Statistics and Probabilities for Computing (CEN1602)

Unit-4

Continuous Random Variables and Probability Distributions

Dr. Ahmad Alzahrani

Outline

1. Continuous random variables and how they differ from discrete random variables
2. Probability Density Functions (PDFs) and probability as area under a curve
3. Cumulative Distribution Functions (CDFs) and PDF-CDF relationship
4. Mean, variance, expected value, and standard deviation of continuous random variables.
5. Continuous uniform distribution and worked examples
6. Normal distribution, empirical rule, standardization, and table use
7. Normal approximation to the binomial distribution
8. Exponential distribution
9. Erlang / Gamma distributions

Learning Objectives for Unit 4

By the end of this unit, students should be able to:

1. **Differentiate** between discrete and continuous random variables
2. **Define** and **interpret** probability density functions (PDFs)
3. **Compute** probabilities using definite integrals over a PDF
4. **Construct** and **use** cumulative distribution functions (CDFs)
5. **Calculate** mean and variance for continuous random variables
6. **Apply** continuous distributions:
   - Uniform
   - Normal
   - Exponential
   - Erlang / Gamma
7. **Standardize** normal variables and **compute** probabilities using Z-tables
8. **Apply** normal approximation to binomial distribution
9. **Use** computational tools (Python libraries) to **evaluate** and **visualize** distributions (practical side)

From Unit 3 to Unit 4

	Unit 3: **Discrete** Random Variables	Unit 4: **Continuous** Random Variables
Values	Countable values	Any value in an interval
Examples	Number of defects, number of arrivals, number of successes	Time, temperature, voltage, latency
Function	**PMF**, $p(x)$	**PDF**, $f(x)$
Probability at one point	$P(X = x)$ can be positive	$P(X = x) = 0$
Probability over a range	**Sum** of point probabilities	**Area** under the density curve
How to compute	Summation	Integration
Covered distributions	Uniform, Binomial, Poisson, Geometric, Negative Binomial, & Hypergeometric	Uniform, Normal, Exponential, Erlang/Gamma
Main interpretation	Counting outcomes	Measuring outcomes

In short:

Discrete: $P(X = x)$ **Continuous**: $P(a < X < b)$

Continuous Random Variables

- A **continuous random variable** X can take any value in an interval of the real line.
- Its values are not countable one by one as in the discrete case.
- Examples include: server response time, network latency, CPU temperature, and voltage.
- For example, latency may be 8.2ms, 8.25ms, or 8.253ms.
- For a continuous random variable, the probability of one exact value is zero:

$$P(X = x) = 0$$

- This does not mean the value is impossible; it means probability is assigned to intervals.
- So, we calculate probabilities such as $P(10 < X < 12)$ rather than $P(X = 10)$
- In this unit, such probabilities will be found using a probability density function (PDF).

Probability Density Function (PDF)

- A **probability density function (PDF)** describes how probability is distributed over a continuous random variable X.
- It is denoted by $f(x)$.
- For a continuous random variable, probabilities are found from areas under the curve of $f(x)$.
- The value $f(x)$ itself is **not** a probability; it is a **density value**.
- If $f(3) = 10$ and $f(8) = 2$,then probability is denser near 3 than near 8. Therefore, equal small intervals around 3 contain more probability than equal small intervals around 8.
- The probability that X lies between a and b is:

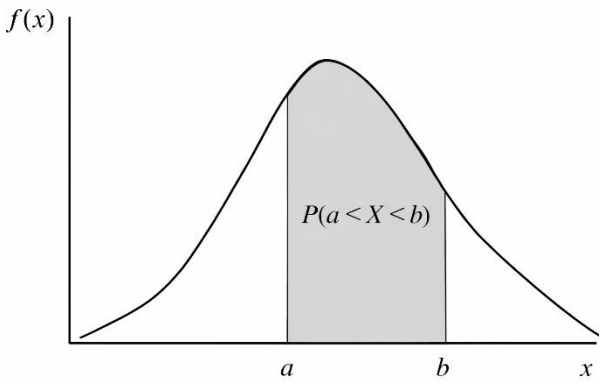
$$P(a < X < b) = \int_a^b f(x)\, dx$$

- A PDF must satisfy the condition

$$f(x) \geq 0 \ \textit{for all } x$$

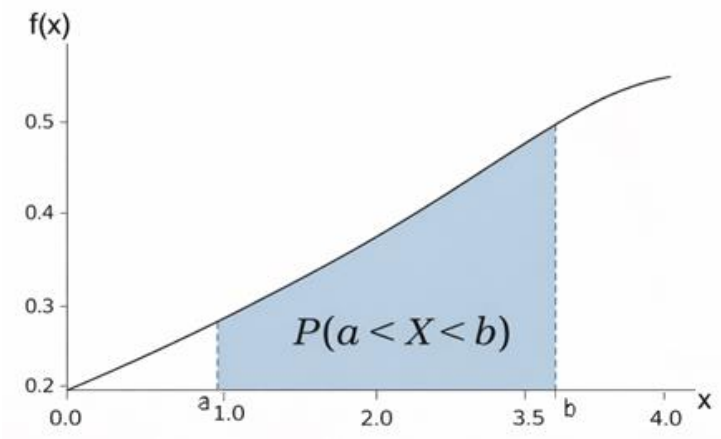
- The total area under the entire density curve = 1:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$



Interpreting Probability as Area

- For continuous random variables, probability is the area under the density curve over an interval.
- The wider the interval or the larger the density over that interval, the larger the probability.



- Because a single point has zero width, its area is zero; therefore, for a continuous random variable X:

$$P(X = a) = 0$$

- As a result, for continuous random variables, there is no difference between $<$ and $\leq$ in probability calculations.
- Therefore, $P(a < X < b) = P(a \leq X \leq b)$

Cumulative Distribution Function (CDF)

- For a continuous random variable X, the **cumulative distribution function (CDF)** gives the probability that X is less than or equal to a given value x.
- The CDF is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f\,(t)\,dt$$

- The CDF accumulates probability from the far left up to x.
- Since probability is area under the PDF, $F(x)$ represents the area under $f(t)$ from $-\infty$ to x.
- The CDF is always between 0 and 1:

$$0 \leq F(x) \leq 1$$

- For a continuous random variable, probabilities over intervals can be found using the CDF:

$$P(a < X \leq b) = F(b) - F(a)$$

- If $F(x)$ is differentiable, then the PDF is obtained from the CDF by

$$f(x) = \frac{dF(x)}{dx}$$

Example: Meaning of the CDF and PDF

Let X be the waiting time (in minutes) for a student shuttle bus. Assume the bus arrives at a random time between 0 and 10 minutes. Then X has a continuous uniform distribution on $[0,10]$.

