Recall: Random variables

Random variable $X$ taking its values in the domain $\Omega$.

- if $\Omega \subseteq \mathbb{N}$, $X$ is discrete
- if $\Omega \subseteq \mathbb{R}$, $X$ is continuous
Recall: Random variables

Random variable $X$ taking its values in the domain $\Omega$.

- if $\Omega \subseteq \mathbb{N}$, $X$ is **discrete**
  the set of events is $\mathcal{P}(\Omega)$

- if $\Omega \subseteq \mathbb{R}$, $X$ is **continuous**
  the set of events is the set of all subsets of $\Omega$ that are a union or intersection of any number of intervals of $\Omega$.

Sometimes the distribution of $X$ is called “law” of $X$. 
Recall: Random variables

Random variable $X$ taking its values in the domain $\Omega$.

- if $\Omega \subseteq \mathbb{N}$, $X$ is discrete
  the set of events is $\mathcal{P}(\Omega)$

- if $\Omega \subseteq \mathbb{R}$, $X$ is continuous
  the set of events is the set of all subsets of $\Omega$ that are a union or intersection of any number of intervals of $\Omega$.

- the distribution of $X$ is the function: $A$ event $\mapsto P(X \in A) \in [0,1]$.

Sometimes the distribution of $X$ is called “law” of $X$. 
Recall: Random variables

Random variable $X$ taking its values in the domain $\Omega$.

- if $\Omega \subseteq \mathbb{N}$, $X$ is discrete
  the set of events is $\mathcal{P}(\Omega)$

- if $\Omega \subseteq \mathbb{R}$, $X$ is continuous
  the set of events is the set of all subsets of $\Omega$ that are a union or intersection of any number of intervals of $\Omega$.

- the distribution of $X$ is the function: $A$ event $\mapsto P(X \in A) \in [0, 1]$
  if $X$ is discrete, it is characterized by the function:

  $$ t \in \Omega \mapsto P(X = t) $$

  if $X$ is continuous, we saw that $\forall t \in \Omega, P(X = t) = 0$

Sometimes the distribution of $X$ is called “law” of $X$. 

Recall: Random variables

Discrete random variable $X$ taking its values in the domain $\Omega$.

- **Expected value, or mean of $X$:**

$$\text{Exp}(X) = \sum_{v \in \Omega} v \cdot P(X = v)$$

(\textit{Exp}(X) is the center of the distribution of $X$)

- if $\mu = \text{Exp}(X)$ is finite, the **variance** of $X$ is:

$$\text{Var}(X) = \text{Exp}((X - \mu)^2) = \sum_{v \in \Omega} P(X = v).(v - \mu)^2$$

- if $\text{Exp}(X)$ is finite, the **standard deviation** of $X$ is:

$$\text{Std}(X) = \sqrt{\text{Var}(X)}$$

(\textit{Std}(X) is the spread around the center)
Recall: Random variables

Discrete random variable

Theorem (Thm 9.8 of the textbook and more)

Let $X$ and $Y$ be two discrete random variables with finite expected values, and $a \in \mathbb{R}$. Then:

(i) $\text{Exp}(X + Y) = \text{Exp}(X) + \text{Exp}(Y)$, and

(ii) $\text{Exp}(aX) = a\text{Exp}(X)$,

(iii) $\text{Exp}(a + X) = a + \text{Exp}(X)$.

Theorem (Thms 9.9, 9.12 of the textbook)

Let $X$ and $Y$ be two independent discrete random variables with finite expected values. Then:

(iv) $\text{Exp}(X \cdot Y) = \text{Exp}(X)\text{Exp}(Y)$, and

(v) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

(Let $X, Y$ be two independent random variable. What can you say about $\text{Var}(X - Y)$?)
Recall: Random variables

Exercise:
Let \( X \) and \( Y \) be two numerical random variables such that 
\[ \text{Exp}(X) = 1, \text{Var}(X) = 4, \text{Exp}(Y) = 3, \text{Var}(Y) = 1. \]

A. What is the value of \( \text{Exp}(2X + 3) \)?

B. What is the value of \( \text{Var}(2X + 3) \)?

C. What is the value of \( \text{Std}(2X + 3) \)?

D. Consider the equation \( \text{Exp}(2X + 3Y) = 2\text{Exp}(X) + 3\text{Exp}(Y) \). Does this hold (a) always; (b) if \( X \) and \( Y \) are independent, but not necessarily otherwise; (c) not necessarily even if \( X \) and \( Y \) are independent?

E. Consider the equation \( \text{Var}(2X + 3Y) = 4\text{Var}(X) + 9\text{Var}(Y) \). Does this hold (a) always; (b) if \( X \) and \( Y \) are independent, but not necessarily otherwise; (c) not necessarily even if \( X \) and \( Y \) are independent?
Recall: Random variables, discrete distribution

The discrete uniform distribution: \( \Omega \) is finite, let \( \Omega = \{1, \ldots, N\} \)

- \( \forall w \in \Omega, P(X = \omega) = \frac{1}{N} \)
- \( \text{Exp}(X) = \frac{N+1}{2} \)
- \( \text{Var}(X) = \) (exercise)

(example: we roll a fair die)
Recall: Random variables, discrete distribution

The discrete uniform distribution: \( \Omega \) is finite, let \( \Omega = \{1, \ldots, N\} \)

- \( \forall w \in \Omega, P(X = \omega) = \frac{1}{N} \)
- \( \text{Exp}(X) = \frac{N+1}{2} \)
- \( \text{Var}(X) = ? \) (exercise)

The discrete uniform distribution (2): \( \Omega \) is finite, let \( \Omega = \{a, \ldots, b\} \)

- \( P(X = \omega) = \frac{1}{b-a+1} \)
- \( \text{Exp}(X) = \frac{a+b}{2} \)
- \( \text{Var}(X) = ? \) (exercise)

(example: we roll a fair die)
Recall: Random variables, discrete distribution

The Bernoulli distribution: depends on a parameter $p \in [0, 1]$. $\Omega = \{0, 1\}$

- $P(X = 0) = (1 - p)$, $P(X = 1) = p$
- $Exp(X) = p$
- $Var(X) = p(1 - p)$

(example: we flip a weighted coin with weight $p$)
Recall: Random variables, discrete distribution

The Bernoulli distribution: depends on a parameter \( p \in [0, 1] \). \( \Omega = \{0, 1\} \)

- \( P(X = 0) = (1 - p) \), \( P(X = 1) = p \)
- \( \text{Exp}(X) = p \)
- \( \text{Var}(X) = p(1 - p) \)

(example: we flip a weighted coin with weight \( p \))

The binomial distribution: depends on two parameters \( n \) and \( p \). \( \Omega = \{0, 1, ..., n\} \)

- for any \( k \in \Omega \), \( P(Y = k) = C(n, k)p^k(1 - p)^{n-k} \).
- \( \text{Exp}(X) = np \)
- \( \text{Var}(X) = np(1 - p) \)

(example: we flip \( n \) times a weighted coin with weight \( p \))
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

**Recall:** the event of $\Omega$ are the intersections/unions of intervals of $\Omega$. 

Example 6:

Experience: We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.

Random variable: $X$

Domain: $\Omega = [a, b]$.

