Continuous Random Variable and Their Distributions

Siman Giri

1.Probability of a Continuous Sample Space.

1.1 Continuous Probability: Intuition

- **Example**:
  - Your friend tells you that she will come arrive for your birthday party sometime after or equal to 1 p.m. to 2 p.m. but she cannot give you any more information as her schedule is quite hectic. Your friend is very dependable, so you are sure that she will stop by your birthday party, but other than that we have no information about the arrival time. Thus, we assume that the arrival time is completely random in the interval. Let be T the arrival time.
    a) What is the sample space S?
    b) What is the probability of P(1.5)?Why?

1.2 Continuous Sample Space.

- What is the sample space S?
  - Since any real number in [1,2) is a possible outcome, thus the sample space is: $\boldsymbol{S = [1, 2)}$

1.3 Continuous Sample Space.

- What is the sample space S?
  - Since any real number in [1,2) is a possible outcome, thus the sample space is: $\boldsymbol{S} = [\mathbf{1}, \mathbf{2})$

Definition: Continuous Sample Space:

A continuous sample space is a sample space containing outcomes defined in the terms of interval and some interval can have infinite number of points.

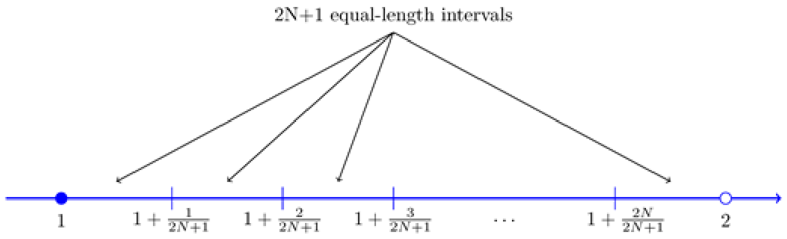
Example: Continuous Sample Space:

Consider {Experiment:} observing the distance a ball can be thrown, say d.
The Sample space is now continuous and defined as:

$$\Omega = \{d | d \in \mathbb{R}_{+} \& \mathrm{d} > 0\}$$

1.4 Assigning Probability: Continuous Sample Space.

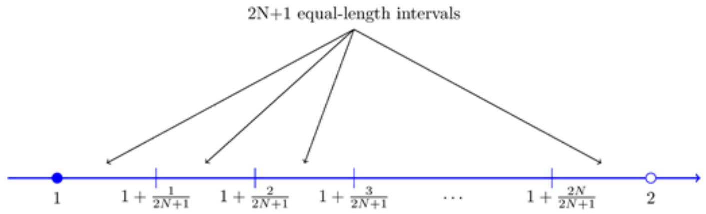
- What is the probability of P(1.5)? Why?
  - $\boldsymbol{P(1.5) \approx 0.}$
- **Why?**
  - Let us divide the sample space $[1,2)$ interval to $2N+1$ equal-length and disjoint intervals as in figure:



    - By the axioms of probabilities: $\boldsymbol{P(\Omega)=1}$ i.e. sum of all the probabilities $\boldsymbol{P(T: T \in [1,2)) = 1.}$
      - If $\boldsymbol{P(T: T \in [1,2))}$ is **non zero** the **sum of all probabilities** will be **finite**.
      - If $P(T: T \in [1,2))$ is **zero** then the **sum of all probabilities** will **vanish**.
      - **How can we solve this paradox?**

1.4 Assigning Probability: Continuous Sample Space.

- What is the probability of P(1.5)? Why?
  - $\boldsymbol{P(1.5) \approx 0.}$
- **Why?**
  - Let us divide the sample space $[\mathbf{1,2})$ interval to $\mathbf{2N+1}$ equal-length and disjoint intervals as in figure:



- By the axioms of probabilities: $\boldsymbol{P(\Omega) = 1}$ i.e. sum of all the probabilities $\boldsymbol{P\big(T\!:T \in [1,2)\big) = 1.}$
  - If $\boldsymbol{P\big(T\!:T \in [1,2)\big)}$ is **non zero** the **sum of all probabilities** will be **finite**.
  - If $P\big(T\!:T \in [1,2)\big)$ is **zero** then the **sum of all probabilities** will **vanish**.
  - How can we solve this **paradox**?
    - This paradox **can be resolved** with the notion that **the sum of a continuum of values** is commonly **replaced** by an **integral**.

