Continuous Random Variables

Vibha Mane
The COSINE Lab
Department of Electrical and Computer Engineering

Topics

- Definition of Continuous Random Variables
- Characterization of Continuous Random Variables
  - Cumulative Distribution Function (CDF)
  - Probability Density Function (PDF)
- Examples of Continuous Random Variables
  - Uniform
  - Exponential
  - Normal (or Gaussian)
  - Computing probabilities with the normal distribution

I. Meaning of a Random Variable

- A **random variable (RV)** is a function whose domain is the sample space Ω.
- To every elementary outcome ω in Ω, we assign a real value X(ω).
- A **discrete random variable** takes values in a **countable set**, such as $0, 1, 2, \cdots$
- A **continuous random variable** takes values in an **uncountable set**, as in the set of real numbers.
- The toss of a coin, with a value 1 assigned to heads and a value 0 assigned to tails, is an example of a discrete random variable.
- The arrival time of a customer in a bank is an example of a continuous random variable.

Definition of a Random Variable

- A random variable is a **measurable function** $X: \Omega \rightarrow E$, from a set of possible outcomes Ω to a **measure space** E.
- The set of integers $\mathcal{Z}$, with counting measure, is an example of measure space
- The set of real numbers $\mathcal{R}$, with a Lebesgue measure is another example of measure space.
- A **discrete random variable** $\boldsymbol{X}$ takes values in a countable set, as in $\mathcal{Z}$.
- A **continuous random variables** $\boldsymbol{X}$ takes values in an uncountable set, as in $\mathcal{R}$.

II. Characterization of Continuous Random Variables

- The continuous random variables are characterized by
  - Cumulative Distribution Function (CDF)
  - Probability Density Function (PDF)

Cumulative Distribution Function

- The **cumulative distribution function (CDF)** of a random variable X **(discrete or continuous)** is the function

  $F_X(x): \mathcal{R} \rightarrow [0,1]$, given by

  $F_X(x) = P(X \leq x)$, defined for both discrete and continuous random variables.

- Note that the uppercase X is used for the random variable and the lower case x is used for a specific value.

- For example, $P(X \leq 3)$ means probability that the random variable X is less than or equal to 3.

Properties of CDF

The CDF of a random variable X has the following properties:

- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to \infty} F_X(x) = 1$
- If $x_1 < x_2$, then $F_X(x_1) \leq F_X(x_2)$, that is, $F_X(x)$ is a **non-decreasing function** of x.
- $F_X(x^+) = F_X(x)$, that is, $F_X(x)$ is right continuous.
- Note that the above properties are necessary and sufficient conditions for $F_X(x)$ to be a CDF.

Additional Properties of CDF

The CDF has the following additional properties:

- $P(X > x) = 1 - P(X \leq x) = 1 - F_X(x)$
- $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$
- $P(X = x) = F_X(x) - F_X(x^-)$

Probability Density Function

- We say that X is a continuous random variable if its CDF $F_X(x) = P(X \leq x)$ can be expressed as

  $$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

  for some integrable function f: $\mathcal{R} \to [0, \infty)$.

- In the above integral, u is a **dummy variable** of integration.
- $f_X(x)$ is called the **probability density function (PDF)** of X.

Properties of PDF

The PDF of a continuous random variable X has the following properties:

- $\int_x f_X(x)dx = 1$, that is the integral of the PDF over the support of X sums up to 1.
- $P(X = x) = 0$.
- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$.
- **Notation**
  - We use uppercase $F_X(x)$ for CDF
  - We use lowercase $f_X(x)$ for PDF
  - We also use $p_X(x)$ for PDF

Difference between Discrete and Continuous Random Variables (1 of 5)

- For discrete random variables, the CDF

  $F_X(x) = \sum_{u=-\infty}^{x} f_X(u)$, that is, **it is a sum**.

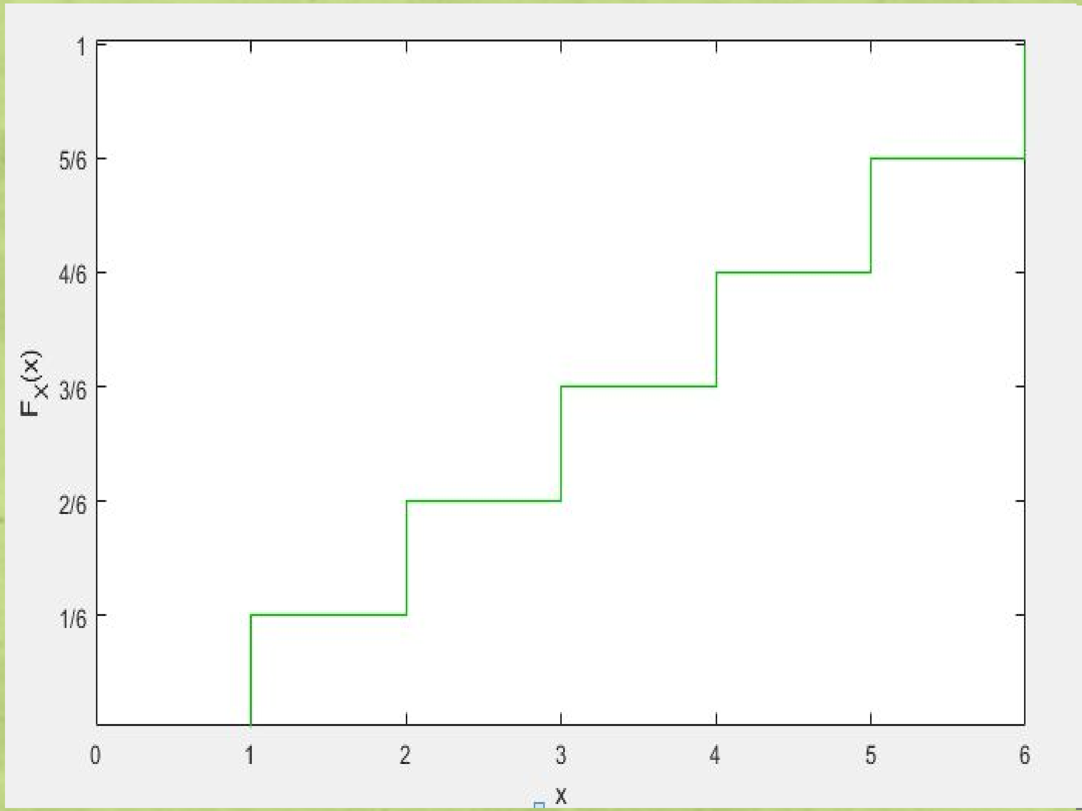
- For continuous random variables, the CDF