Distribution: How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

Intuition: What is $P(X \leq b)$? What is $P(X \leq a + b - a)$? What is $P(X \leq a + b - a^2)$? What is $P(X \leq a + 3b - a^4)$? What is $\ldots$ What is $P(X \leq \omega)$? 

$\Rightarrow$ Cumulative distribution function
Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

Recall: the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

Example 6:
Experience: We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.
Random variable: $X$
Domain: $\Omega = [a, b]$.
Distribution: How to characterize $P(X \in A)$, for an event $A$ of $\Omega$
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

Recall: the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

Example 6:
Experience: We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.
Random variable: $X$
Domain: $\Omega = [a, b]$.
Distribution: How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

Intuition:
What is $P(X \leq b)$?
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

**Recall:** the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

**Example 6:**

**Experience:** We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.

**Random variable:** $X$

**Domain:** $\Omega = [a, b]$.

**Distribution:** How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

**Intuition:**

What is $P(X \leq b)$?

$P(X \leq a + \frac{b-a}{2})$?
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

**Recall:** the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

**Example 6:**

**Experience:** We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.

**Random variable:** $X$

**Domain:** $\Omega = [a, b]$.

**Distribution:** How to characterize $P(X \in A)$, for an event $A$ of $\Omega$.

**Intuition:**

- What is $P(X \leq b)$?
- $P(X \leq a + \frac{b-a}{2})$?
- $P(X \leq a + \frac{b-a}{4})$?

![Diagram](image)
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

Recall: the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

Example 6:
Experience: We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.
Random variable: $X$
Domain: $\Omega = [a, b]$.
Distribution: How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

Intuition:
What is $P(X \leq b)$?
$P(X \leq a + \frac{b-a}{2})$?
$P(X \leq a + \frac{b-a}{4})$?
$P(X \leq a + 3\frac{b-a}{4})$?
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

Recall: the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

Example 6:
Experience: We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.

Random variable: $X$

Domain: $\Omega = [a, b]$.

Distribution: How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

Intuition:
What is $P(X \leq b)$?
$P(X \leq a + \frac{b-a}{2})$?
$P(X \leq a + \frac{b-a}{4})$?
$P(X \leq a + 3\frac{b-a}{4})$?
...
3/ Continuous random variables

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

**Recall:** the event of $\Omega$ are the intersections/unions of intervals of $\Omega$.

**Example 6:**

**Experience:** We pick up a number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear.

**Random variable:** $X$

**Domain:** $\Omega = [a, b]$.

**Distribution:** How to characterize $P(X \in A)$, for an event $A$ of $\Omega$

**Intuition:**

What is $P(X \leq b)$?

$P(X \leq a + \frac{b-a}{2})$?

$P(X \leq a + \frac{b-a}{4})$?

$P(X \leq a + 3\frac{b-a}{4})$?

$\cdots$

$P(X \leq \omega)$?

$\Rightarrow$ Cumulative distribution function
Cumulative distribution function: definition

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

**Definition (def 9.15 of the textbook)**

The *cumulative distribution function* for $X$ is a function $c : \Omega \rightarrow [0, 1]$ s.t.:

- $\forall t \in \Omega$, $P(X \leq t) = P(X < t) = c(t)$,
- if $t_1 \leq t_2$, then $c(t_1) \leq c(t_2)$, (*$c$ is monotonically non decreasing*)
- $\forall 0 < v < 1$, there exists $t$ s.t. $c(t) = v$.

Left: $c(t) = \frac{t-a}{b-a}$. Right: $c(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$, $\mu = 80$ and $\sigma = 15$. 
Cumulative distribution function: definition

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$.

Definition (def 9.15 of the textbook)

The *cumulative distribution function* for $X$ is a function $c : \Omega \rightarrow [0, 1]$ s.t.:

- $\forall t \in \Omega, P(X \leq t) = P(X < t) = c(t)$,
- if $t_1 \leq t_2$, then $c(t_1) \leq c(t_2)$, (*$c$ is monotonically non decreasing*)
- $\forall 0 < v < 1$, there exists $t$ s.t. $c(t) = v$.

Let $[u, v]$ be an interval of $\Omega$ and $c$ be the cdf of $X$. Then:

$$P(X \in [u, v]) = c(v) - c(u)$$

Because: $P(X \leq v) = P((X < u) \cup (u \leq X \leq v))$

$$= P(X < u) + P(X \in [u, v])$$
Cumulative distribution function: Example

Let $X$ be a random variable on $\Omega = [a, b]$, with the cdf:

$$c(t) = \frac{t - a}{b - a}$$

Let $A_1$ be the interval $A_1 = [a_1, b_1]$. Then:

$$P(X \in A_1) = \frac{b_1 - a}{b - a} - \frac{a_1 - a}{b - a} = \frac{b_1 - a_1}{b - a}$$

Let $A_2$ be the interval $A_2 = [a_2, b_2]$, s.t. $b_2 - a_2 = b_1 - a_1$. Then:

$$P(X \in A_1) = P(X \in A_2).$$
**Probability density function: definition**

Let $X$ be a continuous variable with values in $\Omega \subseteq \mathbb{R}$, with cdf $c$.

**Definition (def 9.17 of the textbook)**

If $c$ is piecewise differentiable then $c'(t)$ is called the *probability density function* (pdf) of $X$.

We note $\tilde{P}(X = t)$ the value of the pdf of $X$ at $t$. We call it *probability density at $t$*.

Formally:

$$\tilde{P}(X = t) = \lim_{u \to t, \epsilon \to 0^+} \frac{P(u < X < u + \epsilon)}{\epsilon} = \frac{d}{dt} P(X \leq t)$$

($\tilde{P}(X = t)$ is the probability that $X$ lies in an interval around $t$ getting as small as possible, or infinitely small)
Probability density function: Example

Let $X$ be a random variable on $\Omega = [a, b]$, with the cdf:

$$c(t) = \frac{t - a}{b - a}$$

Then:

$$\forall t \in \Omega, \tilde{P}(X = t) = c'(t) = \frac{1}{b - a}$$

the probability density is $1/(b - a)$ at any $t$: each element of $\Omega$ has the same likelihood to appear.
Probability density function: Example

Let $X$ be a random variable on $\Omega = [a, b]$, with the cdf:

$$c(t) = \frac{t - a}{b - a}$$

Then:

$$\forall t \in \Omega, \tilde{P}(X = t) = c'(t) = \frac{1}{b - a}$$

the probability density is $1/(b - a)$ at any $t$: each element of $\Omega$ has the same likelihood to appear.

Example 6: Experience: We pick up randomly a real number in the interval $[a, b]$; we suppose that each element has the same likelihood to appear. Random variable: $X$ takes the value of the number picked up randomly. Domain: $\Omega = [a, b]$, it is an interval of $\mathbb{R}$. Distribution:

cdf: $P(X \leq t) = \frac{t - a}{b - a}$

pdf: $\tilde{P}(X = t) = \frac{1}{b - a}$
Probability density function

Recall: the events of $\Omega \subseteq \mathbb{R}$ are the union or intersections of intervals of $\Omega$

**Theorem (thm 9.18)**

Let $X$ be a random variable on $\Omega$, and $E$ be a event of $\Omega$. Then

$$P(X \in E) = \int_E \tilde{P}(X = t)dt$$

(analogous to the discrete case, but the sums is now an integral)
3.1/ Joint density function and independence

Let $X, Y$ be two continuous random variables on $\Omega_X, \Omega_Y$.