$$f(x) = \begin{cases} 0.1, & 0 \le x \le 10 \\ 0, & \textit{otherwise} \end{cases}$$

- The PDF is constant because all waiting times between 0 and 10 minutes are equally likely. The probability of waiting between $a$and $b$minutes is the area under the PDF:

$$P(a \le X \le b) = \int_a^b 0.1 \; dx = \frac{b-a}{10}$$

- The CDF gives the probability that the waiting time is at most x minutes:

$$F(x) = P(X \le x) = \int_0^x 0.1 \, dt = \frac{x}{10}$$

Example: $F(4) = P(X \le 4) = \frac{4}{10} = 0.4$

- So, the probability that the bus arrives in 4 minutes or less is 0.4.
- Also,

$$f(x) = F'(x) = 0.1 \qquad \textit{for } 0 \le x \le 10$$

Mean and Variance of a Continuous Random Variable

- For a continuous random variable X with probability density function $f(x)$,the **mean** (or expected value) is the average value of X over many repeated observations.
- The mean is denoted by μ or $E(X)$,and is computed by

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

- The **variance**, denoted by $\mathrm{Var}(X)$ or σ^2, measures how far the values of X tend to spread around the mean, and is computed by

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

- or

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$$

where,

$$E(X^2) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx$$

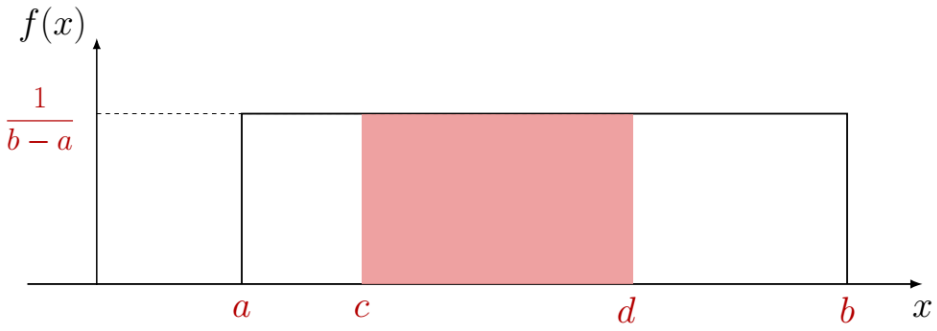
- The **standard deviation** is the square root of the variance:

$$\sigma = \sqrt{\mathrm{Var}(X)}$$

Continuous Uniform Distribution

- The continuous uniform distribution is the simplest continuous distribution.
- It models a quantity that is equally likely to occur anywhere in an interval $[a, b]$.
- This means the density is constant between a and b, and zero outside that interval.
- The probability density function is

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



- Since the density is constant, the probability over any interval depends only on its length.
- For $a \le c \le d \le b$,

$$P(c \le X \le d) = \int_{c}^{d} f(x)\ dx = \frac{d-c}{b-a}$$

- Mean and variance for a continuous uniform distribution are:

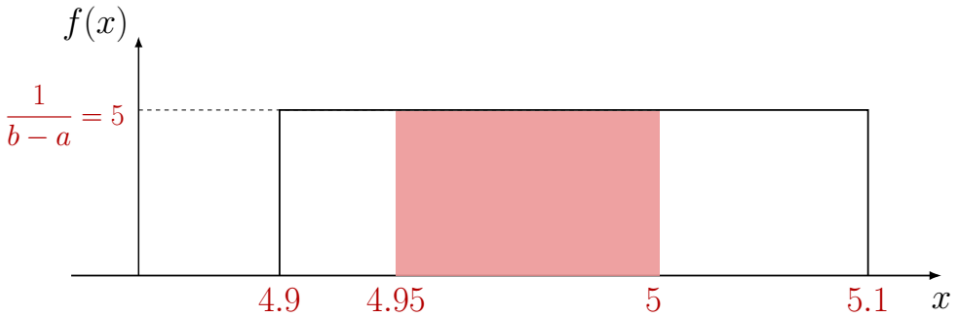
$$\mu = E(X) = \frac{a+b}{2}, \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

Continuous Uniform Distribution (Example)

Suppose X is the electrical current (mA) in a thin copper wire, with range

$$4.9 \leq X \leq 5.1$$

and thus, constant density $f(x) = 5$, Find $P(X < 5)$ and $P(4.95 < X < 5.10)$.



Solution:

$$P(X < 5) = \int_{4.9}^{5} 5 \; dx = 5(5 - 4.9) = 5(0.1) = 0.5$$

Similarly,

$$P(4.95 < X < 5.10) = \int_{4.95}^{5.10} 5 \; dx = 5(5.10 - 4.95) = 5(0.15) = 0.75$$

Since the density is constant, we can observe that probability depends only on the interval length.

Example: Random Number Generated Between 0 and 4

Suppose a program generates a random real number uniformly between 0 and 4. Let X denote the generated value. Then, $X \sim \mathit{Uniform}(0,4)$*

- The PDF is

$$f(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 4 \\ 0, & \mathit{otherwise} \end{cases}$$

- Probability that the generated number is between 0 and 1:

$$P(0 \le X \le 1) = \frac{1}{4} = 0.25$$

- Probability that the generated number is greater than 2.5:

$$P(2.5 < X \le 4) = \frac{4 - 2.5}{4} = \frac{1.5}{4} = 0.375$$

- Expected value, $E(X) = \frac{0+4}{2} = 2$
- Variance, $\mathrm{Var}(X) = \frac{(4-0)^2}{12} = \frac{16}{12} \approx 1.33$
- Standard deviation:

$$\sigma = \sqrt{1.33} \approx 1.15$$

$* \mathit{Note}$: $X \sim \mathit{Uniform}(0,4)$ means the random variable X follows a uniform distribution on the interval $[0,4]$