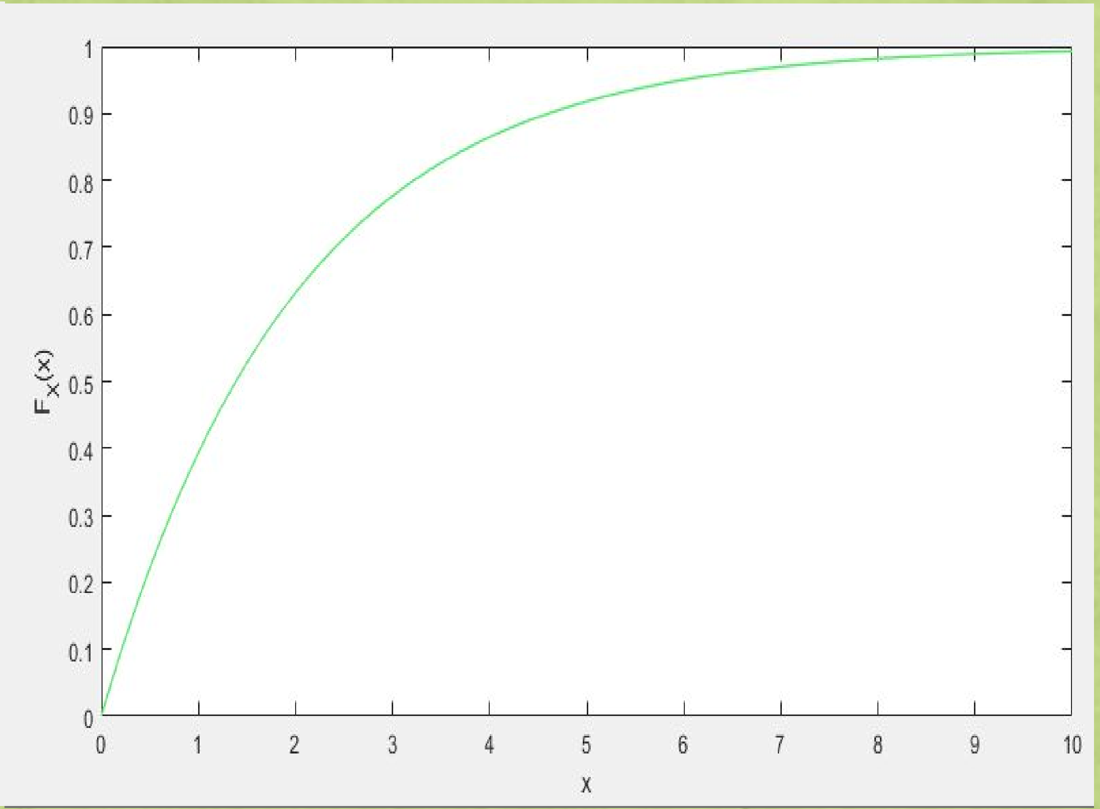
  $F_X(x) = \int_{-\infty}^{x} f_X(u)du$, that is, **it is an integral**.

- The CDF of a **discrete** random variable is a **discontinuous function** (has jumps).
- The CDF of a **continuous** random variable is a **continuous function**.

Difference between Discrete and Continuous Random Variables (2 of 5)

- The following plots show the CDF of a discrete random variable (left) and a continuous random variable (right).





Difference between Discrete and Continuous Random Variables (3 of 5)

- For discrete random variables, the PMF can be obtained from CDF as follows:

$$f_X(x) = P(X = x) = F_X(x) - F_X(x^-).$$

- For continuous random variables,

$$f_X(x) = P(X = x) = F_X(x) - F_X(x^-) = 0.$$

- That is, for a continuous random variable, the **PDF at any point *x* is zero**.
- For a continuous random variable, the PDF can be obtained from the CDF as follows:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Difference between Discrete and Continuous Random Variables (4 of 5)

- Since the PDF $f_X(x)$ for a continuous random variable is zero at any point x, we use probabilities in an interval.
- That is, the probability in a **"small interval" dx** is given by $f_X(x)dx$.
- And as indicated earlier, probability in an interval $[a, b]$ is given by the integral,

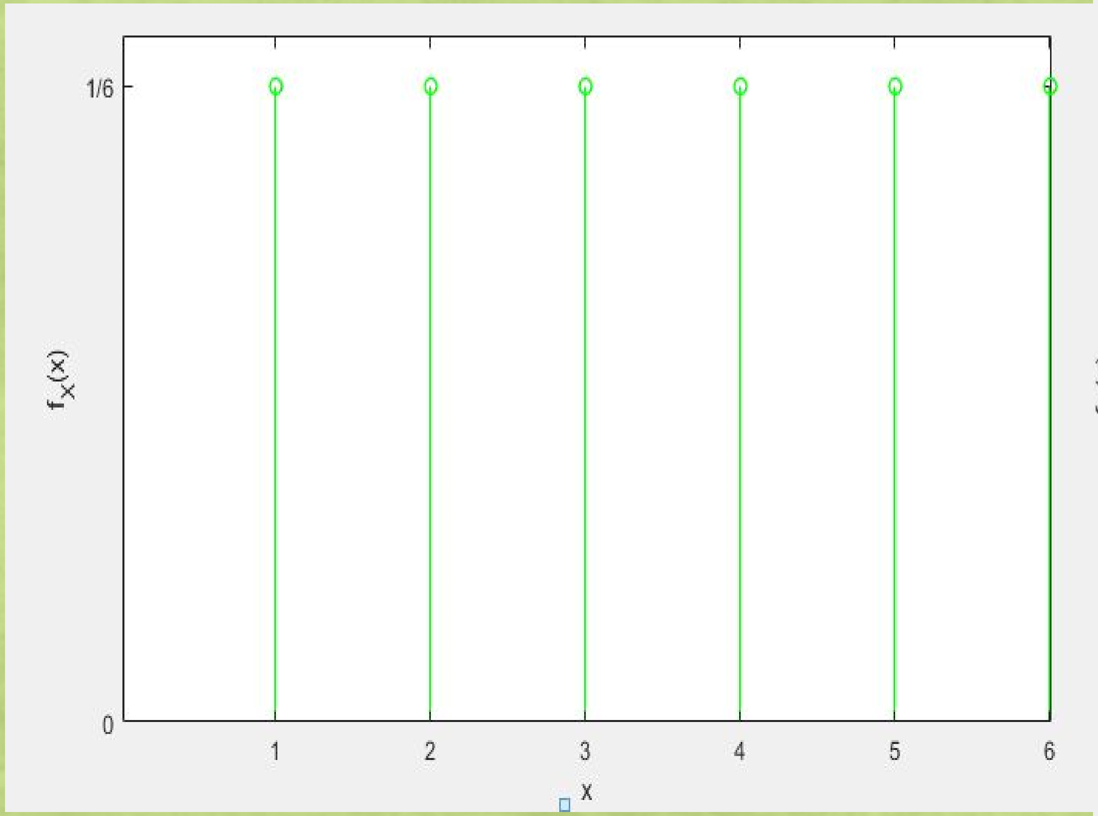
$$P(a \leq X \leq b) = \int_a^b f_X(x)dx.$$