We define the joint probability density of $X, Y$ as

$$\tilde{P}(X = t, Y = u) = \lim_{v \to t, w \to u, \epsilon \to 0^+} P((v \leq X \leq v + \epsilon) \cap (w \leq Y \leq w + \epsilon)) / \epsilon^2$$

where $P$ is the joint distribution of $X, Y$ if this limit exists.

One can extend this definition to a finite set of random variables;

**Theorem (thm 9.25)**

Suppose that the joint probability density of $X$ and $Y$ exists. $X$ and $Y$ are independent if $\forall t \in \Omega_X, \forall u \in \Omega_Y,$

$$\tilde{P}(X = t, Y = u) = \tilde{P}(X = t).\tilde{P}(Y = u)$$
3.2/ Conditional probability

Let $X, Y$ be two continuous random variables on $\Omega_X, \Omega_Y$, such that the joint probability density of $X, Y$ exists.

The \textit{conditional density probability of $X$ knowing $Y$} is defined as

$$\tilde{P}(X = t|Y = u) = \frac{\tilde{P}(X = t, Y = u)}{\tilde{P}(Y = u)}$$
3.3/ Marginal distribution

Let $X, Y$ be two continuous random variables on $\Omega_X, \Omega_Y$, such that the joint probability density of $X, Y$ exists.

One has, for any $t \in \Omega_X$:

$$\tilde{P}(X = t) = \int_{\Omega_Y} \tilde{P}(X = t, Y = u)du$$

and for any $u \in \Omega_Y$

$$\tilde{P}(Y = u) = \int_{\Omega_X} \tilde{P}(X = t, Y = u)dt$$
3.4/ Expected value, variance, standard deviation

Let $X$ be a continuous random variable on $\Omega$.

The expected value, or mean, of $X$ is

$$\text{Exp}(X) = \int_{\Omega} t \tilde{P}(X = t) dt$$
3.4/ Expected value, variance, standard deviation

Let $X$ be a continuous random variable on $\Omega$.

The expected value, or mean, of $X$ is

$$\text{Exp}(X) = \int_{\Omega} t \tilde{P}(X = t) dt$$

it is not necessarily finite!

If $\text{Exp}(X)$ exists (i.e. is finite), the variance of $X$ is

$$\text{Var}(X) = \int_{\Omega} (t - \text{Exp}(X))^2 \tilde{P}(X = t) dt$$

The standard deviation of $X$ is $\sqrt{\text{Var}(X)}$.

In general, the expected value is noted $\mu$ and the standard deviation $\sigma$. 
3.4/ Expected value, variance, standard deviation

Theorem (Thm 9.8’ of the textbook and more)

Let $X$ and $Y$ be two continuous random variables with finite expected values, and $a \in \mathbb{R}$. Then:

(i) $\text{Exp}(X + Y) = \text{Exp}(X) + \text{Exp}(Y)$, and 
(ii) $\text{Exp}(aX) = a\text{Exp}(X)$, 
(iii) $\text{Exp}(a + X) = a + \text{Exp}(X)$.

Theorem (Thms 9.9’, 9.12’ of the textbook)

Let $X$ and $Y$ be two independent continuous random variables with finite expected values. Then:

(iv) $\text{Exp}(X \cdot Y) = \text{Exp}(X) \cdot \text{Exp}(Y)$, and 
(v) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. 
3.4/ Expected value, variance, standard deviation

Example 6: Experience: We pick up randomly a real number in the interval 
\([a, b]\); we suppose that each element has the same likelihood to appear.
Random variable: \(X\) takes the value of the number picked up randomly.
Domain: \(\Omega = [a, b]\), it is an interval of \(\mathbb{R}\).
Distribution:
cdf: \(P(X \leq t) = \frac{t-a}{b-a}\)
pdf: \(\tilde{P}(X = t) = \frac{1}{b-a}\)

\[Exp(X) = \int_a^b t\tilde{P}(X = t)\,dt = \frac{1}{b-a} \int_a^b t\,dt = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.\]
3.4/ Expected value, variance, standard deviation

Example 6: Experience: We pick up randomly a real number in the interval 
$[a, b]$; we suppose that each element has the same likelihood to appear.

Random variable: $X$ takes the value of the number picked up randomly.

Domain: $\Omega = [a, b]$, it is an interval of $\mathbb{R}$.

Distribution:

cdf: $P(X \leq t) = \frac{t-a}{b-a}$
dpdf: $\tilde{P}(X = t) = \frac{1}{b-a}$

$Exp(X) = \int_a^b t \tilde{P}(X = t) \, dt = \frac{1}{b-a} \int_a^b t \, dt = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}$.

$Var(X) = \int_a^b (t - Exp(X))^2 \tilde{P}(X = t) \, dt$

$= \int_a^b t^2 \tilde{P}(X = t) \, dt - 2Exp(X) \int_a^b t \tilde{P}(X = t) \, dt$

$+ Exp(X)^2 \int_a^b \tilde{P}(X = t) \, dt$

$= \frac{b^3-a^3}{3(b-a)} - 2\left(\frac{b+a}{2}\right)^2 + \left(\frac{b+a}{2}\right)^2 \frac{b-a}{b-a} = \ldots = \frac{(b-a)^2}{12}$
3.5/ Continuous distributions

The continuous uniform distribution: $\Omega = [a, b]$

- each $t$ of $\Omega$ has the same density of probability
3.5/ Continuous distributions

The continuous uniform distribution: \( \Omega = [a, b] \)

- each \( t \) of \( \Omega \) has the same density of probability

**Example 6:** Experience: We pick up randomly a real number in the interval \([a, b]\); we suppose that each element has the same likelihood to appear.

**Random variable:** \( X \) takes the value of the number picked up randomly.

**Domain:** \( \Omega = [a, b] \), it is an interval of \( \mathbb{R} \).

**Distribution:**
- cdf: \( P(X \leq t) = \frac{t-a}{b-a} \)
- pdf: \( \tilde{P}(X = t) = \frac{1}{b-a} \)

\[
\begin{align*}
\text{Exp}(X) &= \frac{b+a}{2} \\
\text{Var}(X) &= \frac{(b-a)^2}{12}
\end{align*}
\]
3.5/ Continuous distributions

The continuous uniform distribution: \( \Omega = [a, b] \)

- each \( t \) of \( \Omega \) has the same density of probability

**Example 6:** Experience: We pick up randomly a real number in the interval \([a, b]\); we suppose that each element has the same likelihood to appear.

Random variable: \( X \) takes the value of the number picked up randomly.

Domain: \( \Omega = [a, b] \), it is an interval of \( \mathbb{R} \).

Distribution:
- cdf: \( P(X \leq t) = \frac{t-a}{b-a} \)
- pdf: \( \tilde{P}(X = t) = \frac{1}{b-a} \)

\( \text{Exp}(X) = \frac{b-a}{2} \)

\( \text{Var}(X) = \frac{(b-a)^2}{12} \)
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain $\Omega = \mathbb{R}$.