Normal Distribution

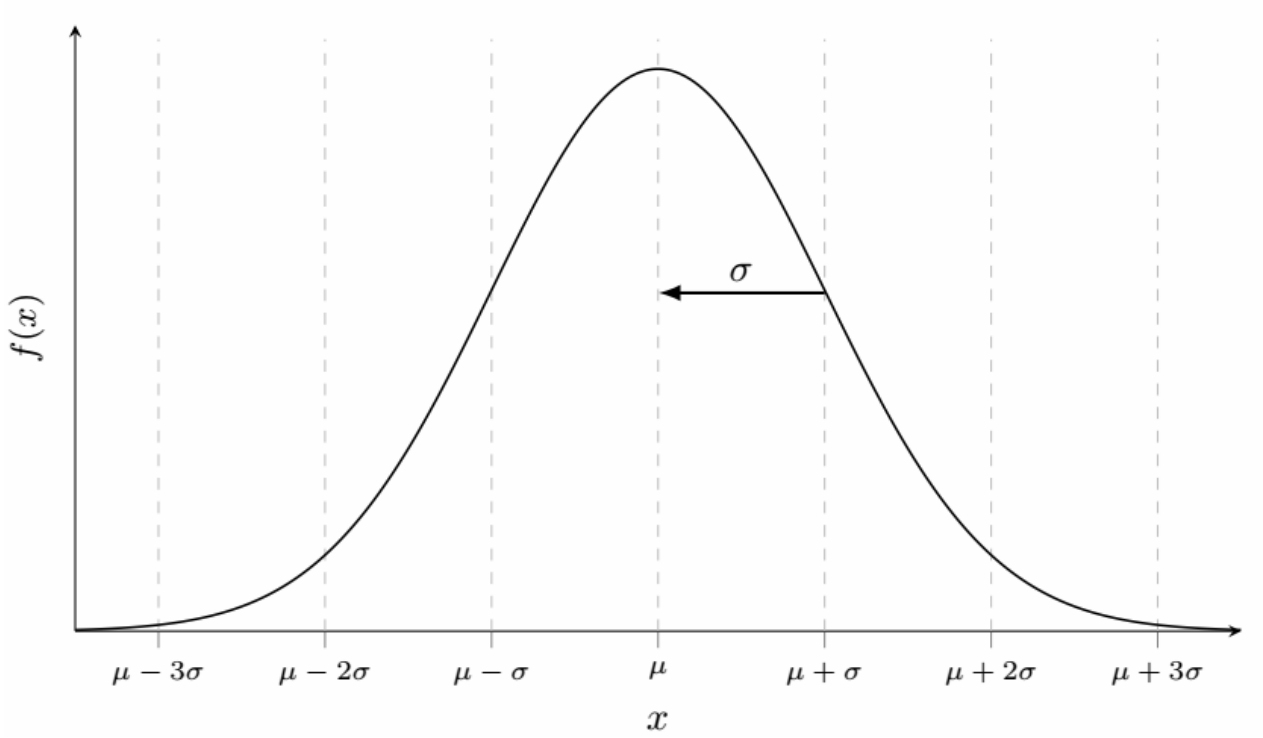
- The **normal distribution** is one of the most important continuous probability distributions.
- It is used to model many real-world quantities such as measurement error, test scores, response times, and natural variation in data.
- The normal distribution is also called the **Gaussian distribution**.
- A continuous random variable X is said to follow a normal distribution with mean μ and variance σ^2 if its PDF is

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

- We write this as $X \sim N(\mu, \sigma^2)$
- The parameters are:
  - μ: the mean, which determines the center of the distribution
  - σ: the standard deviation, which determines the spread
- For a normal random variable,

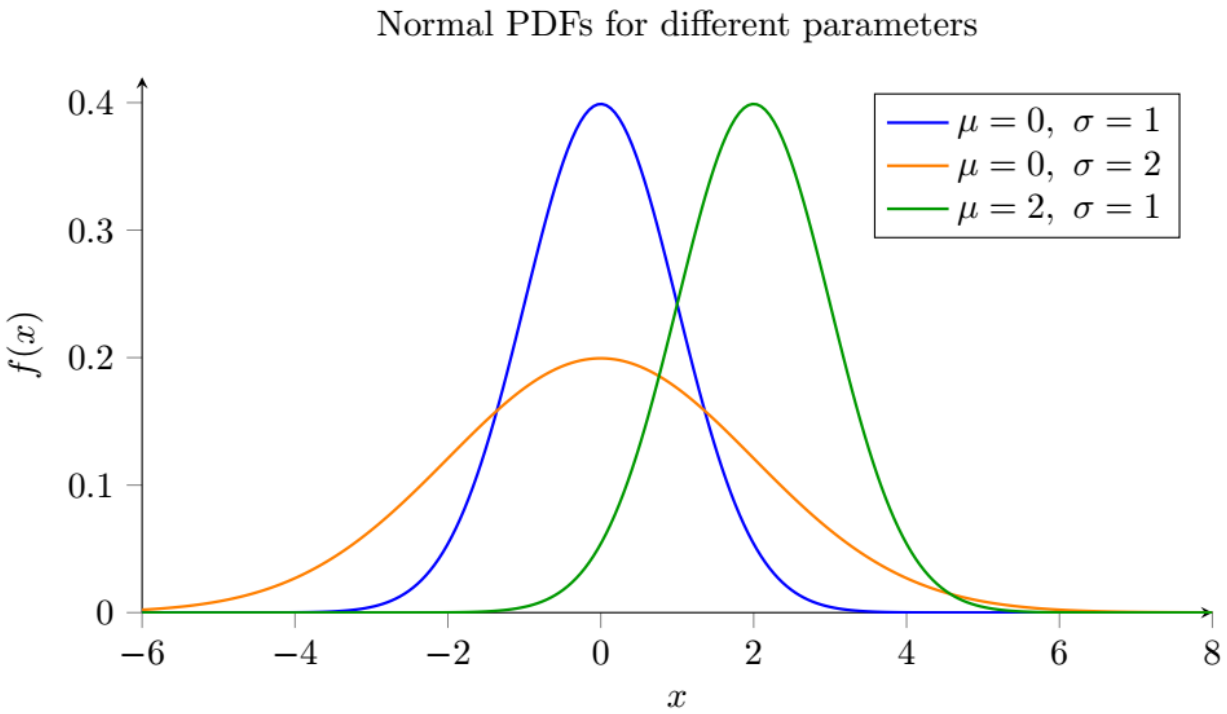
$$E(X) = \mu, \qquad \mathrm{Var}(X) = \sigma^2$$

Normal Distribution Curve



- The normal distribution curve is centered at μ
- The left and right sides are mirror images (symmetric about μ)
- Mean, median, and mode are all equal to μ
- Spread is determined by the standard deviation σ
- Larger σ gives a wider & flatter curve
- Total area under the curve is 1

Normal Distribution Curve



- The **blue** and **orange** curves have the same center because both have $\mu = 0$.
- The **orange** curve is wider and flatter because it has the larger standard deviation $\sigma = 2$.
- The **blue** and **green** curves have the same spread because both have $\sigma = 1$.
- The **green** curve is shifted to the right because its mean is larger.