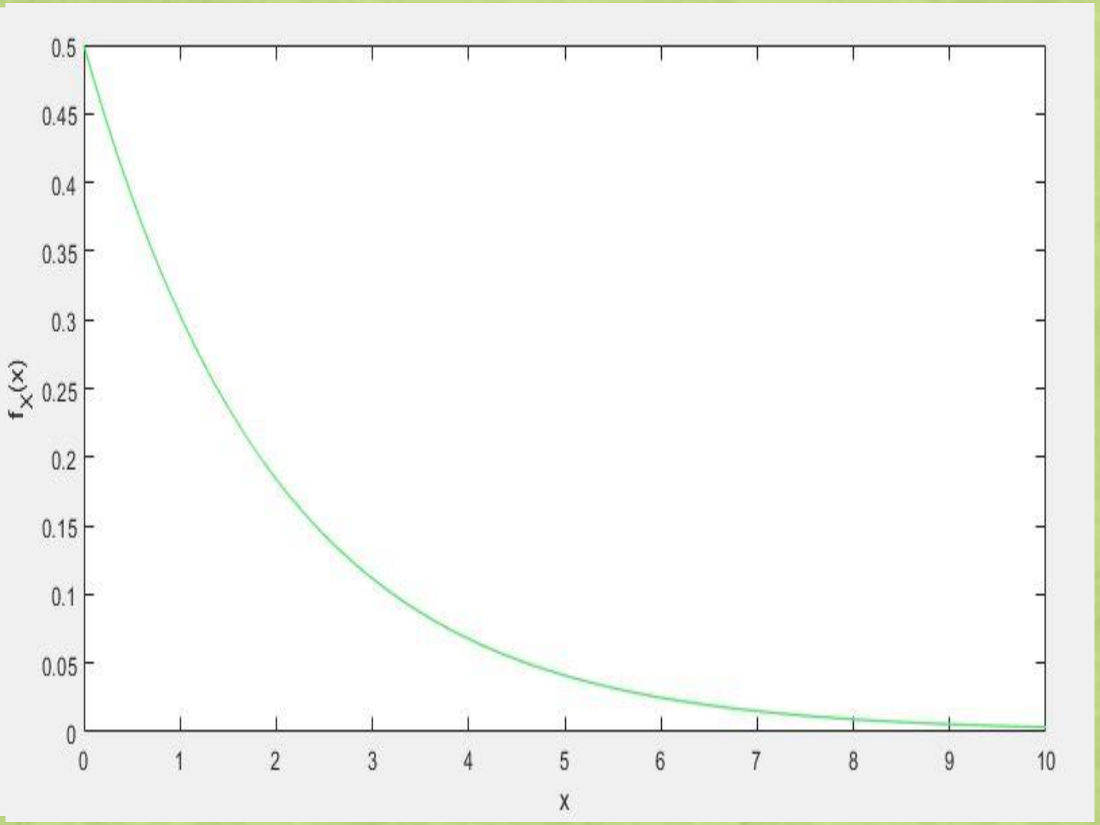
- This can also be shown to be equal to

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Difference between Discrete and Continuous Random Variables (5 of 5)

- The following plots show the PMF of a discrete random variable (left) and the PDF of a continuous random variable (right).





III. Examples of Continuous Random Variables

We will study properties of the following continuous random variables

- Uniform
- Exponential
- Normal (or Gaussian)

Uniform Random Variable (1 of 3)

Example

X is a uniform random variable, written as $\boldsymbol{\mathcal{U}(a,b)}$**, with parameters $\boldsymbol{a}$ and $\boldsymbol{b}$**. The PDF of X is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise.} \end{cases}$$

Determine the CDF $F_X(x)$.

Solution

We consider the following cases:

- $\boldsymbol{x < a}$ $F_X(x) = P(X \le x) = 0.$
- $\boldsymbol{x \ge b}$ $F_X(x) = P(X \le x) = 1.$

Uniform Random Variable (2 of 3)

- $a \le x \le b$ In this case, we determine the CDF as follows:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