Defined by two parameters:

- the mean $\mu$
- the standard deviation $\sigma$

If $X$ follows the Gaussian distribution of parameters $\mu$ and $\sigma$, then:

$$\tilde{P}(X = t) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

and the cdf is

$$P(X \leq t) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

By definition, the mean of $X$ is $\mu$ and the standard deviation is $\sigma$. 
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain $\Omega = \mathbb{R}$. Defined by two parameters:

- the mean $\mu$
- the standard deviation $\sigma$

If $X$ follows the Gaussian distribution of parameters $\mu$ and $\sigma$, then:

$$
\tilde{P}(X = t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{\mu,\sigma}(t)
$$

and the cdf is

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{t} N_{\mu,\sigma}(t) dt
$$

By definition, the mean of $X$ is $\mu$ and the standard deviation is $\sigma$. 

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.

Experience:

We pick up randomly a male inhabitant of Gotham and weight him. Random variable: $X$ is the weight of the inhabitant picked up randomly. Domain: $\Omega = \mathbb{R}$. 

$$
\tilde{P}(X = 90) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(90-\mu)^2}{2\sigma^2}} = \int_{-\infty}^{90} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx 
$$

(use Matlab to compute this...)

$$
P(X \leq 90) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{90} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{90} N_{\mu,\sigma}(90) dt 
$$

(How can we compute this???)
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain $\Omega = \mathbb{R}$.
If $X$ follows the Gaussian distribution of parameters $\mu$ and $\sigma$, then:

$$ \tilde{P}(X = t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} $$

and the cdf is

$$ P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx $$

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.
Experience: We pick up randomly a male inhabitant of Gotham and weight him.
Random variable: $X$ is the weight of the inhabitant picked up randomly.
Domain: $\Omega = \mathbb{R}$
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain \( \Omega = \mathbb{R} \).
If \( X \) follows the Gaussian distribution of parameters \( \mu \) and \( \sigma \), then:

\[
\tilde{P}(X = t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}
\]

and the cdf is

\[
P(X \leq t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx
\]

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean \( \mu = 80 \) kg and standard deviation \( \sigma = 15 \) kg.
Experience: We pick up randomly a male inhabitant of Gotham and weight him.
Random variable: \( X \) is the weight of the inhabitant picked up randomly.
Domain: \( \Omega = \mathbb{R} \)

\[
\tilde{P}(X = 90) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(90-\mu)^2}{2\sigma^2}} = 0.023. \text{ (use Matlab to compute this...)}
\]
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain $\Omega = \mathbb{R}$.
If $X$ follows the Gaussian distribution of parameters $\mu$ and $\sigma$, then:

$$\tilde{P}(X = t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

and the cdf is

$$P(X \leq t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.
Experience: We pick up randomly a male inhabitant of Gotham and weight him.
Random variable: $X$ is the weight of the inhabitant picked up randomly.
Domain: $\Omega = \mathbb{R}$

$$\tilde{P}(X = 90) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(90-\mu)^2}{2\sigma^2}} = 0.023. \text{ (use Matlab to compute this...)}$$

$$P(X \leq 90) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{90} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \text{ (how can we compute this???)}$$
3.5/ Continuous distributions

The Gaussian distribution: or normal distribution. Domain $\Omega = \mathbb{R}$.

If $X$ follows the Gaussian distribution of parameters $\mu$ and $\sigma$,
then:

$$\tilde{P}(X = t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t - \mu)^2}{2\sigma^2}} = 0.023.$$  

(use MATLAB to compute this...)

$$P(X \leq 90) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{90} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \, dx \text{ (how can we compute this???)}$$

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.

Experience: We pick up randomly a male inhabitant of Gotham and weight him.

Random variable: $X$ is the weight of the inhabitant picked up randomly.

Domain: $\Omega = \mathbb{R}$

Left: the cdf of $X$, right: the pdf of $X$. 

$\tilde{P}(X = 90) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(90 - \mu)^2}{2\sigma^2}} = 0.023.$ (use MATLAB to compute this...)

$P(X \leq 90) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{90} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \, dx \text{ (how can we compute this???)}$
3.5/ Continuous distributions

The Gaussian distribution:

\[ P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \]  \( \text{(how can we compute this???)} \)
3.5/ Continuous distributions

The Gaussian distribution:
\[ P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \] (how can we compute this???)

**Theorem**

*If* \( X \) *follows* \( N_{\mu,\sigma} \) *then* \( Y = \frac{X-\mu}{\sigma} \) *follows* \( N_{0,1} \).

Then \( P(X \leq t) = P(Y \leq \frac{t-\mu}{\sigma}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t-\mu}{\sigma}} e^{-\frac{x^2}{2}} \, dx \)
3.5/ Continuous distributions

The Gaussian distribution:

\[ P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \]

*(how can we compute this???)*

**Theorem**

*If* \(X\) *follows* \(N_{\mu,\sigma}\) *then* \(Y = \frac{X-\mu}{\sigma}\) *follows* \(N_{0,1}\).

Then \(P(X \leq t) = P\left(Y \leq \frac{t-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t-\mu}{\sigma}} e^{-\frac{x^2}{2}} \, dx\)

If we know how to compute:

\[ E(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} \, dx = \int_{-\infty}^{t} N_{0,1}(x) \, dx \]

we are all set!
3.5/ Continuous distributions

The Gaussian distribution:
\[ P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \] (how can we compute this???)

Theorem

If \( X \) follows \( N_{\mu,\sigma} \) then \( Y = \frac{X-\mu}{\sigma} \) follows \( N_{0,1} \).

Then \( P(X \leq t) = P\left(Y \leq \frac{t-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t-\mu}{\sigma}} e^{-\frac{x^2}{2}} \, dx \)

If we know how to compute:

\[ E(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} \, dx = \int_{-\infty}^{t} N_{0,1}(x) \, dx \]

we are all set!

In Matlab: \( E(t) \) is obtained with \((1/2)*(1+erf(t/sqrt(2)))\).
3.5/ Continuous distributions

Example 7: The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.

Experience: We pick up randomly a male inhabitant of Gotham and weight him.

Random variable: $X$ is the weight of the inhabitant picked up randomly.

Domain: $\Omega = \mathbb{R}$

$$\tilde{P}(X = 90) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(90-\mu)^2}{2\sigma^2}} = 0.023.$$ (use \textit{Matlab} to compute this...)

$$P(X \leq 90) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{90} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$ (how can we compute this???)

If we know how to compute $E(t)$: $P(X \leq t) = E\left(\frac{t-\mu}{\sigma}\right)$

In Matlab: \((1/2)*(1+\text{erf}\left((t-\mu)/(\text{sigma}\times\text{sqrt}(2))\right))\)

For instance, compute $P(85 \leq X \leq 90)$ as:

- $\mu = 80$;
- $\sigma = 15$;
- $$(1/2)*(1+\text{erf}\left((90-\mu)/(\text{sigma}\times\text{sqrt}(2))\right))$$
- $-(1/2)*(1+\text{erf}\left((85-\mu)/(\text{sigma}\times\text{sqrt}(2))\right))$$
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p = 0.54$. (54% of the inhabitants are female).

Let $Z$ be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p = 0.54$. (54% of the inhabitants are female).