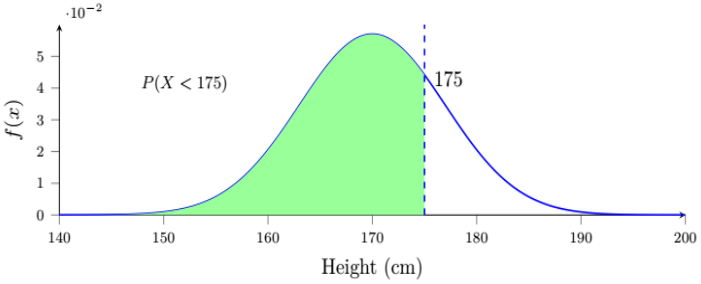
Example: Probabilities for a Normal Distribution

Height is often normally distributed in a population. Based on the national Saudi growth chart study, we use 170 cm as the mean height of Saudi adult males. Using 7 cm as the standard deviation, $X \sim N(170, 7^2)$

So, the PDF for height distribution is $f(x) = \frac{1}{7\sqrt{2\pi}} e^{-\frac{(x-170)^2}{98}}$

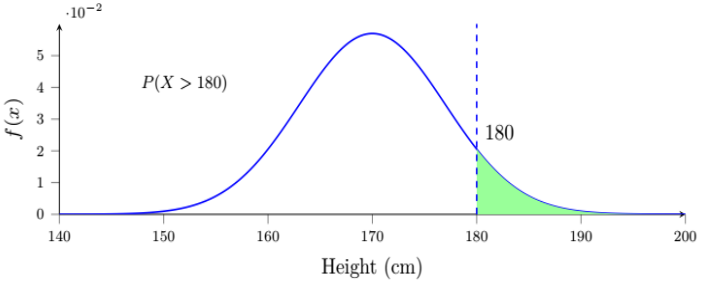
- What is the probability that a male adult is **shorter than 175 cm**?

$$P(X < 175) = \int_{-\infty}^{175} f(x)\, dx \approx 0.7625$$



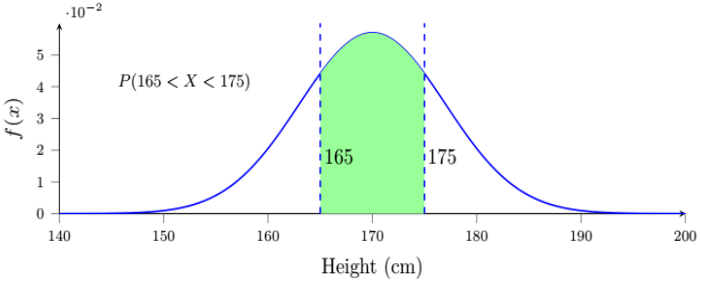
- What is the probability that a male adult is **taller than 180 cm**?

$$P(X > 180) = \int_{180}^{\infty} f(x)\, dx \approx 0.0766$$



- What is the probability that a male height is **between 165 cm and 175 cm**?

$$P(165 < X < 175) = \int_{165}^{175} f(x)\, dx \approx 0.5249$$

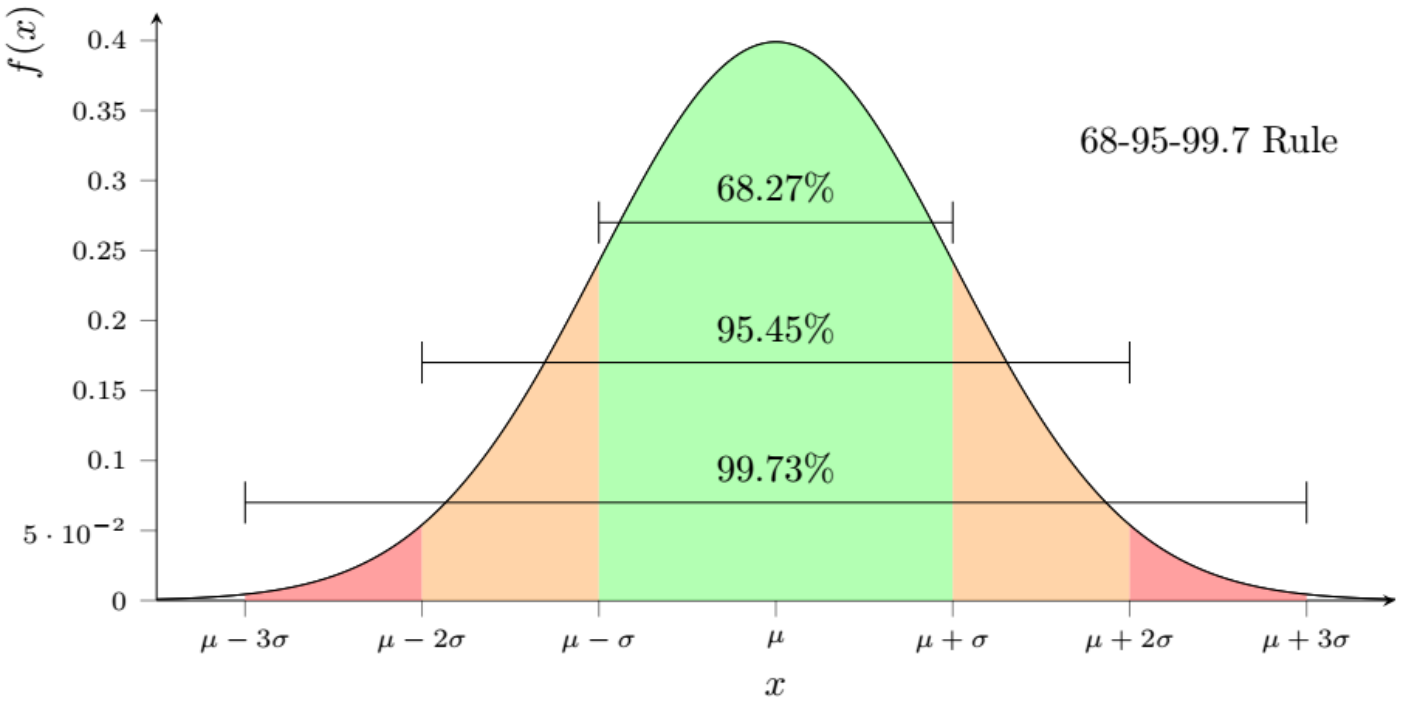


Empirical Rule for the Normal Distribution

- For a normal random variable X , most observations lie close to the mean μ.
- The **Empirical Rule** gives approximate probabilities **within** 1, 2, and 3 standard deviations of the mean as follows:

$$P(\mu - \sigma < X < \mu + \sigma) \approx \mathbf{0.6827}$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx \mathbf{0.9545}$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) \approx \mathbf{0.9973}$$

- In percentage form, this is often called the **68–95–99.7 rule**.