$$= \int_{a}^{x} \frac{1}{b-a} du$$

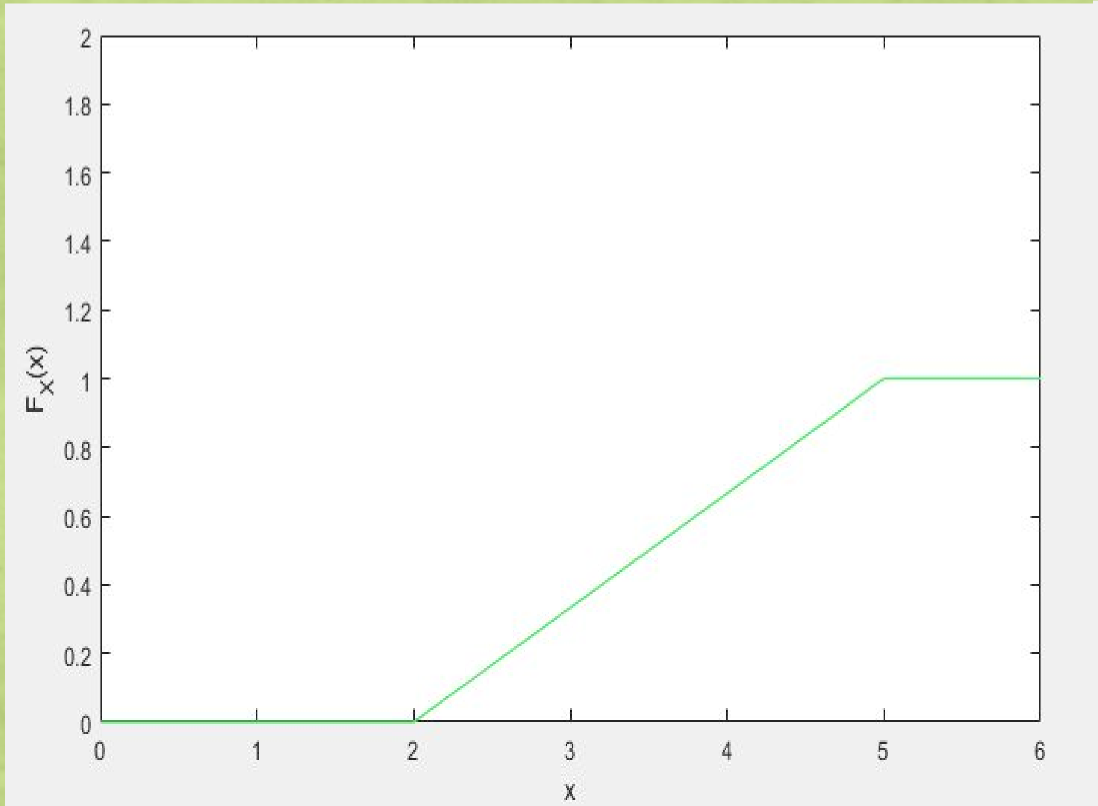
$$= \frac{1}{b-a} [u]_a^x$$

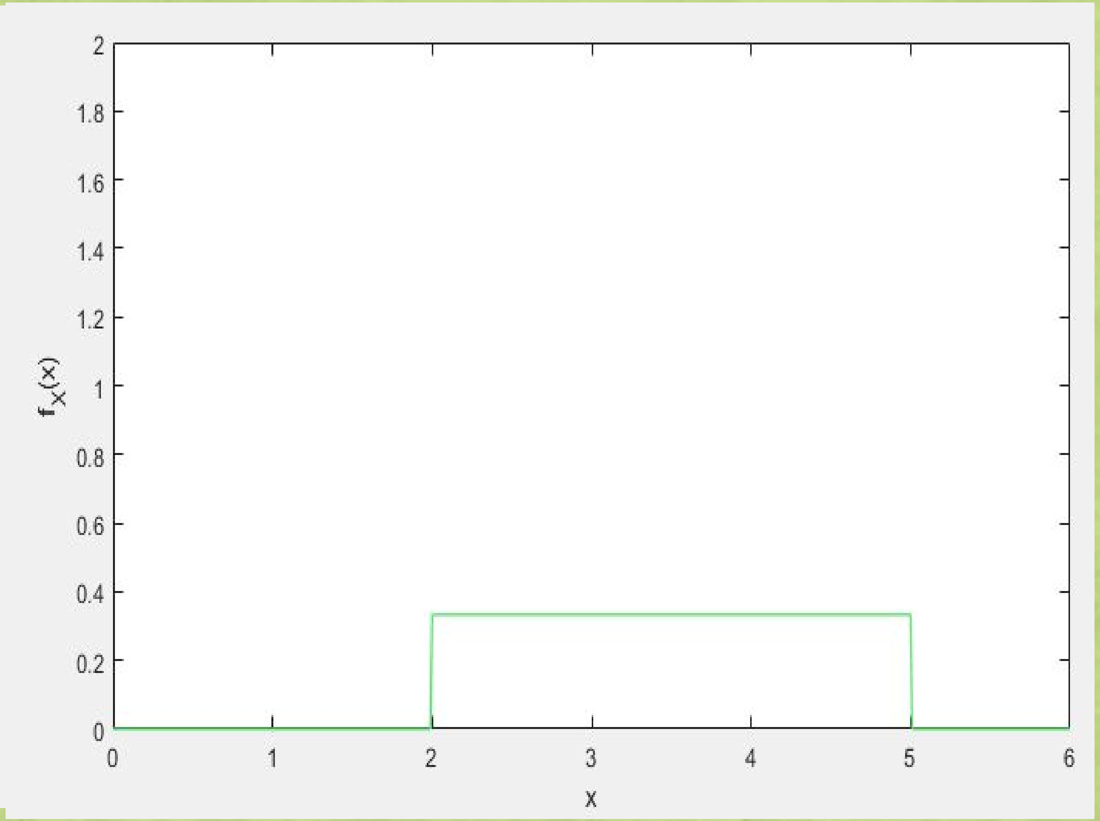
$$= \frac{1}{b-a}(x - a).$$

- Note that, for the PDF, the area under the curve satisfies $\int_a^b \frac{1}{b-a} dx = 1$.

Uniform Random Variable (3 of 3)

- The following plots show the CDF (left) and the PDF (right) of the uniform random variable $\mathcal{U}(2,5)$.





Example - Phase Angle

Example

The phase angle of a received signal is **uniformly distributed** between zero and 2π radians. Find the probability that the angle is between $\pi/4$ and $\pi/2$ radians.

Solution

Let X be the phase angle of the signal. Using $a = 0$ and $b = 2\pi$, for the parameters in the expression for the uniform PDF, we obtain,

$$f_X(x) = \begin{cases} \frac{1}{2\pi}, & 0 \le x \le 2\pi \\ 0, & \text{otherwise.} \end{cases}$$

Example - Phase Angle, cont.

The required probability is given by

$$P\left(\frac{\pi}{2} \leq X \leq \frac{\pi}{4}\right) = \int_{\pi/4}^{\pi/2} f_X(x)dx$$

$$= \int_{\pi/4}^{\pi/2} \frac{1}{2\pi} dx$$

$$= \frac{1}{8}.$$

Exponential Random Variable (1 of 3)

- The **PDF of an exponential random variable** X is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Note that it has a **single parameter** $\boldsymbol{\lambda}$, which satisfies the condition $\lambda > 0$.

- Next, we compute its CDF from the PDF.

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

$$= \int_{0}^{x} \lambda e^{-\lambda u} du.$$

- Note that after substituting for $f_X(u)$, the lower limit of the integral changed to 0. This is due to the support of X being $[0, \infty)$.

