Mathematical Techniques for Computer Science Applications

remi.imbach@nyu.edu

webpage:
https://cs.nyu.edu/courses/spring19/CSCI-GA.1180-001/
Course 10 - Continuous Random Variables

I Random variables
   1 Example
   2 Definitions
   3 Density of probability of a continuous r.v.
   4 Expected value, variance, standard deviation

II Two continuous distributions

III Joint distribution
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b] \) (it is infinite and not countable)
**Random number**

We pick up a number in the interval \([a, b]\), where \(a < b\).

We suppose that each number appears with the same “likelihood”.

**Outcomes:** \(\Omega = [a, b]\)  (it is infinite and not countable)

**Events:** how to construct a set of events (i.e. a \(\sigma\)-algebra) \(\mathcal{U}\) on \(\Omega\)?
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Outcomes: $\Omega = [a, b]$ (it is infinite and not countable)

Events: how to construct a set of events (i.e. a $\sigma$-algebra) $\mathcal{U}$ on $\Omega$?

Probability function: how to define a probability function on $\Omega, \mathcal{U}$?

Recall: It must satisfy

P1. for any event $A \in \mathcal{U}$, $0 \leq P(A) \leq 1$

P2. for $k \in \mathbb{N} \cup \{+\infty\}$, if $A_1, \ldots, A_k$ are pairwise mutually exclusive events in $\mathcal{U}$,

$$P\left( \bigcup_{1 \leq i \leq k} A_i \right) = \sum_{1 \leq i \leq k} P(A_i)$$

P3. $P(\Omega) = 1$

→ the probability of any element in $\Omega$ should be 0
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

**Outcomes:** \(\Omega = [a, b]\) (it is infinite and not countable)

**Events:** how to construct a set of events (i.e. a \(\sigma\)-algebra) \(\mathcal{U}\) on \(\Omega\)?

**Probability function:** how to define a probability function on \(\Omega, \mathcal{U}\)?

\[ \rightarrow \text{the probability of any element in } \Omega \text{ should be } 0 \]

\[ \rightarrow \text{suppose } \mathcal{U} \text{ contains the intervals of } [a, b]; \]

Consider the interval \([l, u]\) where \(a \leq l < u \leq b\):

\[ P([l, u]) = \frac{l - u}{b - a} = \frac{\text{width of } [l, u]}{\text{width of } [a, b]} \]

(define in the same way \(P([l, u]), P([l, u[]), P([l, u[]))\).
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\)  (it is infinite and not countable)

Events: how to construct a set of events (i.e. a \(\sigma\)-algebra) \(\mathcal{U}\) on \(\Omega\)?

Probability function: how to define a probability function on \(\Omega, \mathcal{U}\)?
→ the probability of any element in \(\Omega\) should be 0

→ suppose \(\mathcal{U}\) contains the intervals of \([a, b]\);

Consider the interval \([l, u]\) where \(a \leq l < u \leq b\):

\[
P([l, u]) = \frac{l - u}{b - a} = \frac{\text{width of } [l, u]}{\text{width of } [a, b]}
\]

(define in the same way \(P([l, u]), P([l, u]), P([l, u])\)).

→ we construct \(\mathcal{U}\) and \(P\) on intervals
\(\sigma\)-algebras over \([a, b]\): sets of intervals

Recall: a \(\sigma\)-algebra \(\mathcal{U}\) over \(\Omega\) is a family of sets satisfying:

- C1. \(\Omega\) is in \(\mathcal{U}\)
- C2. if \(A\) is in \(\mathcal{U}\), then \(A^c\) is in \(\mathcal{U}\),
- C3. for \(k \in \mathbb{N} \cup \{+\infty\}\), if \(A_1, \ldots, A_k\) are in \(\mathcal{U}\) then \(\bigcup_{1 \leq i \leq k} A_k\) is in \(\mathcal{U}\)

Definition: The \textit{Borel set} of \([a, b]\) is the smallest set that

- contains all the intervals \([l, u]\) where \(a \leq l < u \leq b\)
- satisfies C1, C2, C3.
σ-algebras over \([a, b]\): sets of intervals

Recall: a σ-algebra \(\mathcal{U}\) over \(\Omega\) is a family of sets satisfying:

C1. \(\Omega\) is in \(\mathcal{U}\)

C2. if \(A\) is in \(\mathcal{U}\), then \(A^c \in \mathcal{U}\),

C3. for \(k \in \mathbb{N} \cup \{+\infty\}\), if \(A_1, \ldots, A_k\) are in \(\mathcal{U}\) then \(\bigcup_{1 \leq i \leq k} A_k\) is in \(\mathcal{U}\)

Definition: The \textit{Borel set} of \([a, b]\) is the smallest set that

- contains all the intervals \([l, u]\) where \(a \leq l < u \leq b\)

- satisfies C1, C2, C3.