1.5 Assigning Probability: Continuous Sample Space.

- Continuous sample spaces do **not** have **distinct outcomes** to which probabilities can be assigned.
  - Instead, probabilities are assigned to ***intervals*** of the **sample space**, and these probabilities are described using a real-valued *probability function*, such as: $f_X(x)$.
- **Axioms of probability** now written as:
  1. **Non-negativity**: integration over any region of the sample space must never produce a negative value: $f_X(x) \geq 0$ for all values of x.
  2. **Exhaustive:** Over the whole sample space, the probability function must integrate to one: $\int_{\Omega} f(x)\, dx = 1$
  3. **Additive:** The probability of the union of any non-overlapping regions is the sum of the individual regions.

Cautions!!!
The probability function $f_X(x)$ does **not** give the probability of observing the value $X = x$; because the sample space is continuous, the **probability** of observing any **single point** is almost($\approx$)0. Probabilities are computed for intervals.
This implies that $f_X(x) > 1$ may be true for some values of x, provided the total area over the sample space is 1.

A. Preliminary Concepts!!!

Anti-Derivative - Integration.

Disclaimer!!!
We will assume you can compute the most familiar forms of derivatives and integrals by hand!!!

A.1 Integration: Notation

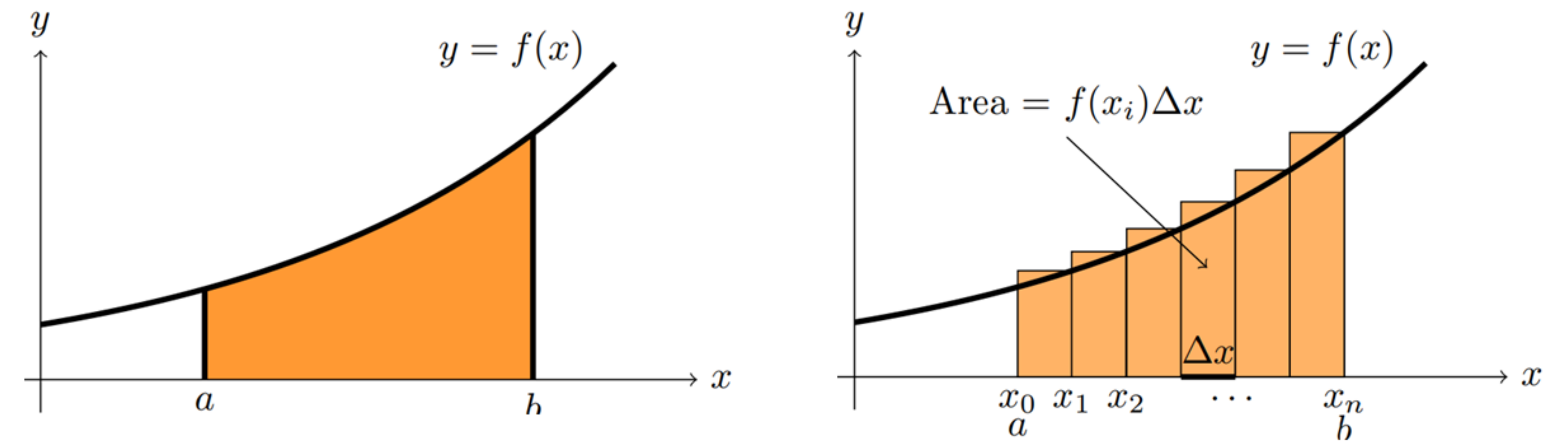
- **Definite Integral**:
  - For a function $\boldsymbol{f(x)}$, considered between $\boldsymbol{x = a}$ **and** $\boldsymbol{x = b}$, the limit we obtain for the exact area is denoted by: $\int_a^b f(x)\, dx$ and is called the **definite integral**.
    - {a.k.a integral of $f(x)$ w.r.to x from $x = a$ and $x = b$.}
    - The symbol $\int$ is called an **integral sign.**
    - The numbers **a** and **b** are called the **terminals or endpoints** of the integral
    - The function $\boldsymbol{f(x)}$ is called the **integrand**.
- **Indefinite Integral**:
  - An **indefinite integral** is an integral written **without terminals**;
    - it simply asks us to find a general anti-derivative of the integrand.
    - It is not one function but a family of functions, differing by constants; and so the answer must have a '+ constant' term to indicate all anti-derivatives.
      - Such as: $\int x^2\, dx = \frac{x^3}{3} + c$

A.2 Definite Integral: Introduction.