Exponential Random Variable (2 of 3)

$$= \lambda \left[\frac{e^{-\lambda u}}{-\lambda}\right]_0^x$$

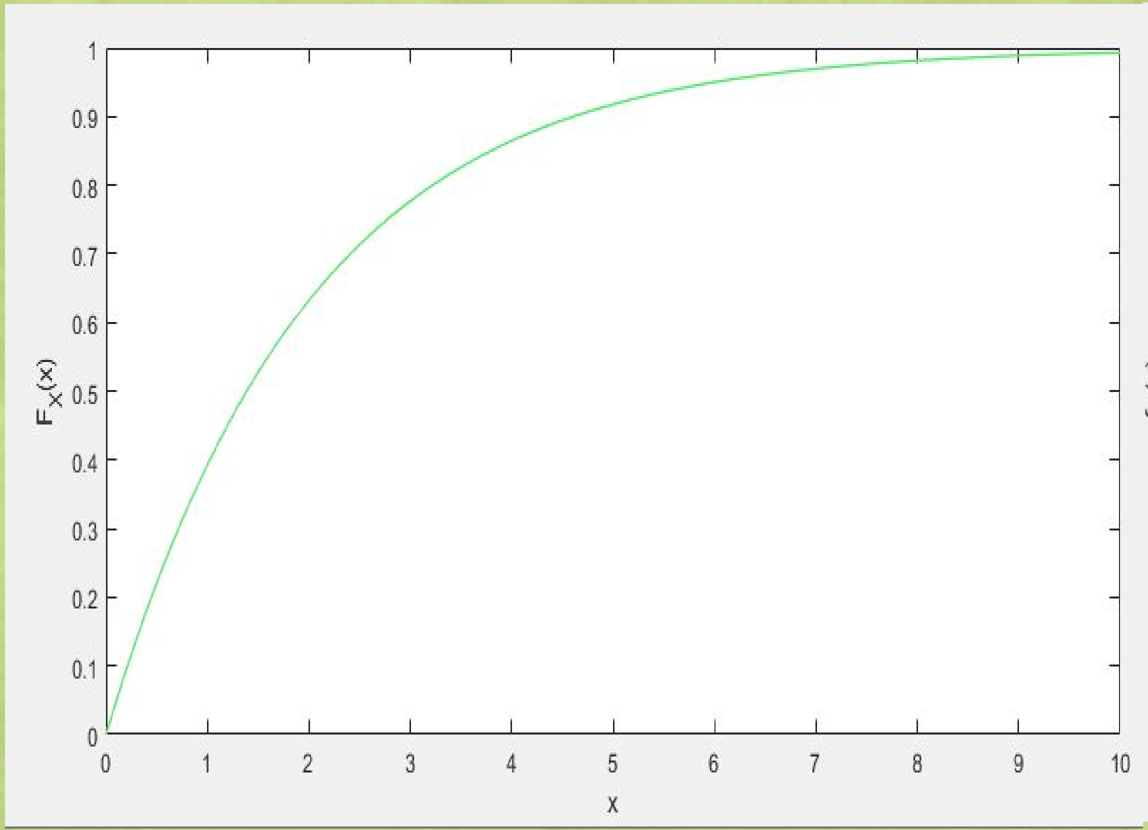
$$= (-1)\left[e^{-\lambda x} - 1\right]$$

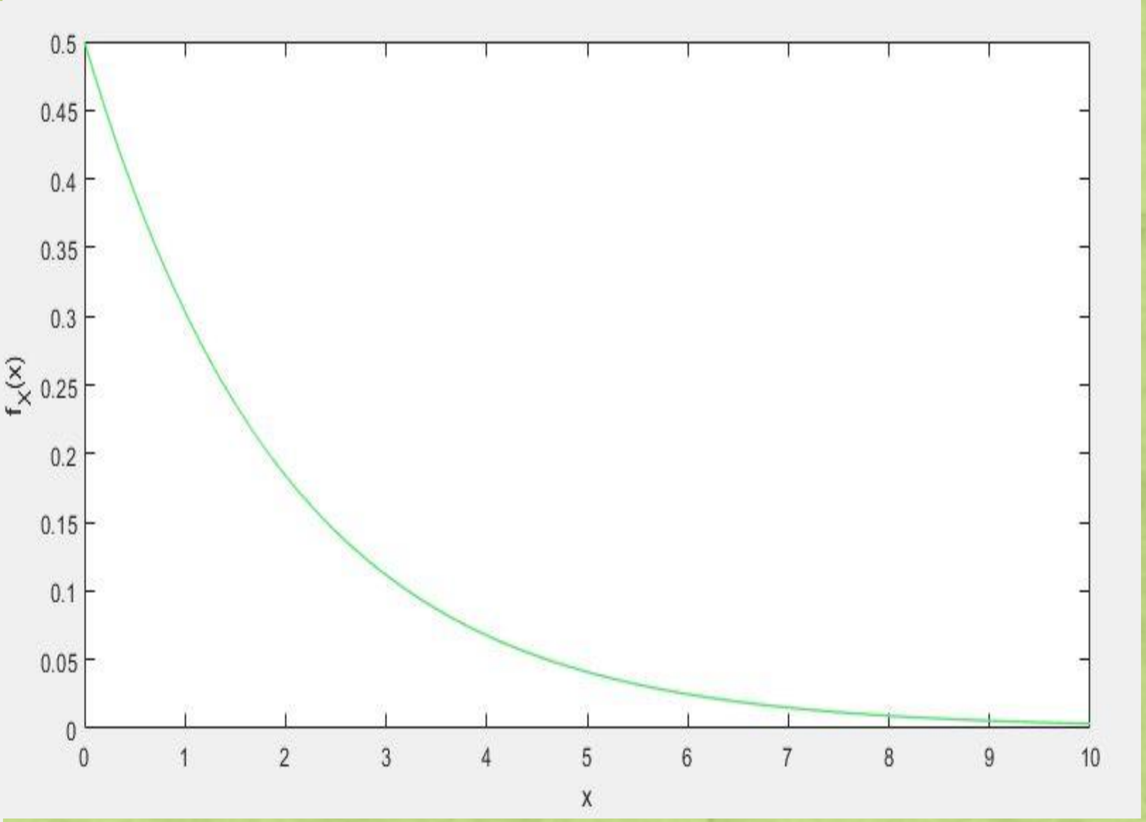
$$= 1 - e^{-\lambda x}$$

- Therefore, the CDF is $F_X(x) = 1 - e^{-\lambda x}$, $x \geq 0$.
- Note that, $F_X(x = \infty) = 1$
- **Exercise**
  - Verify that the PDF satisfies $\int_x f_X(x)dx = 1$.
  - Compute the PDF from the CDF.