Empirical Rule: Examples

Example 1: Fasting Blood Glucose

Fasting blood glucose in a healthy adult reference group is approximately normally distributed with $\mu = 90$ mg/dL , and $\sigma = 10$ mg/dL

- Using the Empirical Rule:
  - About **68%** of healthy measurements lie between 80 and 100 mg/dL
  - About **95%** lie between 70 and 110 mg/dL
  - About **99.7%** lie between 60 and 120 mg/dL
- A result around **125 mg/dL** is unusually high relative to this healthy reference distribution and may need a follow-up.

Example 2: CPU Power Consumption

The electrical power used by a processor during a standard benchmark is approximately normally distributed with $\mu = 65$ W, and $\sigma = 4$ W

- Using the Empirical Rule:
  - About **68%** of runs use between 61 W and 69 W
  - About **95%** of runs use between 57 W and 73 W
  - About **99.7%** of runs use between 53 W and 77 W
- If repeated runs show **79 W**, that is outside the usual 3σ range and may suggest overheating, voltage instability, or hardware fault.

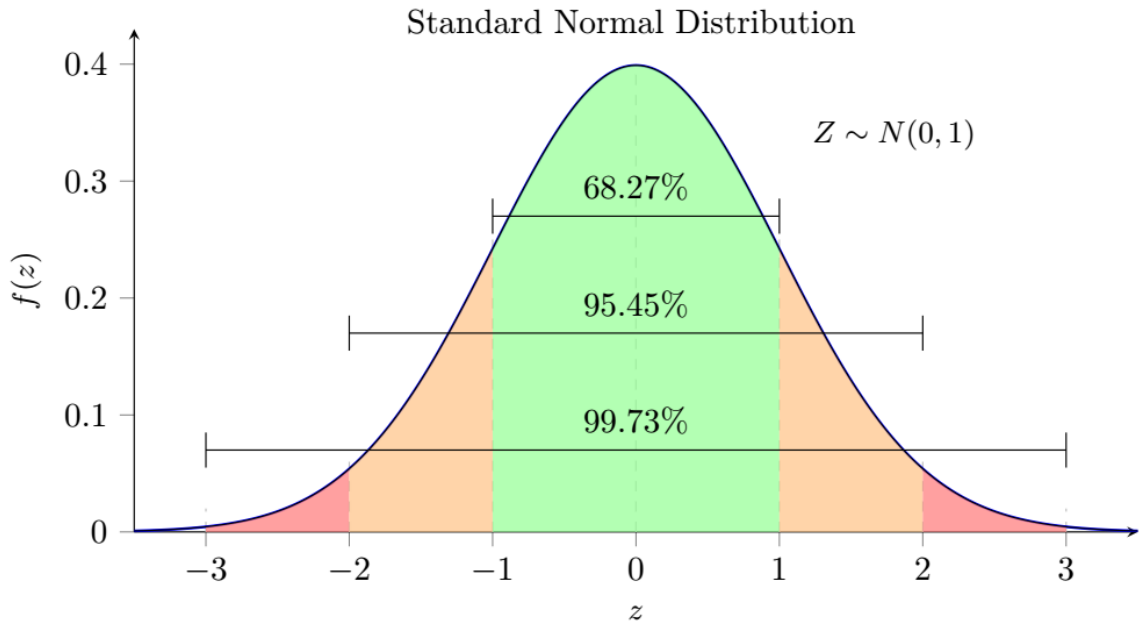
Standard Normal Random Variable

- A **standard normal random variable** is a normal random variable with $\mu = 0$ and $\sigma = 1$.
- It is usually denoted by $Z \sim N(0,1)$
- Its probability density function is

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \qquad -\infty < z < \infty$$

- It is used as a common reference form for all normal distributions.
- Probabilities for Z are written using the cumulative distribution function:

$$\Phi(z) = P(Z \leq z)$$



Standardizing a Normal Random Variable

- Any normal random variable X can be converted to a standard normal random variable Z using:

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

- This process is called **standardization**.
- It tells us how many standard deviations a value x is from the mean μ.
  - If $z = 0$, then $x = \mu$
  - If $z = 1$, then x is one standard deviation above the mean
  - If $z = -2$, then x is two standard deviations below the mean
- So probabilities for X can be converted into probabilities for Z.

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

$$P(a < X < b) = P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right)$$

- Standardization changes the scale, but not the underlying probability.

Example: Using the Standard Normal Table

From a previous example, Saudi adult male height is approximately normally distributed with mean $\mu = 170$cm and standard deviation $\sigma = 7$cm. Find:

1. The probability that a randomly selected male is shorter than 175 cm.

- Step 1: Standardize $z = \frac{x-\mu}{\sigma} = \frac{175-170}{7} = 0.71$
- Step 2: Use the standard normal table

$$P(X < 175) = P(Z < 0.71)$$

From the standard normal table,

$$P(Z < 0.71) = 0.76115$$
$$P(X < 175) \approx 0.76115$$

2. The probability that a randomly selected male is taller than 180 cm.

- Step 1: Standardize $z = \frac{x-\mu}{\sigma} = \frac{180-170}{7} = 1.43$
- Step 2: Use the standard normal table

$$P(X > 180) = P(Z > 1.43)$$

From the standard normal table,

$$P(Z < 1.43) = 0.92364$$
$$P(Z > 1.43) = 1 - 0.92364 = 0.07636$$
$$P(X > 180) \approx 0.07636$$

Normal Approximation to the Binomial Distribution

- When n is large enough, the binomial distribution can be approximated by a normal distribution.
- The mean is $\mu = np$, and variance is $\sigma^2 = npq$
- The approximation is usually good when

$$np > 5 \quad \text{and} \quad nq > 5$$

- Because the binomial distribution is **discrete** and the normal distribution is **continuous**, we use a **continuity correction**.