- We do not expect you to be calculus whizzes, but at least be comfortable with following two views of **definite integral**:
  - $\int_a^b f(x)\,dx$ = **area under the curve** $y = f(x)$
  - $\int_a^b f(x)\,dx$ = **"sum of** $f(x)dx$**".**
- **Definite Integral** provides an **estimate** of **an area under the graph** between endpoints/terminals.

A.3 Definite Integral: area under the curve.

- The connection between the two points from previous slide is:
  - **area ≈ sum of rectangle areas =** $\boldsymbol{f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x = \sum_1^n f(x_i)\Delta x}$.



Area is approximately the sum of rectangles.

- As the width $\boldsymbol{\Delta x}$ of the intervals gets **smaller** the **approximation** becomes **better**.

A.3 Definite Integral: Theorem and Properties.

Fundamental Theorem of Calculus:

Let $f(x)$ be a continuous real-valued function on the interval $[a, b]$.

Then $\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a)$ where $F(x)$ is any anti-derivative of $f(x)$.

The notation $[F(x)]_a^b$: shorthand to substitute $x = a$ and $x = b$ into $F(x)$.

Linearity of Integration:

a) If f and g are continuous functions on the interval $[a, b]$, then

$$\int_a^b f(x) \pm g(x)\, dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx.$$

b) If f is a continuous functions on $[a, b]$ and k is a real constant, then

$$\int_a^b kf(x)\; dx = k\int_a^b f(x)dx.$$

Additivity of Integration:

Let $f: [a, b] \to \mathbb{R}$ be a continuous , where **a, b** are **real numbers**. Let **c** be a **real number** between **a** and **b**. Then:

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

Integral: Cheat Sheet.

RULE: PROPERTIES OF THE DEFINITE INTEGRAL

If the limits of integration are the same, the integral is just a line and contains no area:

$$\int_a^b f(x)\,dx = 0$$

If the limits are reversed, then place a negative sign in front of the integral:

$$\int_b^a f(x)\,dx = -\int_a^b f(x)dx$$

Power rule of Integration:

$$\text{a) } \int x^n\,dx = \frac{x^{n+1}}{n+1} + C$$

for any real constants C with $n \neq -1$.

$$b)\ \int (ax+b)^n = \frac{(ax+b)^{n+1}}{a(n+1)} + C$$

for a, b, c, n for any real constants with $a \neq 0$, $n \neq -1$.

Integration by parts:

$$\int u dv = uv - \int v du$$

u = function of $u(x)$.
dv = variable dv.
v = function of $v(x)$.
du = variable du

Integral: Practice Questions.

a) Find $\int_0^2 (-x^2 - x + 2)dx$.

b) Compute $\int_{-2}^{2} (x^3 - x)\, dx$.

c) Find $\int_5^0 (-2x - 3)dx$.

d) Find $f(x) = 6x^3 - 4x^2 + 2x - 3$ over the interval [1,3] .

Integral: Practice Questions.

a) Find $\int_0^2 (-x^2 - x + 2)dx$.

b) Compute $\int_{-2}^{2} (x^3 - x)\, dx$.

c) Find $\int_5^0 (-2x - 3)dx$.

d) Find $f(x) = 6x^3 - 4x^2 + 2x\ - 3$ over the interval [1,3] .

a) solution:

$$\int_0^2 (-x^2 - x + 2)dx = \left[-\frac{1}{3}x^3 - \frac{1}{2x^2} + 2x\right]_0^2 = \left(-\frac{8}{3} - 2 + 4\right) - (0) = -\frac{2}{3}.$$

b) solution:

$$\int_{-2}^{2} (x^3 - x)dx = \int_{-2}^{2} x^3 dx - \int_{-2}^{2} x dx = \left[\frac{x^4}{4}\right]_{-2}^{2} - \left[\frac{x^2}{2}\right]_{-2}^{2}$$

2.Continuous Random Variable and Probability Density Function.

Introduction.

2.1 Continuous Random Variables and Probability Density Function.

- A continuous random variable takes a range of values, which may be finite or infinite in extent.
  - Here are a few examples of ranges: $[0,1], [0,\infty), (-\infty,\infty), [a,b]$.

Definition: Continuous Random variable

A random variable $\boldsymbol{X}$ is continuous if there is a function $\boldsymbol{f_X(x)}: \mathbb{R} \to [\mathbf{0},\infty)$ such that for any $\mathbf{a} \leq \boldsymbol{b}$ we have

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

The function $\boldsymbol{f_X}$ is called the probability density function(pdf) of $\boldsymbol{X}$.