Exponential Random Variable (3 of 3)

- The following plots show the CDF (left) and the PDF (right) of the exponential random variable, with parameter $\lambda = 0.5$





Example - Phone Call (1 of 3)

Example

Suppose that the length of a phone call, in minutes, is an exponential random variable with parameter

$\lambda = \dfrac{1}{10 \text{ minutes}}$.

If someone arrives immediately ahead of you in a public telephone booth, find the probability that you will have to wait

(a) more than 10 minutes.

(b) between 10 and 20 minutes.

Example - Phone Call (2 of 3)

Solution

Let X be the length of the phone call made by the person.

(a) We want to compute $P(X > 10)$

We have the CDF of exponential random variable,

$$F_X(x) = P(X \leq x) = 1 - e^{-\lambda x}$$

Therefore,

$$P(X > x) = 1 - P(X \leq x) = e^{-\lambda x}$$

$$P(X > 10) = e^{-\left(\frac{1}{10}\right)10}$$

$$= e^{-1} = 0.368$$

Example - Phone Call (3 of 3)

(b) We want to compute $P(10 \leq X \leq 20)$

$$P(10 \leq X \leq 20) = F_X(20) \text{-} F_X(10)$$

$$= \left(1 - e^{-\left(\frac{1}{10}\right)20}\right) - \left(1 - e^{-\left(\frac{1}{10}\right)10}\right)$$

$$= -e^{-2} + e^{-1} = 0.233$$

Normal Distribution

- The **normal distribution** is the most widely used distribution in many practical scenarios.
- Some examples of data which follows this pattern are:
  - Heights of people
  - Salaries of people
  - Test Scores
- This distribution is also known as **Gaussian distribution** or the **bell curve**.

Normal Random Variable - PDF

- The **PDF of a normal random variable** X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

- It has **two parameters,** μ **and** $\boldsymbol{\sigma^2}$, which are the **mean and the variance of** $\boldsymbol{X}$.
- We write this as $X \sim \mathcal{N}(\mu, \sigma^2)$.
- If **μ=0 and** $\boldsymbol{\sigma^2 = 1}$, we call it the **standard normal distribution**.

Normal Random Variable - PDF, cont.

- Let the variable Z have a standard normal distribution, that is, $Z \sim \mathcal{N}(0,1)$. Its PDF is given by

  $$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty.$$

- Using the transformation of variables (done in a later lecture), it can be shown that when

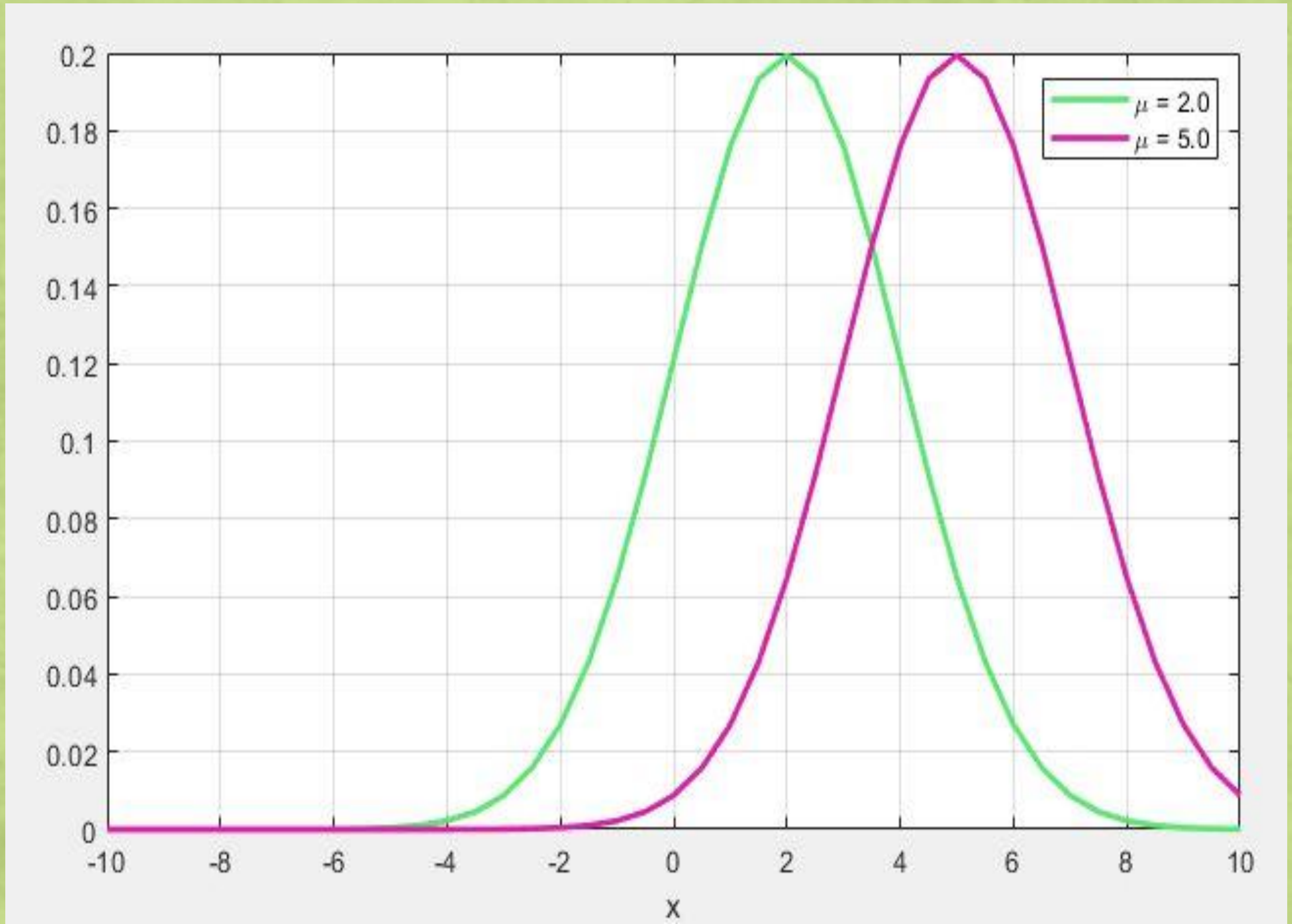
  $X \sim \mathcal{N}(\mu, \sigma^2)$, then

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

  has a standard normal distribution $\mathcal{N}(0,1)$.