$$P(X \leq x) \approx P(Y \leq x + 0.5)$$
$$P(X \geq x) \approx P(Y \geq x - 0.5)$$
$$P(a \leq X \leq b) \approx P(a - 0.5 < Y < b + 0.5)$$

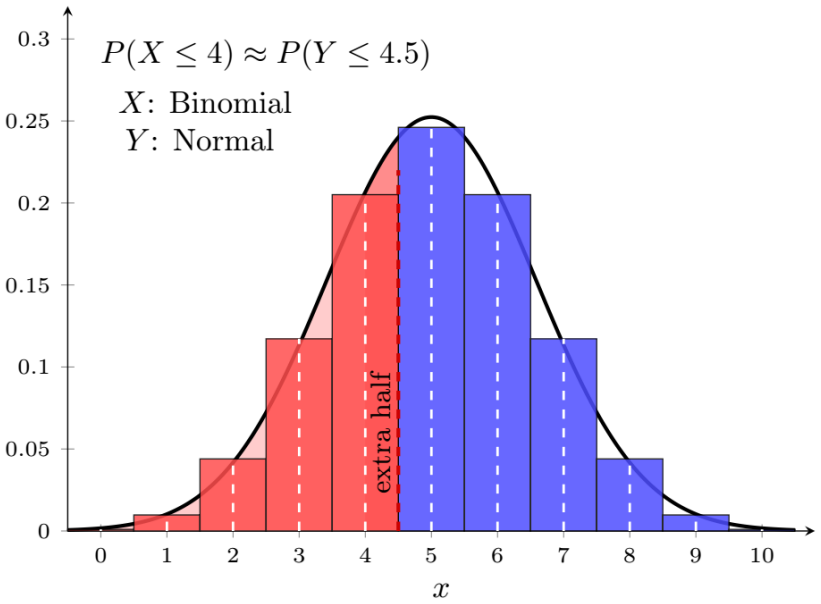
- where Y is the equivalent normal variable.

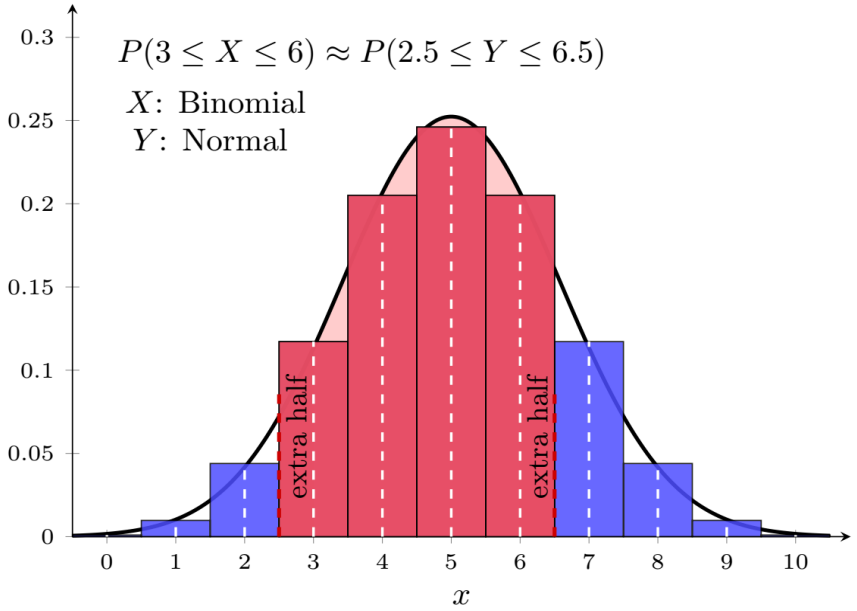
Normal Approximation to the Binomial Distribution: Example

A computer lab has 10 PCs. For each PC, the probability that it is ON at a certain time is 0.5. Let X = number of PCs that are ON. Find:

1. the probability that **at most 4** PCs are ON.
2. the probability that **from 3 to 6** PCs are ON.

$$\mu = np = 5, \qquad \sigma = \sqrt{npq} = \sqrt{2.5} \approx 1.58$$





1. $P(X \le 4) \approx P(Y \le 4.5)$

$$\approx P\left(Z \le \frac{4.5-5}{1.58}\right)$$

$$\approx P(Z \le -0.32) \approx 0.375$$

2. $P(3 \le X \le 6) \approx P(2.5 \le Y \le 6.5)$

$$\approx P\left(\frac{2.5-5}{1.58} \le Z \le \frac{6.5-5}{1.58}\right)$$

$$\approx P(-1.58 \le Z \le 0.95)$$

$$\approx P(Z \le 0.95) - P(Z \le -1.58)$$

$$\approx 0.772$$

Normal Approximation to the Binomial Distribution: Example

A factory produces memory chips. Each chip has a **10%** chance of being defective. A box contains **100 chips**. Let X be the number of defective chips in the box. Find the probability that the box contains **at least 13 defective chips**.

$$\mu = np = 100(0.1) = 10$$
$$\sigma^2 = npq = 100(0.1)(0.9) = 9$$

Check the conditions:

$$np = 10 > 5 \quad , \quad nq = 90 > 5$$

So the approximation is reasonable.

$$P(X \geq 13)$$

Using continuity correction:

$$P(X \geq 13) \approx P(Y \geq 12.5)$$

Standardize:

$$z = \frac{12.5 - 10}{3} = 0.83$$

So:

$$P(X \geq 13) \approx P(Z \geq 0.83) = 1 - P(Z < 0.83)$$

From the standard normal table:

$$P(Z < 0.83) = 0.79673$$

Therefore,

$$P(X \geq 13) \approx 1 - 0.79673 = 0.20327$$

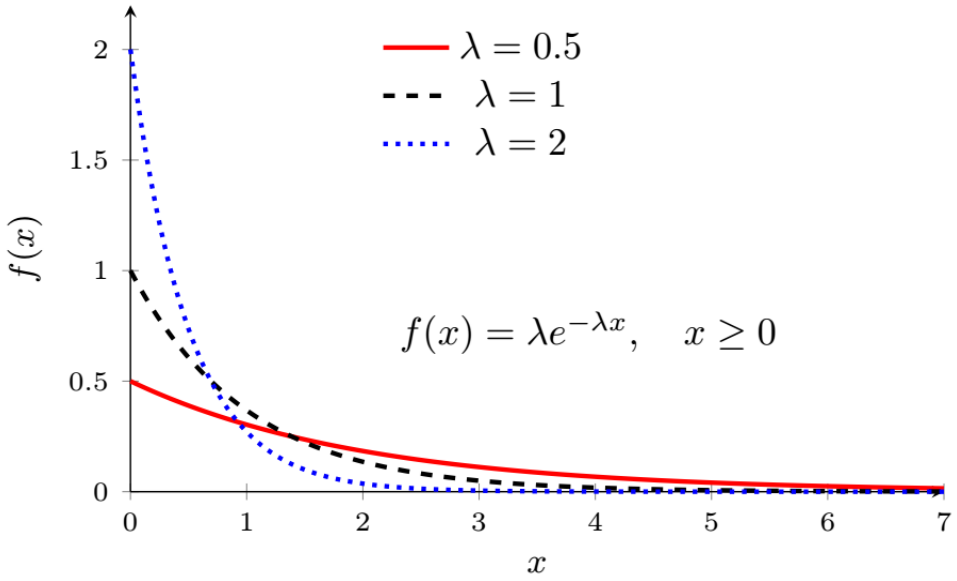
Exponential Distribution

- The **exponential distribution** models **waiting time until the next event**.
- Examples:
  - time until the next customer arrives
  - time until the next phone call
  - time until the next network packet arrives
  - time until a network switch fails
- Let X= waiting time.
- If events happen at an average rate $\lambda > 0$, then X may follow an exponential distribution.

- Its probability density function is:

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

- λ is called the **rate parameter**.
- Larger λ means events happen more often, so the average waiting time is smaller.
- The shown curves are **right-skewed**: short waits are more common than long waits.



Example: Lifetime of a Server Hard Drive

A data center uses a common model of server hard drive. During its normal operating life, engineers model the time until failure by an exponential distribution. If the average lifetime is 5 years. Then $\lambda = \frac{1}{5} = 0.2$ failure per year.

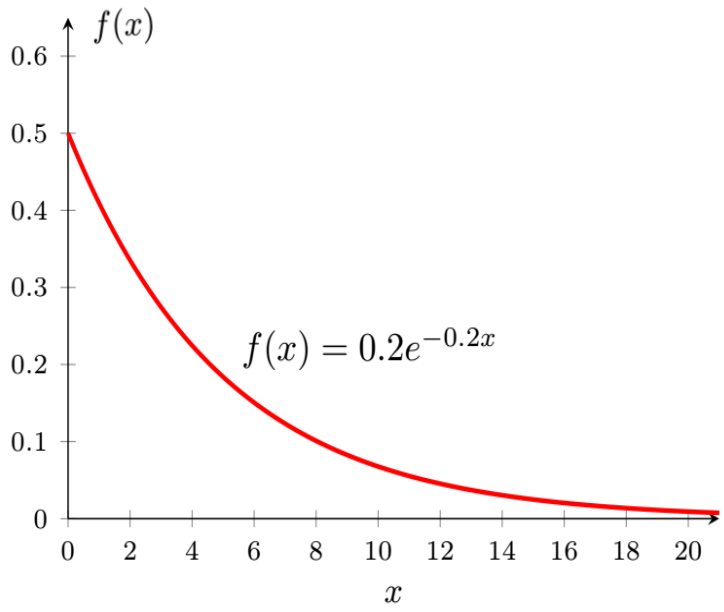
Let X= time until failure of one hard drive (in years**)**.

1. What is the probability that a hard drive fails within 3 years:

$$P(X \leq 3) = \int_0^3 0.2\, e^{-0.2x}\, dx \approx 0.4512$$

2. What is the probability that a hard drive lasts more than 3 years:

$$P(X > 3) = \int_3^\infty 0.2\, e^{-0.2x}\, dx \approx 0.5488$$



Thus:

- About **45.1%** of such drives fail within 3 years.
- About **54.9%** last more than 3 years.

This information can help decide when to replace hard drives, how reliable they are, and how many spare drives to keep ready.

Exponential Distribution: Mean, Variance, Standard Deviation, and CDF

For an exponential random variable,

$$E(X) = \frac{1}{\lambda} \quad , \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2} \quad , \qquad \sigma = \sqrt{\mathrm{Var}(X)} = \frac{1}{\lambda}$$

and

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

For the previous example of lifetime of a server hard drive :

1. Find the mean, variance, standard deviation of X.

$$E(X) = \frac{1}{0.2} = 5 \; years \quad , \qquad Var(X) = \frac{1}{(0.2)^2} = 25 \; years^2 \quad , \qquad \sigma = \sqrt{25} = 5 \; years$$

This large standard deviation means the lifetimes can differ a lot from one hard drive to another.

2. Find the probability that a hard drive fails within 1 year.

$$F(1) = P(X \leq 1) = 1 - e^{-0.2(1)} = 1 - e^{-0.2} \approx 0.1813$$

There is about an 18.1% chance that the hard drive fails during the first year.

3. Find the probability that a hard drive fails within 8 years.

$$F(8) = P(X \leq 8) = 1 - e^{-0.2(8)} = 1 - e^{-1.6} \approx 0.7981$$

There is about a 79.8% chance that the hard drive fails within 8 years.