2.2 pdf: Properties

- Not any function can serve as a pdf. The pdf of any random variable will always have the following properties:
  1. $\boldsymbol{f(x) \geq 0}$ {f is nonnegative}
  2. $\boldsymbol{\int_{-\infty}^{+\infty} f(x)dx = 1}$ {equivalent to: $P(-\infty < X < \infty) = 1$}

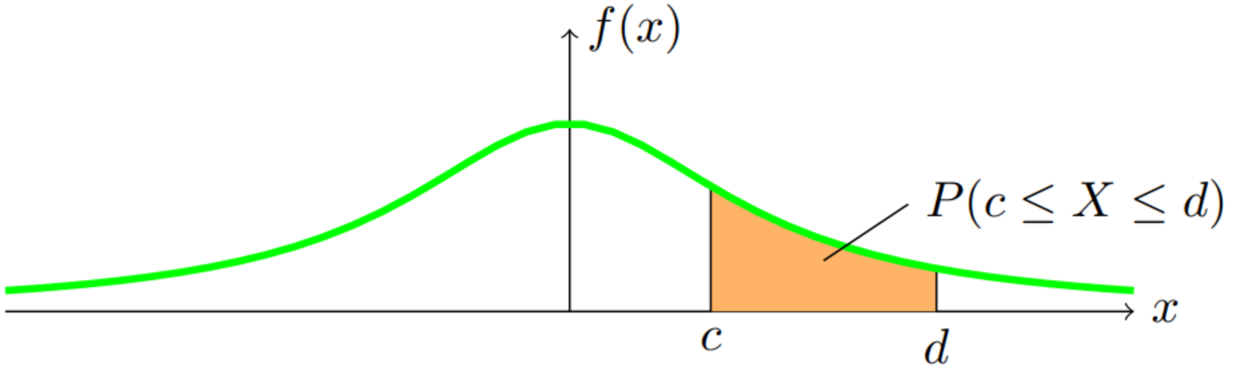
2.3 pdf vs. pmf

- The probability density function $f_X(x)$ of a continuous random variable is the analogue of the probability mass function $P_X(x)$ of a discrete random variable.
- Important Differences are:
  - Unlike $p(x)$; $f(x)$; is not a probability. We integrate it to get a probability.
  - Since $f(x)$; is not a probability, there is no restriction that $f(x)$; be less than or equal to 1.

Distribution	Discrete	Continuous
Definition	$P(x) = P\{X = x\}$ (pmf)	$f(x) = F'(x)$ (pdf)
Computing probabilities	$\boldsymbol{P}\{X \in A\} = \sum_{x \in A} P(x)$	$\boldsymbol{P}\{X \in A\} = \int_A f(x)dx$
Cumulative distribution function	$F(x) = \boldsymbol{P}\{X \leq x\} = \sum_{y \leq x} P(y)$	$F(x) = \boldsymbol{P}\{X \leq x\} = \int_{-\infty}^{x} f(y)dy$
Total probability	$\sum_x P(x) = 1$	$\int_{-\infty}^{\infty} f(x)dx = 1$

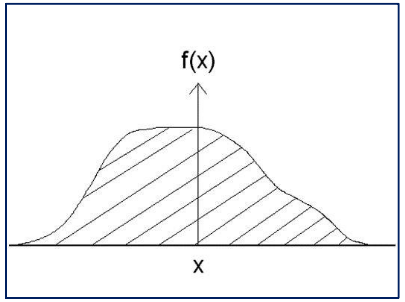
2.4 Graphical View of pdf.

- If you graph the probability density function of a continuous random variable X then:
- **$P(c \leq X \leq d)$ = area under the graph between c and d.**

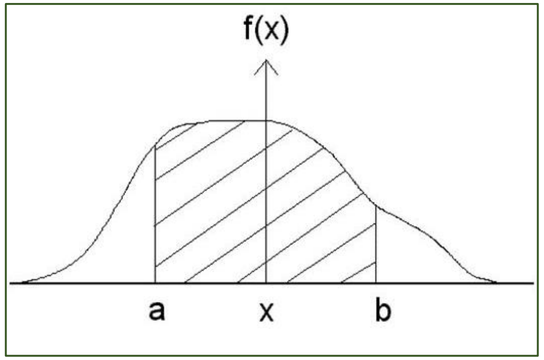


- **What is the total area under the pdf $f_X(x)$?**

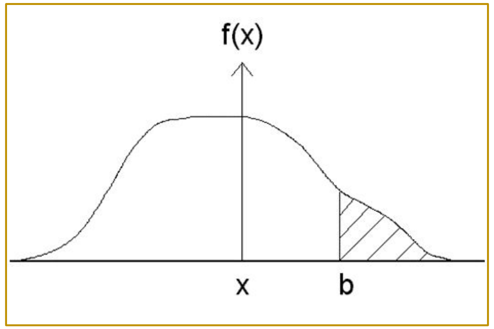
2.5 Graphical View of pdf: Hint for determining the area.