Remark: The \textit{Borel set} of \([a, b]\) contains \(\emptyset\), \([a, b]\), and all the reunions and intersections of intervals of \([a, b]\).
**σ-algebras over \( \mathbb{R} \): sets of intervals**

**Recall:** a σ-algebra \( \mathcal{U} \) over \( \Omega \) is a family of sets satisfying:

C1. \( \Omega \) is in \( \mathcal{U} \)

C2. if \( A \) is in \( \mathcal{U} \), then \( A^c \in \mathcal{U} \),

C3. for \( k \in \mathbb{N} \cup \{+\infty\} \), if \( A_1, \ldots, A_k \) are in \( \mathcal{U} \) then \( \bigcup_{1 \leq i \leq k} A_k \) is in \( \mathcal{U} \)

**Definition:** The *Borel set* of \( \mathbb{R} \) is the smallest set that

- contains all the intervals \([l, u]\) where \( l < u \)
- satisfies C1, C2, C3.
σ-algebras over $\mathbb{R}$: sets of intervals

Recall: a $\sigma$-algebra $\mathcal{U}$ over $\Omega$ is a family of sets satisfying:

C1. $\Omega$ is in $\mathcal{U}$

C2. if $A$ is in $\mathcal{U}$, then $A^c \in \mathcal{U}$,

C3. for $k \in \mathbb{N} \cup \{+\infty\}$, if $A_1, \ldots, A_k$ are in $\mathcal{U}$ then $\bigcup_{1 \leq i \leq k} A_k$ is in $\mathcal{U}$

Definition: The Borel set of $\mathbb{R}$ is the smallest set that

- contains all the intervals $[l, u]$ where $l < u$

- satisfies C1, C2, C3.

Remark: The Borel set of $\mathbb{R}$ contains $\emptyset$, $\mathbb{R}$, and all the reunions and intersections of intervals.

Examples:

- $[l, u], [l, u[, ]l, u], ]l, u[\text{ where } l, u \in \mathbb{R}$

- $([-1, 2] \cup [3, 3.3[ \cup \ldots) \cap ([0, 3.2] \cup \ldots)$
**σ-algebras over \( \mathbb{R} \): sets of intervals**

**Recall:** a \( \sigma \)-algebra \( \mathcal{U} \) over \( \Omega \) is a family of sets satisfying:

C1. \( \Omega \) is in \( \mathcal{U} \)

C2. if \( A \) is in \( \mathcal{U} \), then \( A^c \in \mathcal{U} \),

C3. for \( k \in \mathbb{N} \cup \{+\infty\} \), if \( A_1, \ldots, A_k \) are in \( \mathcal{U} \) then \( \bigcup_{1 \leq i \leq k} A_k \) is in \( \mathcal{U} \)

**Definition:** The *Borel set* of \( \mathbb{R} \) is the smallest set that

- contains all the intervals \([l, u]\) where \( l < u \)
- satisfies C1, C2, C3.

**Remark:** The *Borel set* of \( \mathbb{R} \) contains \( \emptyset \), \( \mathbb{R} \), and all the reunions and intersections of intervals.

**Remark:** (for later) it is possible to compute integrals of functions over such subsets of \( \mathbb{R} \)
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\)  (it is infinite and not countable)

Events:  \(\mathcal{U}:\) the borel set over \([a, b]\)

Probability function:  \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b-a}\)

\(\rightarrow P\) is a legitimate probability function over \(\Omega, \mathcal{U}\)
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Outcomes: $\Omega = [a, b]$ (it is infinite and not countable)

Events: $\mathcal{U}$: the borel set over $[a, b]$

Probability function: $\forall I \in \mathcal{U}, \ P(I) = \frac{\text{width of } I}{b-a}$

Random variable:

$$X : \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto \omega$$

$(X(\Omega)$ is infinite, not countable)$

$P(X = t) = P(\{\omega | X(\omega) = t\}) = 0$

$P(X \in I) = P(\{\omega | X(\omega) \in I\}) = \frac{\text{width of } I}{b-a}$
Course 10 - Continuous Random Variables

I Random variables
   1 Example
   2 Definitions
   3 Density of probability of a continuous r.v.
   4 Expected value, variance, standard deviation

II Two continuous distributions

III Joint distribution
Random Variables

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** A function \(X : \Omega \to \mathbb{R}\) is a *random variable (r.v.)* on \((\Omega, \mathcal{U}, P)\) if

\[
\forall I \text{ interval of } \mathbb{R}, \{\omega \in \Omega | X(\omega) \in I\} \in \mathcal{U}
\]

We call *domain* of \(X\) the set \(X(\Omega)\): we suppose \(X(\Omega) \subseteq \mathbb{R}\).
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Outcomes: $\Omega = [a, b]$ (it is infinite and not countable)

Events: $\mathcal{U}$: the borel set over $[a, b]$

Probability function: $\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b-a}$

Random variable:

$$X : \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto \omega$$

$(X(\Omega)$ is infinite, not countable)$

$$P(X = t) = P(\{\omega | X(\omega) = t\}) = 0$$

$$P(X \in I) = P(\{\omega | X(\omega) \in I\}) = \frac{\text{width of } I}{b-a}$$