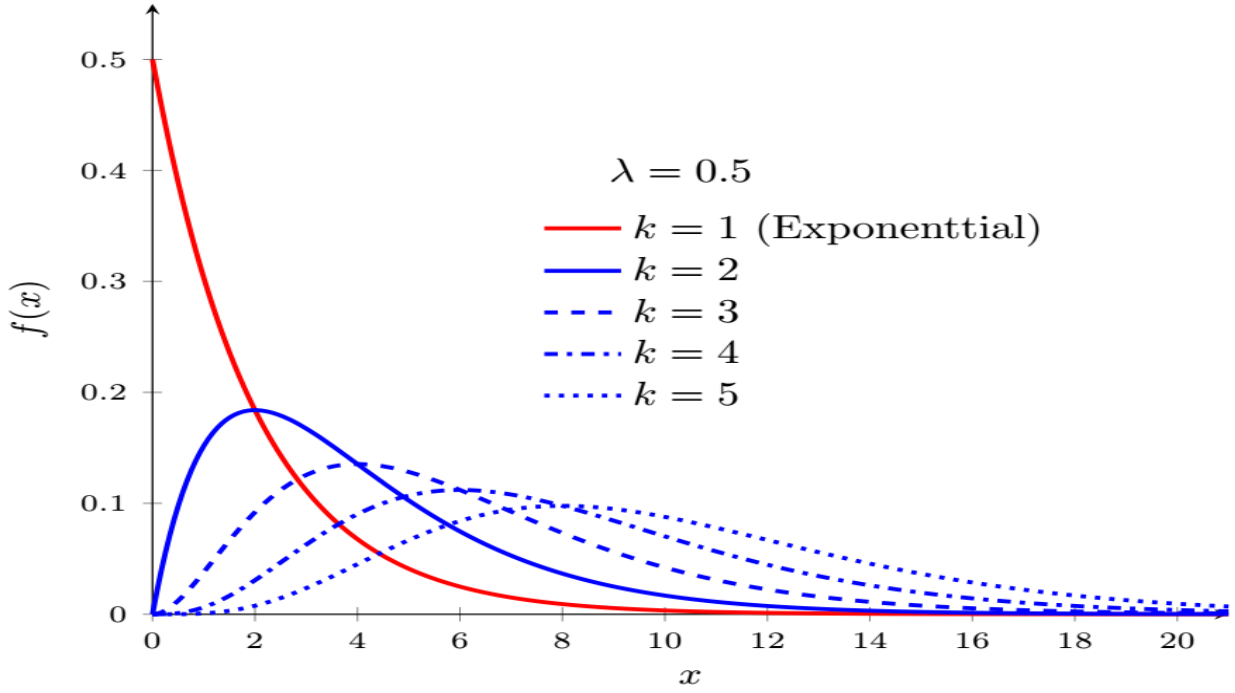
Erlang / Gamma Distribution

- The exponential distribution models waiting time until the **first event**.
- The **Erlang / Gamma distribution** models waiting time until the k-*th* event.
- It is used when events occur randomly over time and we wait for many events to happen.
- This is similar to the discrete case:
  - **Geometric**: number of trials until the first success
  - **Negative binomial**: number of trials until the r-*th* success
- So, Erlang / Gamma is the continuous version of that same idea.
- If X is the waiting time until k events occur at rate λ, then its PDF is

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- **Gamma** distribution: k can be any positive **real** number.
- **Erlang** distribution: k is any positive **integer** (so Erlang is a special case of Gamma).

Erlang vs. Exponential Distributions



- Again, x is the time we wait until k events happen.
- Exponential ($k = 1$): waiting time for the **first event**
- Erlang ($k \geq 2$): waiting time for ***k*** **events**
- In **exponential**, very small waiting times are likely, but in **Erlang** they are unlikely because multiple events are required.
- As k increases → waiting time becomes longer
- More events → the curve becomes less skewed and more spread around an average (mean).

Erlang / Gamma Distribution: Example

A server receives requests randomly at an average rate of $\lambda = 2$ requests per second. Let X be the waiting time until **3 requests** arrive.

Thus, X can be modeled by an Erlang distribution with $k = 3$ and $\lambda = 2$

So, its PDF is

$$f(x) = \frac{2^3 x^{3-1} e^{-2x}}{(3-1)!} = 4x^2 e^{-2x}, \quad x \geq 0$$

1. Find the probability that 3 requests arrive within 2 seconds.

$$P(X \leq 2) = \int_0^2 4\, x^2 e^{-2x}\, dx \approx 0.762$$

2. Find the probability that 3 requests arrive within 5 seconds.

$$P(X \leq 5) = \int_0^5 4\, x^2 e^{-2x}\, dx \approx 0.9972$$

Thus:

- There is about 76.2% chance that 3 requests arrive within 2 seconds.
- There is about 99.72% chance that 3 requests arrive within 5 seconds.
- Almost certainly, 3 requests will arrive within 5 seconds.
- This model is useful for estimating how quickly requests arrive at a server.

Erlang / Gamma Distribution: Mean, Variance, and Standard Deviation

- For the waiting time until k events:

$$E(X) = \frac{k}{\lambda} \quad , \qquad Var(X) = \frac{k}{\lambda^2} \quad , \qquad \sigma = \sqrt{Var(X)} = \frac{\sqrt{k}}{\lambda}$$

- Using the previous example (3 requests):

$$k = 3 \quad \text{and } \lambda = 2$$

1. Find the mean (average waiting time):

$$E(X) = \frac{3}{2} = 1.5 \textit{ seconds}$$

2. Find the variance:

$$\mathrm{Var}(X) = \frac{3}{2^2} = \frac{3}{4} = 0.75 \textit{ seconds}^2$$

3. Find the standard deviation:

$$\sigma = \frac{\sqrt{3}}{2} \approx 0.87 \textit{ seconds}$$

Observations:

- On average, it takes 1.5 seconds for 3 requests to arrive.
- The waiting time varies because arrivals are random.
- More required events (k) → longer average waiting time.

How to Choose a Distribution

If the random variable represents …	Use …
a value between (a) and (b), with all values equally likely	Continuous Uniform
a measurement centered around an average (bell-shaped)	Normal
the waiting time until the first event	Exponential
the waiting time until the k-th event	Erlang / Gamma

- First ask what the random variable X means.
- Is X a measurement, or is it a waiting time?
- If it is a waiting time, ask whether it is for one event or several events.
- The meaning of X helps us choose the right distribution.

Simple rule

Choose the distribution based on what the random variable represents.

Assessment: Choose the Appropriate Distribution

For each of the following situations, choose the most appropriate distribution (Continuous Uniform, Normal, Exponential, or Erlang / Gamma)

1. A random delay is chosen between 1 ms and 5 ms, and all values are equally likely.
2. The time until the next user arrives at a website.
3. The total waiting time until 4 requests arrive at a server.
4. Sensor measurement errors are centered around 0, with small errors more common than large ones.
5. The lifetime of electronic components, assuming failures occur randomly at a constant rate.
6. A random number between 10 and 20, with values near 15 more common than values near 10 or 20.
7. The waiting time until the first failure in a system.
8. The time until 5 print jobs arrive in a printer queue.

Summary for Unit-4

- A continuous random variable can take any value in an interval, and $P(X = x) = 0$.
- A probability density function (PDF) describes how probability is distributed over continuous values.
- Probabilities are found as areas under the curve, not at single points.
- The cumulative distribution function (CDF) gives $P(X \leq x)$.
- The expected value gives the long-run average, while variance and standard deviation measure spread.
- Different continuous distributions model different situations:
  - **Continuous Uniform**: all values in an interval are equally likely
  - **Normal**: values are centered around a mean
  - **Exponential**: waiting time until the first event
  - **Erlang / Gamma**: waiting time until the k-th event
- Standardization converts any normal variable to the standard normal Z.
- We can sometimes use the normal distribution to approximate a binomial distribution, but only if the conditions are right and we use continuity correction.
- Choosing the correct distribution depends on what the random variable represents.