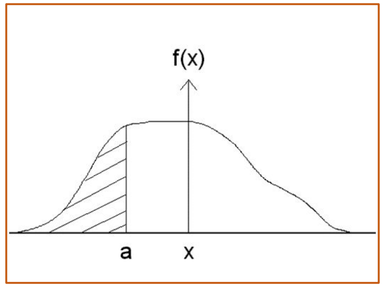
Total area = $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$



Area $P\left[X \in [a, b]\right] = \int_a^b f_X(x)\,dx$



Area $P_X(x \geq b) = \int_b^{\infty} f_X(x)\,dx$



Area $P_X(x \leq a) = \int_{-\infty}^{a} f_X(x)\,dx$

2.6 Expectation and Variance of crv.

Definition: Expectation of Continuous Random Variable

For a **continuous random variable X** with pdf $f_X(x)$ we define the expectation $\mathbb{E}[X]$as:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

as long as $\int_{-\infty}^{\infty} |x f_X(x)|\, dx < \infty$. When this is $+\infty$ expectation is not defined.

Definition: Variance of Continuous Random Variable

The definition of the variance and the standard deviation are analogous to their discrete versions:

$$Var[X] = \int_{-\infty}^{\infty} (y - \mu_X)^2 f_X(x)\, dx$$

Here $\mu_X = \mathbb{E}[X]$
and $SD[X] = \sqrt{Var[X]}$

3. Cumulative Distribution Function.

3.1 CDF: Continuous Definition

Definition: Continuous Random Variable

For a continuous random variable X the Cumulative Distribution Function, written $F(x)$ is:

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

- Why is the CDF the probability that a random variable takes on a value **less than** the input value as opposed to greater than?
  - It is a matter of convention. But it is a useful convention. Most probability questions can be solved simply by knowing the CDF (and taking advantage of the fact that the integral over the range $-\infty$ *to* ∞ is 1.) For Example:

Probability Query	Solution	Explanation
$P(X < a)$	$F(a)$	Definition of CDF.
$P(X \leq a)$	$F(a)$	$P(X = a) = 0$
$P(X > a)$	$1 - F(a)$	$P(X < a) + P(X > a) = 1$
$P(a < X < b)$	$F(b) - F(a)$	$F(a) + P(a < X < b) = F(b)$

3.2 CDF: Discrete Definition

- For Discrete Random Variable we define the CDF as
  - $\boldsymbol{F(x) = P(X \leq x)\ for\ any\ x \in \mathbb{R}.}$
- but did not have much occasion to use it.

3.3 CDF: Properties

- Summarry of the most important properties of cumulative distribution functions:
  - $\boldsymbol{F(x)}$ is non decreasing i.e. if $\boldsymbol{a \leq b}$ then $\boldsymbol{F(a) \leq F(b)}$.
  - $\boldsymbol{0 \leq F(x) \leq 1}$.
  - The maximum of $\boldsymbol{F(x)}$ is: $\boldsymbol{\lim_{x \to \infty} F(x) = 1}$.
  - The minimum of $\boldsymbol{F(x)}$ is: $\boldsymbol{\lim\limits_{x \to \infty} F(x) = 0}$.
  - $\boldsymbol{P(a \leq X \leq b) = F(b) - F(a)}$
  - $\boldsymbol{F'(x) = f(x)}$.
    - **{PDF → CDF: Integration**
    - **CDF → PDF: Derivative }**

pdf & cdf: Example

Let **X** be a random variable with PDF:

$$f_X(x) = \begin{cases} C(4x - 2x^2) & when\ 0 < x < 2 \\ 0 & otherwise \end{cases}$$

Find following:

a) **Value of constant C.**

b) $P(X > 1)$?