Let $Z$ be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

$$\tilde{P}(G = 0 | Z = t) = \frac{\tilde{P}(Z = t | G = 0) \times P(G = 0)}{\tilde{P}(Z = t | G = 0) \times P(G = 0) + \tilde{P}(Z = t | G = 1) \times P(G = 1)}$$
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex \( G \) of the inhabitants of Gotham follows the Bernoulli law of parameter \( p = 0.54 \). (54% of the inhabitants are female).

Let \( Z \) be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

\[
\tilde{P}(G = 0|Z = t) = \frac{\tilde{P}(Z = t|G = 0) \times P(G = 0)}{\tilde{P}(Z = t|G = 0) \times P(G = 0) + \tilde{P}(Z = t|G = 1) \times P(G = 1)}
\]
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex \( G \) of the inhabitants of Gotham follows the Bernoulli law of parameter \( p = 0.54 \). (54% of the inhabitants are female).

Let \( Z \) be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

\[
\tilde{P}(G = 0|Z = t) = \frac{\tilde{P}(Z = t|G = 0) \times 0.46}{\tilde{P}(Z = t|G = 0) \times 0.46 + \tilde{P}(Z = t|G = 1) \times 0.54}
\]
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p = 0.54$. (54% of the inhabitants are female).

Let $Z$ be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

$$\tilde{P}(G = 0|Z = t) = \frac{0.0161 \times 0.46}{0.0161 \times 0.46 + \tilde{P}(Z = t|G = 1) \times 0.54}$$

⇒ the inhabitant is probably a female!
3.6/ Baye’s law on an example:

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p = 0.54$. (54% of the inhabitants are female).

Let $Z$ be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

$$\tilde{P}(G = 0|Z = t) = \frac{0.0161 \times 0.46}{0.0161 \times 0.46 + 0.0242 \times 0.54} \approx 0.3617$$

⇒ the inhabitant is probably a female!
3.6/ Baye’s law on an example: Bayesian classifier

Example 7:
The weight of the male inhabitants of Gotham follows a Gaussian distribution of mean 80 kg and standard deviation 15 kg.

Example 8:
The weight of the female inhabitants of Gotham follows a Gaussian distribution of mean 75 kg and standard deviation 10 kg.

The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p = 0.54$. (54% of the inhabitants are female).

Let $Z$ be the weight of an inhabitant of Gotham picked up randomly. Determine the probability that this inhabitant is a male knowing its weight.

Baye’s law:

$$\tilde{P}(G = 0|Z = t) = \frac{0.0161 * 0.46}{0.0161 * 0.46 + 0.0242 * 0.54} \approx 0.3617$$

$\Rightarrow$ the inhabitant is probably a female!
4/ Statistics
4.1/ Tschebyscheff’s inequality

**Theorem (Thms 9.11 of the textbook)**

Let $X$ be a random variable with mean $\mu$ and standard deviation $\sigma$. Then for any $w > \sigma$, one has:

$$P(|X - \mu| \geq w) \leq \frac{\sigma^2}{w^2}$$

(if $w$ is greater than the standard deviation $\sigma$, the probability that $X$ is at a distance more than $w$ from $\mu$ decreases at least as $\frac{\sigma^2}{w^2}$.)
4.1/ Tschebyscheff’s inequality

**Theorem (Thms 9.11 of the textbook)**

Let $X$ be a random variable with mean $\mu$ and standard deviation $\sigma$. Then for any $w > \sigma$, one has:

$$P(|X - \mu| \geq w) \leq \frac{\sigma^2}{w^2}$$

**Example 7:**
The weight $X$ of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.

- let $w_1 = 20$: the proba. that $X \leq \mu - w_1$ or $\mu + w_1 \leq X$ is less than

  $$\frac{15^2}{20^2} = \frac{225}{400} \sim \frac{1}{2}$$
4.1/ Tschebyscheff’s inequality

Theorem (Thms 9.11 of the textbook)

Let $X$ be a random variable with mean $\mu$ and standard deviation $\sigma$. Then for any $w > \sigma$, one has:

$$P(|X - \mu| \geq w) \leq \frac{\sigma^2}{w^2}$$

Example 7:
The weight $X$ of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = 80$ kg and standard deviation $\sigma = 15$ kg.

- let $w_1 = 20$: the proba. that $X \leq \mu - w_1$ or $\mu + w_1 \leq X$ is less than
  $$\frac{15^2}{20^2} = \frac{225}{400} \sim \frac{1}{2}$$

- let $w_2 = 25$: the proba. that $X \leq \mu - w_2$ or $\mu + w_2 \leq X$ is less than
  $$\frac{15^2}{25^2} = \frac{225}{625} \sim \frac{1}{3}$$
4.2/ Fitting a distribution to a sample

Example 7:
The weight $X$ of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = ?$ kg and standard deviation $\sigma = ?$ kg.

You have a sample of inhabitants of Gotham; for each inhabitant in this sample you know its weight.
How to determine $\mu$ and $\sigma$?
4.2/ Fitting a distribution to a sample

Example 7:
The weight $X$ of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = ?$ kg and standard deviation $\sigma = ?$ kg.

You have a sample of inhabitants of Gotham; for each inhabitant in this sample you know its weight.
How to determine $\mu$ and $\sigma$?

Intuition:
Determine the mean and the standard deviation of the weights in the sample.
4.2/ Fitting a distribution to a sample

Example 7:
The weight $X$ of the male inhabitants of Gotham follows a Gaussian distribution of mean $\mu = ?$ kg and standard deviation $\sigma = ?$ kg.

You have a sample of inhabitants of Gotham; for each inhabitant in this sample you know its weight.

How to determine $\mu$ and $\sigma$?

Intuition:
Determine the mean and the standard deviation of the weights in the sample.

Good news: this intuition is correct.
The problem is known as the Maximum Likelihood Estimation (MLE): see chapter 14 of the textbook for more details.
4.2/ Fitting a distribution to a sample

Example 8:
The sex $G$ of the inhabitants of Gotham follows the Bernoulli law of parameter $p=?$.

You have a sample of inhabitants of Gotham; for each inhabitant in this sample you know its sex.
How to determine $p$?
4.3/ An application of statistics: polls

Problem:
We want to determine the fraction $f$ of a population that prefers $A$ to $B$. 
4.3/ An application of statistics: polls

Problem:
We want to determine the fraction $f$ of a population that prefers $A$ to $B$.

Statistical approach:
$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.
We want to determine:

- the law of $f$
- its mean $\tilde{f}$
- its standard deviation $\sigma_f$
4.3/ An application of statistics: polls

Problem:
We want to determine the fraction \( f \) of a population that prefers \( A \) to \( B \).

Statistical approach:
\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
We want to determine:

- the law of \( f \)
- its mean \( \tilde{f} \)
- its standard deviation \( \sigma_f \)

We poll \( n \) individuals chosen randomly in the population.
Let \( X_1, \ldots, X_n \) be \( n \) independent discrete variables, taking value in domain \( \Omega_X = \{0, 1\} \), following the Bernoulli law of parameter \( \tilde{f} \).

\( X_i \) takes value 1 if the \( i \)-th individual prefers \( A \).
4.3/ An application of statistics: polls

Problem:
We want to determine the fraction \( f \) of a population that prefers \( A \) to \( B \).

