	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{2-p}{\sqrt{1-p}}$	$\frac{p^2-6p+6}{1-p}$
Poisson	λ	λ	$\frac{1}{\sqrt{\lambda}}$	$\frac{1}{\lambda}$
Uniform	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	0	$-\frac{6}{5}$
Exponential	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	2	6
Gaussian	μ	σ^2	0	0

Table: The first few moments of commonly used random variables.

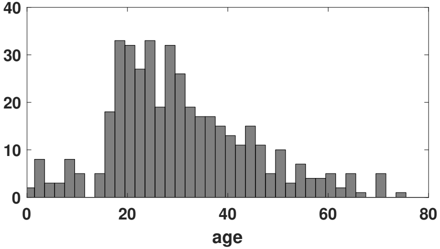
Example: Titanic

On April 15, 1912, RMS Titanic sank after hitting an iceberg. This has killed 1502 out of 2224 passengers and crew. A hundred years later, we want to analyze the data. On `https://www.kaggle.com/c/titanic/` there is a dataset collecting the identities, age, gender, etc of the passengers.

Statistics	Group 1 (Died)	Group 2 (Survived)
Mean	30.6262	28.3437
Standard Deviation	14.1721	14.9510
Skewness	0.5835	0.1795
Excess Kurtosis	0.2652	-0.0772

Example: Titanic

- Mean and standard deviation cannot tell the difference.
- Skewness and kurtosis can tell the difference.



Group 1 (died) Group 2 (survived)

Figure: The Titanic dataset `https://www.kaggle.com/c/titanic/`.

Origin of Gaussian

- Where does Gaussian come from?
- Why are they so popular?
- Why do they have bell shapes?

What is the origin of Gaussian?

- When we **sum** many independent random variables, the resulting random variable is a Gaussian.
- This is known as the **Central Limit Theorem**. The theorem applies to *any* random variable.
- Summing random variables is equivalent to **convolving** the PDFs. Convolving PDFs infinitely many times yields the bell shape.

The experiment of throwing many dices

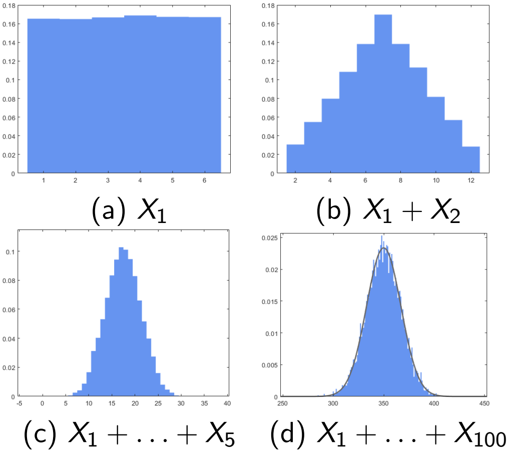


Figure: When adding uniform random variables, the overall distribution is becoming like a Gaussian.

Sum of X and Y = Convolution of f_X and f_Y

Example: Two rectangles to give a triangle:

We will show this result in a later lecture:

$$(f_X * f_X)(x) = \int_{-\infty}^{\infty} f_X(\tau) f_X(x - \tau) d\tau.$$

If you convolve infinitely many times

Then in Fourier domain you will have

$$\mathcal{F}\left\{(f_X * f_X * \ldots * f_X)\right\} = \mathcal{F}\{f_X\} \cdot \mathcal{F}\{f_X\} \cdot \ldots \cdot \mathcal{F}\{f_X\}.$$

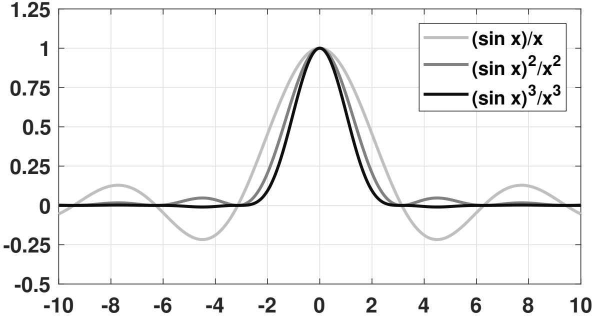


Figure: Convolving the PDF of a uniform distribution is equivalent to multiplying their Fourier transforms in the Fourier space. As the number of convolution grows, the product is gradually becoming Gaussian.

Origin of Gaussian

What happens if you convolve a PDF infinitely many times?

- You will get a Gaussian.
- This is known as the central limit theorem.

Why are Gaussians everywhere?

- We seldom look at individual random variables. We often look at the sum/average.
- Whenever we have a sum, **Central Limit Theorem** kicks in.
- Summing random variables is equivalent to **convolving** the PDFs. Convolving PDFs infinitely many times yields the bell shape.
- This result applies to *any* random variable, as long as they are independently summed.

Questions?