{Hint: PDF sum must be 1}

pdf & cdf: Example

a) solution:

$$\int_0^2 C(4x - 2x^2)dx = 1$$

$$C\left[2x^2 - \frac{2x^3}{3}\right]_0^2 = 1$$

$$C\left[\left(8 - \frac{16}{3}\right) - 0\right] = 1$$

$$\mathrm{C} = \frac{3}{8}.$$

pdf & cdf: Example

a) solution:

$$\int_0^2 C(4x-2x^2)dx = 1$$

$$C\left[2x^2-\frac{2x^3}{3}\right]_0^2 = 1$$

$$C\left[\left(8-\frac{16}{3}\right)-0\right] = 1$$

$$C = \frac{3}{8}.$$

b) Solution:

$$P(X>1) = \int_1^{\infty} f_X(x)dx$$
$$= \int_1^2 \frac{3}{8}(4x-2x^2)dx$$
$$= \frac{3}{8}\left[2x^2-\frac{2x^3}{3}\right]_1^2$$
$$= \frac{3}{8}\left[\left(8-\frac{16}{3}\right)-\left(2-\frac{2}{3}\right)\right]$$
$$= \frac{1}{2}.$$

Popular Continuous Distribution.

4. Uniform Distribution.

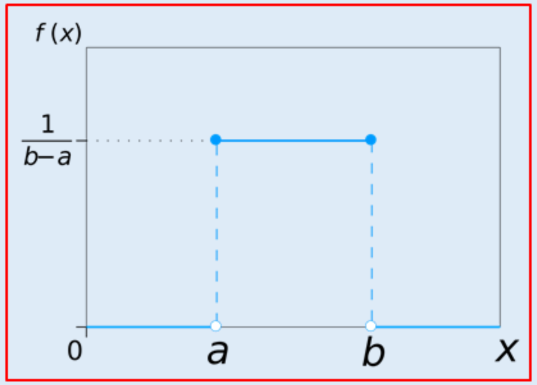
4.1 Uniform Distribution : pdf

- The continuous uniform distribution has a constant pdf over a given range.

Definition Uniform Distribution

A continuous random variable X is said to have a **Uniform distribution** over the interval $[a, b]$, shown as $X \sim Uniform(a, b)$, if its **PDF** is given by:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & x < a \text{ or } x > b \end{cases}$$



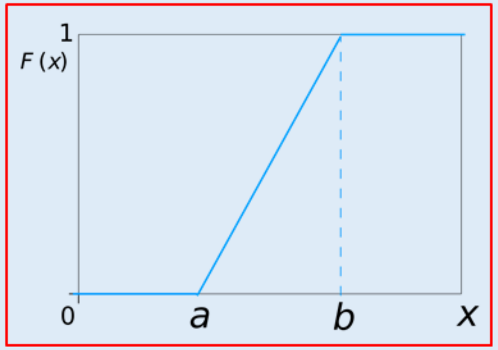
Probability Density Function

4.2 Uniform Distribution : cdf

Definition Uniform Distribution (CDF):

For a random variable X with the **continuous distribution** given in previous slide the distribution function is:

$$F_X(x; a, b) = \begin{cases} 0 & for\ x < a; \\ \dfrac{x-a}{b-a} & for\ a \le x \le b; \\ 1 & for\ x > b. \end{cases}$$



Continuous Distribution Function.

4.3 Uniform Distribution : cdf properties.

Expectation of Uniform Distribution Function:

$$\mathbb{E}[X] = \frac{a+b}{2}.$$

Proof(Optional):

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-\infty}^{0} 0x dx + \int_{a}^{b} \frac{x}{b-a} dx + \int_{b}^{\infty} 0x dx$$

$$\mathbb{E}[X] = \left[\frac{x^2}{2(b-a)}\right]_b^a = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

Variance of Uniform Distribution Function:

$$var(X) = \frac{(b-a)^2}{12}.$$

Proof(Optional):

To find the variance, we first find $\mathbb{E}[X^2]$using **LOTUS** i.e.

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_{a}^{b} x^2 \left(\frac{1}{b-a}\right) dx = \frac{a^2+ab+b^2}{3}$$

Thus variance is :

$$var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{a^2+ab+b^2}{3} \Big/ \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

4.4 Uniform Distribution: Example

- You are running to the bus stop. You don't know exactly when the bus arrives. You believe all times between and are equally likely. You show up at 2.15 pm.
- What is $P(wait < 5mins.)$?

4.4 Uniform Distribution: Example

- You are running to the bus stop. You don't know exactly when the bus arrives. You believe all times between and are equally likely. You show up at 2.15 pm.
  - What is $P(wait < 5mins.)$?