Statistical approach:
\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
We want to determine:

- the law of \( f \)
- its mean \( \tilde{f} \)
- its standard deviation \( \sigma_f \)

We poll \( n \) individuals chosen randomly in the population.
Let \( X_1, \ldots, X_n \) be \( n \) independent discrete variables, taking value in domain \( \Omega_X = \{0, 1\} \), following the Bernoulli law of parameter \( \tilde{f} \).

\( X_i \) takes value 1 if the \( i \)-th individual prefers \( A \).
4.3/ An application of statistics: polls

Let $f$ be a random variable taking values in $\Omega_f = \mathbb{R}$. $X_1, \ldots, X_n$ are $n$ independent variables, taking values in $\Omega_X = \{0, 1\}$, following the Bernoulli law of parameter $\tilde{f}$.

Average of $X_1, \ldots, X_n$: We call

$$Y = \frac{X_1 + \ldots + X_n}{n}$$

the average of $X_1, \ldots, X_n$. 

Theorem (Thm 9.13 of the textbook)

Let $X_1, \ldots, X_n$ be $n$ independent variables, each of which has mean $\mu$ and standard deviation $\sigma$. Let $Y = X_1 + \ldots + X_n$ be the average. Then

(a) $\mathbb{E}(Y) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = n \mu$;
(b) $\text{Var}(Y) = \text{Var}(X_1) + \ldots + \text{Var}(X_n) = \frac{\sigma^2}{n}$;
(c) $\text{Std}(Y) = \frac{\sigma}{\sqrt{n}}$.

Theorem (Central Limit Theorem; Thm 9.29 of the textbook)

Let $X_1, \ldots, X_n$ be $n$ independent variables, each of which has mean $\mu$ and standard deviation $\sigma$. Let $Y = X_1 + \ldots + X_n$ be the average. Then, when $n$ grows, the law of $Y$ converges to the Gaussian distribution of mean $\mu$ and standard deviation $\sigma/\sqrt{n}$. 
4.3/ An application of statistics: polls

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.

$X_1, \ldots, X_n$ are $n$ independent variables, taking value in $\Omega_X = \{0, 1\}$, following the bernoulli law of parameter $\tilde{f}$.

Average of $X_1, \ldots, X_n$: We call

$$Y = \frac{X_1 + \ldots + X_n}{n}$$

the average of $X_1, \ldots, X_n$.

→ when $n$ is the size of the population, $Y = f$!!!
→ $Y$ converges to $f$
→ we use $Y$ to estimate $f$
→ what can we say about the mean and standard dev. of $Y$?
4.3/ An application of statistics: polls

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.

$X_1, \ldots, X_n$ are $n$ independent variables, taking value in $\Omega_X = \{0, 1\}$, following the Bernoulli law of parameter $\tilde{f}$.

**Theorem (Thm 9.13 of the textbook)**

Let $X_1, \ldots, X_n$ be $n$ independent variables, each of which has mean $\mu$ and standard deviation $\sigma$. Let $Y = \frac{X_1 + \ldots + X_n}{n}$ be the average. Then

(a) $\text{Exp}(Y) = \frac{\text{Exp}(X_1) + \ldots + \text{Exp}(X_n)}{n} = \frac{n\mu}{n} = \mu$

(b) $\text{Var}(Y) = \text{Exp}(\left(\frac{X_1 + \ldots + X_n}{n} - \text{Exp}(Y)\right)^2) = \ldots = \frac{\sigma^2}{n}$

(c) $\text{Std}(Y) = \frac{\sigma}{\sqrt{n}}$
4.3/ An application of statistics: polls

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.

$X_1, \ldots, X_n$ are $n$ independent variables, taking value in $\Omega_X = \{0, 1\}$, following the bernoulli law of parameter $\tilde{f}$.

**Average of $X_1, \ldots, X_n$:** We call

$$Y = \frac{X_1 + \ldots + X_n}{n}$$

the average of $X_1, \ldots, X_n$.

$\rightarrow$ when $n$ is the size of the population, $Y = f$!!

$\rightarrow$ $Y$ converges to $f$

$\rightarrow$ we use $Y$ to estimate $f$

$\rightarrow$ what can we say about the mean and standard dev. of $Y$?

$\rightarrow$ $\text{Exp}(Y) = \tilde{f}$, $\text{Std}(Y) = \sqrt{\frac{\tilde{f}(1-\tilde{f})}{n}}$

$\rightarrow$ what can we say about the law of $f$?
4.3/ An application of statistics: polls

\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
\( X_1, \ldots, X_n \) are \( n \) independent variables, taking value in \( \Omega_X = \{0, 1\} \), following the bernoulli law of parameter \( \tilde{f} \).

**Theorem (central limit theorem; Thm 9.29 of the textbook)**

Let \( X_1, \ldots, X_n \) be \( n \) independent variables, each of which has mean \( \mu \) and standard deviation \( \sigma \). Let \( Y = \frac{X_1 + \ldots + X_n}{n} \) be the average. Thus the law of \( Y \), when \( n \) grows, converges to the gaussian distribution of mean \( \mu \) and standard deviation \( \frac{\sigma}{\sqrt{n}} \).

→ when \( n \) is the size of the population, \( Y = f!!! \)
→ \( Y \) converges to \( f \)
→ we use \( Y \) to estimate \( f \)
→ what can we say about the mean and standard dev. of \( Y \)?
→ \( \text{Exp}(Y) = \tilde{f}, \text{Std}(Y) = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}} \)
→ what can we say about the law of \( f \)?
4.3/ An application of statistics: polls

\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
\( X_1, \ldots, X_n \) are \( n \) independent variables, taking value in \( \Omega_X = \{0, 1\} \), following the Bernoulli law of parameter \( \tilde{f} \).

\( \rightarrow \) estimate \( \tilde{f} \) on the sample.
4.3/ An application of statistics: polls

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.

$X_1, \ldots, X_n$ are $n$ independent variables, taking value in $\Omega_X = \{0, 1\}$, following the Bernoulli law of parameter $\tilde{f}$.

→ estimate $\tilde{f}$ on the sample.

First conclusion: if $n$ is sufficiently big:

- knowing $\tilde{f}$ observed on the sample, $f$ follows a Gaussian law of mean $\tilde{f}$ and standard deviation $\sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}$, i.e.:

$$
\tilde{P}(f = t|\text{the poll}) = N_{\tilde{f}, \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}}(t)
$$

Then knowing $\tilde{f}$ observed on the sample, the most probable value for $f$ is $\text{Exp}(f) = \tilde{f}$.
4.3/ An application of statistics: polls

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.

$X_1, \ldots, X_n$ are $n$ independent variables, taking value in $\Omega_X = \{0, 1\}$, following the Bernoulli law of parameter $\tilde{\theta}$.

→ estimate $\tilde{\theta}$ on the sample.

First conclusion: if $n$ is sufficiently big:

knowing $\tilde{\theta}$ observed on the sample, $f$ follows a Gaussian law of mean $\tilde{\theta}$ and standard deviation $\sigma_f = \frac{\sqrt{\tilde{\theta}(1-\tilde{\theta})}}{\sqrt{n}}$, i.e.:

$$\tilde{P}(f = t | \text{the poll}) = N_{\tilde{\theta}, \frac{\sqrt{\tilde{\theta}(1-\tilde{\theta})}}{\sqrt{n}}}(t)$$

Then knowing $\tilde{\theta}$ observed on the sample, the most probable value for $f$ is $\text{Exp}(f) = \tilde{\theta}$.