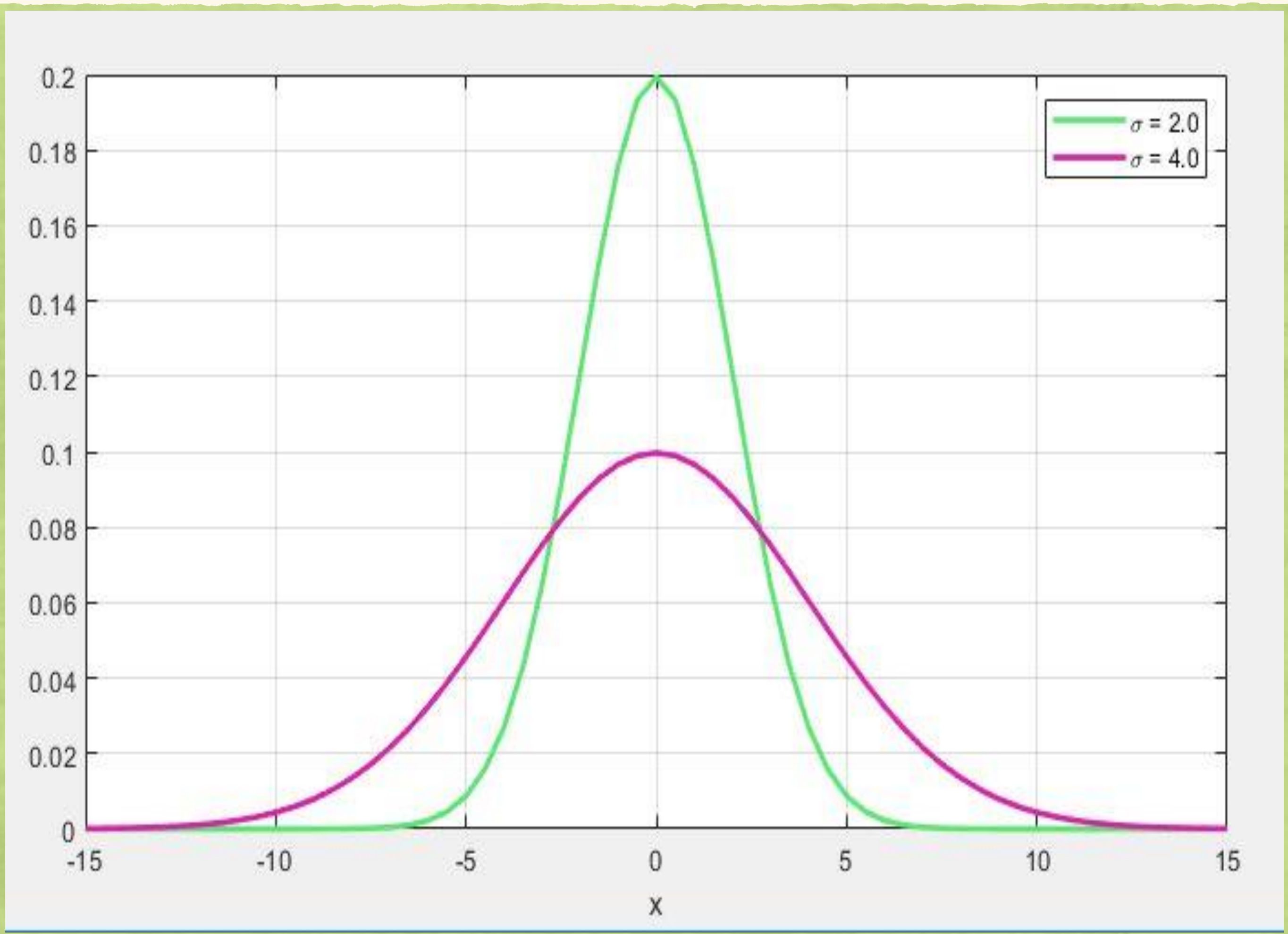
- This relationship will be utilized in computing the probabilities for normal distribution.