Let T be the time in minutes; after 2 pm that the bus arrives.
As given: $T \sim U(a = 0 \; and \; b = 30)$.
The probability of wait for 5 mins. is equal to the probability bus shows up between 2:15 and 2:20 i.e:

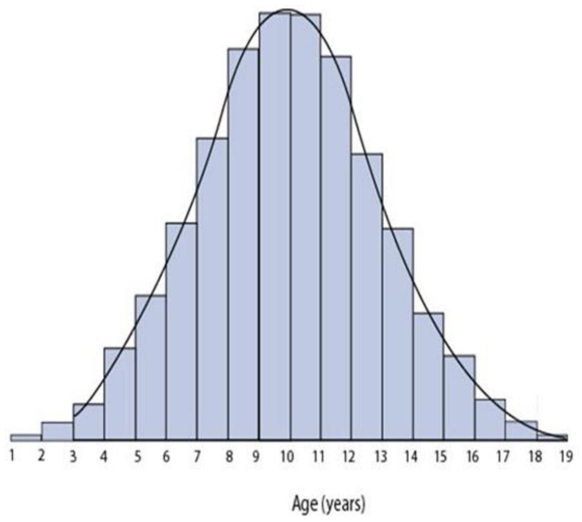
$$P(wat\ under\ 5\ mins) = P(15 < T < 20)$$
$$= \int_{15}^{20} f_T(x)dx$$
$$= \int_{15}^{20} \frac{1}{b-a} dx$$
$$= \left[\frac{x}{30}\right]_{15}^{20}$$
$$= \frac{20}{30} - \frac{15}{30}$$
$$= 5/30.$$

Popular Continuous Distribution.

5. Normal Distribution.

5.1 Normal Distributions: Introduction

- The most well-known **continuous distribution** is probably the ***normal distribution*** (**or *Gaussian distribution named after Carl Friedrich Gauss***), sometimes called the ***bell-shaped*** ***(reference to the shape)*** distribution.



5.2 Normal Distributions: pdf

Definition: Normal Probability Density Function.

If a continuous random variable X has the **pdf**:

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

then X has a normal distribution.

Two parameters are the mean μ such that $-\infty < \mu < \infty$; and the variance σ^2 such that $\sigma^2 > 0$.

We write $X \sim N(\mu, \sigma^2)$.

5.3 Normal Distributions: cdf

Definition: Normal Probability Distribution Function.(Optional)

For a random variable $\boldsymbol{X}$ with the normal distribution given in the previous slide distribution function is:

$$F_X(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left\{\frac{(t-\mu)^2}{2\sigma^2}\right\} dt = \frac{1}{2}\left\{1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right\}$$

where erf(.) is the error function given by:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x} \exp(-t^2)\, dt \;\; for\; x \in \mathbb{R},$$

It is computed using computing software.

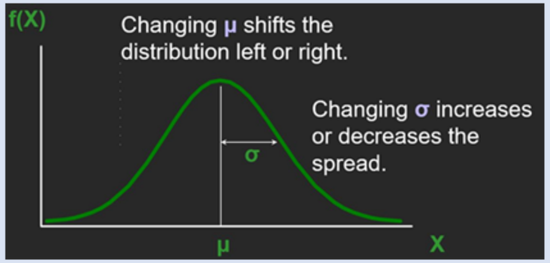
5.4 Normal Distributions: cdf properties.

Normal Distribution Properties (Expectation and Variance):

If $X \sim N(\mu, \sigma^2)$ then:

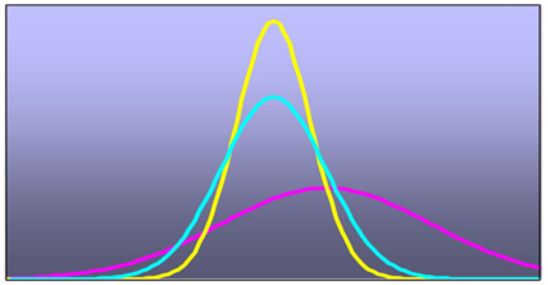
1. $\mathbb{E}[X] = \mu$
2. $var(X) = \sigma^2$

Normal Distribution Properties (shape):



$\boldsymbol{\mu}$ controls the location.

$\boldsymbol{\sigma}$ controls the spread.



By varying the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ we can obtain different normal distribution.