Second “conclusion”: Can we get a probabilistic estimate of the error?
4.4/ Confidence intervals

\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
Its probability distribution knowing the poll follows a gaussian law of mean \( \tilde{f} \) and standard deviation \( \sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}} \)

**Definition (def 9.15 of the textbook)**

Let \([t_1, t_2]\) be an interval in \( \Omega_f \), such that:

\[ \tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) \geq 95\% \]

\([t_1, t_2]\) is called a confidence interval at the 95% level.

*(it means: knowing the poll, the probability that the true value for \( f \) is between \( t_1 \) and \( t_2 \) is more than 95%.)

*(it does not mean: the probability that the true value for \( f \) is between \( t_1 \) and \( t_2 \) is more than 95%.)
4.4/ Confidence intervals

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.
Its probability distribution knowing the poll follows a gaussian law of mean $\tilde{f}$ and standard deviation $\sigma_f = \sqrt{\frac{\tilde{f}(1-\tilde{f})}{\sqrt{n}}}$

Problem: be given $p$ find $t_1, t_2$ such that

$$\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p$$

1. let $t_1 = \tilde{f} - q\sigma_f$ and $t_2 = \tilde{f} + q\sigma_f$
### 4.4/ Confidence intervals

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$. Its probability distribution knowing the poll follows a gaussian law of mean $\tilde{f}$ and standard deviation $\sigma_f = \sqrt{\frac{\tilde{f}(1-\tilde{f})}{n}}$

**Problem:** be given $p$ find $t_1, t_2$ such that

$$\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p$$

1. let $t_1 = \tilde{f} - q\sigma_f$ and $t_2 = \tilde{f} + q\sigma_f$
2. then $\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = P(-q \leq \frac{f-\tilde{f}}{\sigma_f} \leq q)$
   $$= E(q) - E(-q) = 2E(q) - 1$$

Recall: $E(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{t} N_{0,1}(x)dx$

In Matlab: $E(t) = (1/2)*(1+\text{erf}(t/\text{sqrt}(2)))$
4.4/ Confidence intervals

\( f \) is a random variable taking values in \( \Omega_f = \mathbb{R} \).
Its probability distribution \textit{knowing the poll} follows a gaussian law of mean \( \tilde{f} \) and standard deviation \( \sigma_f = \sqrt{\frac{\tilde{f}(1-\tilde{f})}{n}} \)

Problem: be given \( p \) find \( t_1, t_2 \) such that

\[
\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p
\]

1. let \( t_1 = \tilde{f} - q\sigma_f \) and \( t_2 = \tilde{f} + q\sigma_f \)
2. then \( \tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = P(-q \leq \frac{f-\tilde{f}}{\sigma_f} \leq q) = E(q) - E(-q) = 2E(q) - 1 \)
3. \( \tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p \iff E(q) = \frac{p+1}{2} \iff \text{erf}(q/sqrt(2)) = p \)

Recall: \( E(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{t} N_{0,1}(x) dx \)
In Matlab: \( E(t) = (1/2)*(1+\text{erf}(t/sqrt(2))) \)
4.4/ Confidence intervals

$f$ is a random variable taking values in $\Omega_f = \mathbb{R}$.
Its probability distribution knowing the poll follows a gaussian law of mean $\tilde{f}$ and standard deviation $\sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}$

Problem: be given $p$ find $t_1, t_2$ such that

$$\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p$$

1. let $t_1 = \tilde{f} - q\sigma_f$ and $t_2 = \tilde{f} + q\sigma_f$
2. then $\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = P(-q \leq \frac{f - \tilde{f}}{\sigma_f} \leq q)$
   $$= E(q) - E(-q) = 2E(q) - 1$$
3. $\tilde{P}(t_1 \leq f \leq t_2 | \text{the poll}) = p$ $\iff$ $E(q) = \frac{p+1}{2}$
   $$\iff \text{erf}(q/\sqrt{2}) = p$$
4. $E$ is has $E^{-1}$ as inverse: $q = E^{-1}(\frac{p+1}{2})$, $t_1 = \tilde{f} - q\sigma_f$ and $t_2 = \tilde{f} + q\sigma_f$.
In Matlab: $q = \text{erfinv}(p) \times \sqrt{2}$
4.4/ Confidence intervals: Example in Matlab

\[
\begin{align*}
n &= 1000; \text{ }\% \text{ the size of the sample} \\
\mu &= 0.574; \text{ }\% \text{ the mean observed on the poll} \\
\sigma &= \sqrt{\mu(1-\mu)/n}; \text{ }\% \text{ the std of } f \text{ knowing the poll} \\
\end{align*}
\]

\% we want a confidence interval at the 95\% level
\[
\begin{align*}
p &= 0.95; \\
q &= \text{erfinv}(p)\sqrt{2}; \\
t1 &= \mu - \sigma q; \\
t2 &= \mu + \sigma q; \\
\text{disp}(['\text{confidence interval at the }', \text{num2str}(p), ' \text{ level: [', \text{num2str}(t1), ', ', \text{num2str}(t2), ']'}]);
\end{align*}
\]

\% we want a confidence interval at the 99\% level
\[
\begin{align*}
p &= 0.99; \\
q &= \text{erfinv}(p)\sqrt{2}; \\
t199 &= \mu - \sigma q; \\
t299 &= \mu + \sigma q; \\
\text{disp}(['\text{confidence interval at the }', \text{num2str}(p), ' \text{ level: [', \text{num2str}(t199), ', ', \text{num2str}(t299), ']'}]);
\end{align*}
\]

\% we want to reduce the size of the 95\% level:
\[
\begin{align*}
factor &= 2*2; \\
n &= n*factor; \\
\sigma &= \sqrt{\mu(1-\mu)/n}; \text{ }\% \text{ the std of } f \text{ knowing the poll} \\
p &= 0.95; \\
q &= \text{erfinv}(p)\sqrt{2}; \\
t1r &= \mu - \sigma q; \\
t2r &= \mu + \sigma q; \\
\text{disp}(['\text{confidence interval at the }', \text{num2str}(p), ' \text{ level: [', \text{num2str}(t1r), ', ', \text{num2str}(t2r), ']'}]); \\
\text{ratio} &= (t2-t1)/(t2r-t1r); \\
\text{disp}(['\text{ratio old/new: }', \text{num2str(ratio)} ]); \\
\text{disp}(['\text{sqrt of factor: }', \text{num2str(sqrt(factor)) }]);
\end{align*}
\]
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area $a$ of the set $A$ defined as:

$$\left\{ \begin{array}{l}
0 \leq X \leq 10, 0 \leq Y \leq 10 \\
-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1
\end{array} \right. \quad (1)$$

Statistical approach: let $f$ be a random variable with mean $\tilde{f}$ and std. $\sigma_f$ $\tilde{f}$ is the fraction of points in $[0, 10] \times [0, 10]$ that are in $A$.
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area \( a \) of the set \( \mathcal{A} \) defined as:

\[
\begin{cases}
0 \leq X \leq 10, 0 \leq Y \leq 10 \\
-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1
\end{cases}
\]  