Normal Distribution Plots - PDF



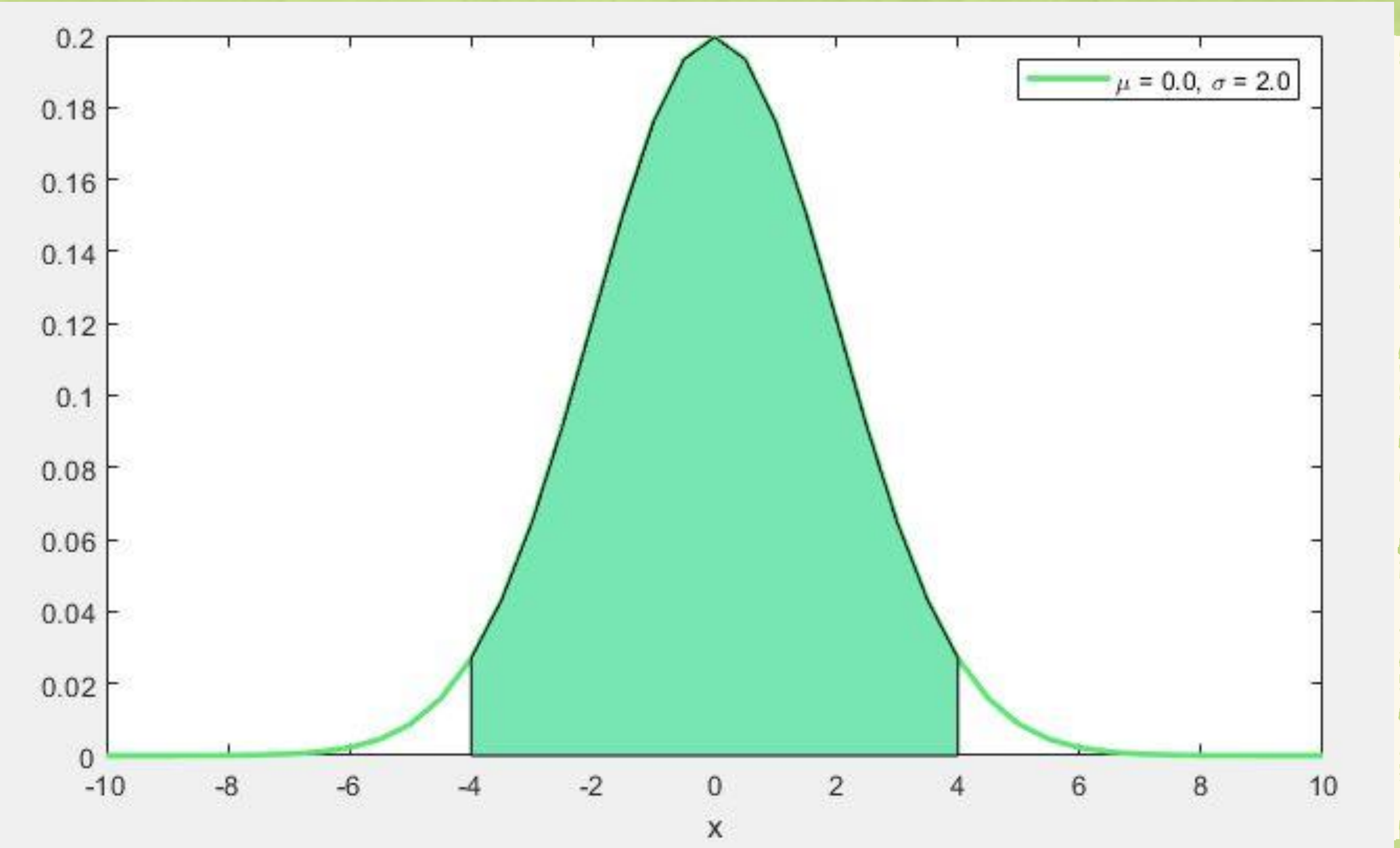
This plot shows the PDF of a normal distribution with two different values of mean μ and same variance σ^2.

Normal Distribution Plots - PDF, cont.



This plot shows the PDF of a normal distribution with the same mean μ and two different values of variance σ^2.

Normal Distribution - Area Under the Curve



- This plot shows the normal distribution with parameters $\mu=0.0$ and $\sigma=2.0$
- The **area under the curve** represents the probabilities.
- Approximately 95% of the area of the normal distribution is within two standard deviations of its mean. In this case, this is between -4.0 and +4.0
- This means that 95% of the data with this specific distribution will lie in the range -4.0 to +4.0

Computing Probabilities with Normal Distribution (1 of 3)

- The CDF of the normal distribution $\mathcal{N}(\mu, \sigma^2)$ is given by

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du\,, \quad -\infty < x < \infty.$$

- In the above equation, ***u* is a dummy variable of integration**. This integral cannot be evaluated analytically and must be computed numerically.

- In order to do this, we first transform the normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ to a standard normal distribution $\mathrm{Z} \sim \mathcal{N}(0,1)$, using the following transformation

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

Computing Probabilities with Normal Distribution (2 of 3)

- We use a special notation **$\Phi(z)$ for the CDF of the standard normal distribution**, and it is given by

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du\,, \quad -\infty < z < \infty.$$

- We utilize the tables for $\Phi(z)$ to compute the probabilities for normal distribution (given in Appendix A).

- With the above transformation between Z and X, we obtain

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(\sigma Z + \mu \leq x) \end{aligned}$$

Computing Probabilities with Normal Distribution (3 of 3)

$$F_X(x) = P\left(Z \leq \frac{x-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{x-\mu}{\sigma}\right)$$

- Note that the table only gives $\Phi(z)$ for positive values of z.
- We can obtain $\Phi(z)$ for negative values of z, from the following expression

  $\Phi(z) + \Phi(-z) = 1.$
- The proof of this expression will be done in the assignment.

Example - Computing Probability, Normal Distribution (1 of 4)

Example

X is a normal random variable, $\mathcal{N}(\mu, \sigma^2)$, with parameters $\mu = 3$ and $\sigma^2 = 4$. Determine the following probabilities: (a) $P(X > 3)$ (b) $P(2 \leq X \leq 3)$

Solution

(a) We make use of the **transformation** from normal random variable X to a standard normal random variable Z, as follows:

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

Therefore,

$$P(X > 3) = P(\sigma Z + \mu > 3)$$

$$= P\left(Z > \frac{3-\mu}{\sigma}\right)$$

Example - Computing Probability, Normal Distribution (2 of 4)

That is,

$$P(X > 3) = P\left(Z > \frac{3-\mu}{\sigma}\right)$$

$$= 1 - \Phi\left(\frac{3-\mu}{\sigma}\right)$$

$$= 1 - \Phi\left(\frac{3-3}{2}\right), \quad \text{substituting for } \mu = 3 \text{ and } \sigma = 2$$

$$= 1 - \Phi(0),$$

$= 0.5,$ **utilizing the table** in Appendix A for $\Phi(z)$.

Example - Computing Probability, Normal Distribution (3 of 4)

(b) $P(2 \leq X \leq 3) = F_X(3) - F_X(2)$,

utilizing the CDF property $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$.

Therefore,

$$P(2 \leq X \leq 3) = \Phi\left(\frac{3-\mu}{\sigma}\right) - \Phi\left(\frac{2-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{3-3}{2}\right) - \Phi\left(\frac{2-3}{2}\right),$$ substituting for μ and σ

$$= \Phi(0) - \Phi(-0.5).$$

Example - Computing Probability, Normal Distribution (4 of 4)

From the table of $\Phi(z)$,

$\Phi(0) = 0.5$ and $\Phi(0.5) = 0.6915$

Also, $\Phi(-0.5) = 1 - \Phi(0.5)$

$$= 1 - 0.6915$$

$$= 0.3085$$

Therefore, $P(2 \leq X \leq 3) = \Phi(0) - \Phi(-0.5)$

$$= 0.5 - 0.3085$$

$$= 0.1915$$

Appendix A: Table of Φ(·) Function

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

References

1. Charles Boncelet, Probability, Statistics and Random Signals, Oxford University Press, 2016.
2. Sheldon Ross, A First Course in Probability, Macmillan Publishing Company, 1988.
3. R. D. Yates, et al., Probability and Stochastic Processes, John Wiley, 2005.