5.5 Standard Normal Distributions.

- The **standard normal distribution** is a **normal distribution** with **mean 0** and **variance 1**.
- We reserve $\boldsymbol{Z}$ for a **standard normal variable**, and pdf is given by
  - $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\}$
- and cdf is :
  - $F_X(x) = \frac{1}{2}\left\{1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right\} = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt = \Phi(z).$
    - erf(.) is a error function.
- **Standard normal distribution properties:**
  - **E(Z) = 0**
  - **var(Z) = 1.**

5.6 Normal to Standard Normal Distribution.

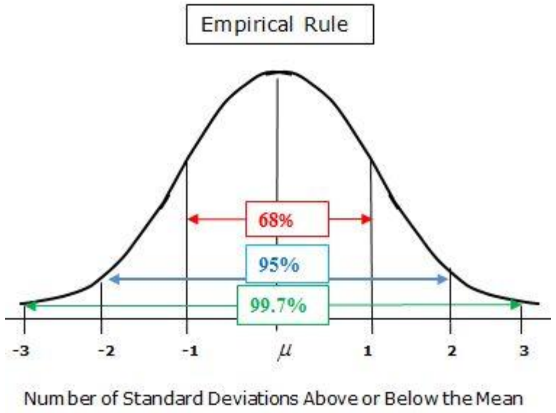
- Every Normal Random variable can be converted to Standard Normal Distribution using z-score defined as:
  - $\boldsymbol{z = \frac{x-\mu}{\sigma}}$.
- **Example:**
  - Following are the marks scored by
    - student A: 65 marks in Math
    - student B: 80 marks in English.
  - Who performed better given the following distributions:
    - Math ~ $N(\mu = 60\ and\ \sigma = 4)$ and English~$N(\mu = 79\ and\ \sigma = 2)$.

5.6 Normal to Standard Normal Distribution.

- **Example:**
  - Following are the marks scored by
    - student A: 65 marks in Math
    - student B: 80 marks in English.
  - Who performed better given the following distributions:
    - Math ~ $N(\mu = 60\ and\ \sigma = 4)$ and English~$N(\mu = 79\ and\ \sigma = 2)$.
- Solution:
  - $z_{score}A = 1.25$
  - $Z_{score}B = 0.5$

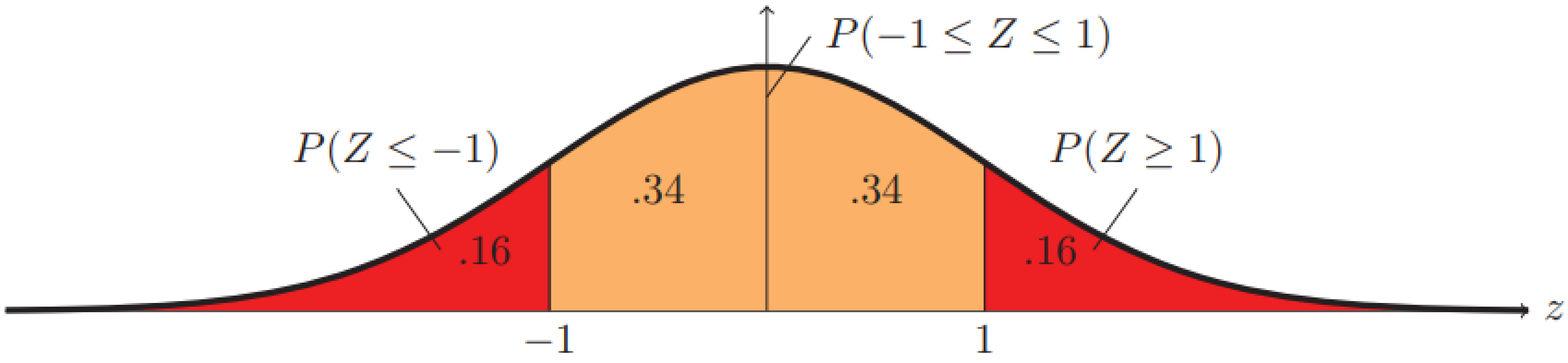
5.7 Normal Probabilities Distribution

- **Empirical rule:**
  - approx 68% of sample values are in the interval $[\boldsymbol{\mu} - \boldsymbol{\sigma}, \boldsymbol{\mu} + \boldsymbol{\sigma}]$.
  - approx 95% are in interval $[\boldsymbol{\mu} - 2\boldsymbol{\sigma}, \boldsymbol{\mu} + 2\boldsymbol{\sigma}]$.
  - all most all are in the interval $[\boldsymbol{\mu} - 3\boldsymbol{\sigma}, \boldsymbol{\mu} + 3\boldsymbol{\sigma}]$.



Tutorial!!!

- Problem solving and How to use table to solve different cases of z-score.



Thank You any Question!!!