(1)

Statistical approach: let \( f \) be a random variable with mean \( \tilde{f} \) and std. \( \sigma_f \)

\( \tilde{f} \) is the fraction of points in \([0, 10] \times [0, 10] \) that are in \( \mathcal{A} \)

1. generate \( n \) points uniformly in \( \mathcal{B} = \{0 \leq X \leq 10, 0 \leq Y \leq 10\} \)

for each point, test the inequalities \(-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1\)

\( n \) discrete random variable \( X_1, \ldots, X_n \), Bernoulli law of parameter \( \tilde{f} \)
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area $a$ of the set $A$ defined as:

$$
\begin{cases}
0 \leq X \leq 10, 0 \leq Y \leq 10 \\
-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1
\end{cases}
$$

(1)

Statistical approach: let $f$ be a random variable with mean $\tilde{f}$ and std. $\sigma_f$

$\tilde{f}$ is the fraction of points in $[0, 10] \times [0, 10]$ that are in $A$

1. generate $n$ points uniformly in $B = \{0 \leq X \leq 10, 0 \leq Y \leq 10\}$
   for each point, test the inequalities $-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1$

2. $\tilde{f}$ is the ratio of the num. of points that are in $A$ on $n$
   \[
   \frac{X_1 + \ldots + X_n}{n}
   \]
   converges to Gauss. of mean $\tilde{f}$ and std. $\sigma_f = \sqrt{\frac{\tilde{f}(1-\tilde{f})}{n}}$
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area $a$ of the set $A$ defined as:

$$
\begin{align*}
0 & \leq X \leq 10, 0 \leq Y \leq 10 \\
-0.1 & \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1
\end{align*}
$$

Statistical approach: let $f$ be a random variable with mean $\tilde{f}$ and std. $\sigma_f$.

$\tilde{f}$ is the fraction of points in $[0, 10] \times [0, 10]$ that are in $A$.

1. generate $n$ points uniformly in $B = \{0 \leq X \leq 10, 0 \leq Y \leq 10\}$
   for each point, test the inequalities $-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1$
   $n$ discrete random variable $X_1, \ldots, X_n$, Bernoulli law of parameter $\tilde{f}$

2. $\tilde{f}$ is the ratio of the num. of points that are in $A$ on $n$
   $\frac{X_1 + \ldots + X_n}{n}$ converges to Gauss. of mean $\tilde{f}$ and std. $\sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}$

3. let $b$ be the area of $B$; we estimate $a$ as $\tilde{f} b$
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area $a$ of the set $\mathcal{A}$ defined as:

$$\left\{ \begin{array}{lcr} 0 \leq X \leq 10, 0 \leq Y \leq 10 \\ -0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1 \end{array} \right. \quad (1)$$

Statistical approach: let $f$ be a random variable with mean $\tilde{f}$ and std. $\sigma_f$

$\tilde{f}$ is the fraction of points in $[0, 10] \times [0, 10]$ that are in $\mathcal{A}$

1. generate $n$ points uniformly in $B = \{0 \leq X \leq 10, 0 \leq Y \leq 10\}$
   for each point, test the inequalities $-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1$

2. $\tilde{f}$ is the ratio of the num. of points that are in $\mathcal{A}$ on $n$

   $\frac{X_1 + \ldots + X_n}{n}$ converges to Gauss. of mean $\tilde{f}$ and std. $\sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}$

3. let $b$ be the area of $B$; we estimate $a$ as $\tilde{f} b$

4. a confidence interval at the $p$ level for $f$ is $[t_1, t_2] = [\tilde{f} - \sigma_f E^{-1}\left(\frac{p+1}{2}\right), \tilde{f} + \sigma_f E^{-1}\left(\frac{p+1}{2}\right)]$
4.5/ An application of statistics: Monte Carlo methods

Problem: estimate the area $a$ of the set $A$ defined as:

$$
\begin{cases}
0 \leq X \leq 10, 0 \leq Y \leq 10 \\
-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1
\end{cases}
$$

(1)

Statistical approach: let $f$ be a random variable with mean $\tilde{f}$ and std. $\sigma_f$

$\tilde{f}$ is the fraction of points in $[0, 10] \times [0, 10]$ that are in $A$

1. generate $n$ points uniformly in $B = \{0 \leq X \leq 10, 0 \leq Y \leq 10\}$
   for each point, test the inequalities $-0.1 \leq \cos(\pi X) \cdot \cos(\pi Y) \leq 0.1$
   $n$ discrete random variable $X_1, \ldots, X_n$, Bernoulli law of parameter $\tilde{f}$

2. $\tilde{f}$ is the ratio of the num. of points that are in $A$ on $n$
   $\frac{X_1+\ldots+X_n}{n}$ converges to Gauss. of mean $\tilde{f}$ and std. $\sigma_f = \frac{\sqrt{\tilde{f}(1-\tilde{f})}}{\sqrt{n}}$

3. let $b$ be the area of $B$; we estimate $a$ as $\tilde{f}b$

4. a confidence interval at the $p$ level for $f$ is $[t_1, t_2] = [\tilde{f} - \sigma_f E^{-1}(\frac{p+1}{2}), \tilde{f} + \sigma_f E^{-1}(\frac{p+1}{2})]$

5. a confidence interval at the $p$ level for $a$ is $[bt_1, bt_2]$
4.5/ An application of statistics: Monte Carlo methods (2)

Problem: estimate the integral \( a = \int_0^8 \sin(1 + \ln(x))\,dx \)

- \( a \) is the *signed* area under the graph of \( g(x) = \sin(1 + \ln(x)) \)
- for any \( x \in [0, 8] \), \(-1 \leq g(x) \leq 1\)
4.5/ An application of statistics: Monte Carlo methods (2)

Problem: estimate the integral \( a = \int_{0}^{8} \sin(1 + \ln(x))dx \)

- \( a \) is the signed area under the graph of \( g(x) = \sin(1 + \ln(x)) \)
- for any \( x \in [0, 8], -1 \leq g(x) \leq 1 \)

Statistical approach: let \( f \) be a random variable with mean \( \tilde{f} \) and std. \( \sigma_f \)
\( \tilde{f} \) is the fraction of points in \([0, 8] \times [-1, 1]\) that are in \( A \)

1. generate \( n \) points uniformly in \( B = [0, 8] \times [-1, 1] \)
   for each point, test if it is between the \( x \)-axis and the graph
4.5/ An application of statistics: Monte Carlo methods (2)

Problem: estimate the integral \( a = \int_{0}^{8} \sin(1 + \ln(x))\,dx \)

- \( a \) is the *signed* area under the graph of \( g(x) = \sin(1 + \ln(x)) \)
- for any \( x \in [0, 8], -1 \leq g(x) \leq 1 \)

Statistical approach: let \( f \) be a random variable with mean \( \tilde{f} \) and std. \( \sigma_f \)
\( \tilde{f} \) is the fraction of points in \( [0, 8] \times [-1, 1] \) that are in \( A \)

1. generate \( n \) points uniformly in \( B = [0, 8] \times [-1, 1] \)
   for each point, test if it is between the \( x \)-axis and the graph
2. ...