Let $I$ be an interval of $\mathbb{R}$: $\{\omega | X(\omega) \in I\} = I \in \mathcal{U}$
Random Variables

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** A function \(X : \Omega \rightarrow \mathbb{R}\) is a *random variable (r.v.)* on \((\Omega, \mathcal{U}, P)\) if

\[
\forall I \text{ interval of } \mathbb{R}, \{\omega \in \Omega | X(\omega) \in I\} \in \mathcal{U} \quad \leftarrow \text{ measurability condition}
\]

We call *domain* of \(X\) the set \(X(\Omega)\): we suppose \(X(\Omega) \subseteq \mathbb{R}\).

**Remark:** we will never try to verify that a given \(X\) is a r.v. We will assume it.
Random Variables

Let \( (\Omega, \mathcal{U}, P) \) be a probability space.

**Definition:** A function \( X : \Omega \rightarrow \mathbb{R} \) is a *random variable (r.v.)* on \( (\Omega, \mathcal{U}, P) \) if

\[
\forall I \text{ interval of } \mathbb{R}, \{ \omega \in \Omega | X(\omega) \in I \} \in \mathcal{U}
\]

We call *domain* of \( X \) the set \( X(\Omega) \): we suppose \( X(\Omega) \subseteq \mathbb{R} \).

**Remark:** we will never try to verify that a given \( X \) is a r.v. We will assume it.

**Remark:** when \( X(\Omega) \) is *finite or countably infinite*, \( X \) is a *discrete r.v.*

an other kind of r.v. are the *continuous r.v.*, that will be defined soon.
Random Variables

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** A function \(X : \Omega \rightarrow \mathbb{R}\) is a *random variable (r.v.)* on \((\Omega, \mathcal{U}, P)\) if

\[
\forall I \text{ interval of } \mathbb{R}, \{\omega \in \Omega | X(\omega) \in I\} \in \mathcal{U}
\]

We call *domain* of \(X\) the set \(X(\Omega)\): we suppose \(X(\Omega) \subseteq \mathbb{R}\).

**Remark:** we will never try to verify that a given \(X\) is a r.v. We will assume it.

**Remark:** when \(X(\Omega)\) is *finite or countably infinite*, \(X\) is a *discrete r.v.*

an other kind of r.v. are the *continuous r.v.*, that will be defined soon.

**Notation:** Let \(t \in \mathbb{R}\) and \(I\) be an interval of \(\mathbb{R}\). We note

- \([X = t]\) the event \(\{\omega | X(\omega) = t\}\) and \(P(X = t)\) for \(P([X = t])\)
- \([X \in I]\) the event \(\{\omega | X(\omega) \in I\}\) and \(P(X \in I)\) for \(P([X \in I])\)
- \([X \leq t]\) the event \(\{\omega | X(\omega) \leq t\}\) and \(P(X \leq t)\) for \(P([X \leq t])\)
- \([X < t], [X \geq t], [X > t]\)
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\) (it is infinite and not countable)

Events: \(\mathcal{U}\) : the borel set over \([a, b]\)

Probability function: \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b-a}\)

Random variable:

\[
X : \Omega \rightarrow \mathbb{R} \\
\omega \mapsto \omega
\]

\((X(\Omega)\) is infinite, not countable) \[P(X = t) = P(\{\omega | X(\omega) = t\}) = 0\]

\[P(X \in I) = P(\{\omega | X(\omega) \in I\}) = \frac{\text{width of } I}{b-a}\]

the notion of probability distribution makes no sens here!
Cumulative distribution function

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$, and consider the function:

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto P(X \leq t)$$

$F_X$ is the *Cumulative Distribution Function (CDF)* of $X$. 

Example: We roll a fair die. $X$ takes the value of the number obtained. The CDF of $X$ is:
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\)  (it is infinite and not countable)

Events: \(\mathcal{U}\) : the borel set over \([a, b]\)

Probability function: \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b - a}\)

Random variable:
\[
X : \Omega \rightarrow \mathbb{R} \\
\omega \mapsto \omega
\]

Cumulative distribution function:
What is \(P(X \leq t)\)?
\(P(X \leq a + \frac{b-a}{2})\)?
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\)  (it is infinite and not countable)

Events:  \(\mathcal{U}\) : the borel set over \([a, b]\)

Probability function: \[\forall I \in \mathcal{U}, \ P(I) = \frac{\text{width of } I}{b-a}\]

Random variable:
\[X : \Omega \rightarrow \mathbb{R} \quad \omega \mapsto \omega\]

Cumulative distribution function:
What is \(P(X \leq t)\)?
\[P(X \leq a + \frac{b-a}{2})?\]
\[P(X \leq a + \frac{b-a}{4})?\]
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\) (it is infinite and not countable)

Events: \(\mathcal{U}\) : the borel set over \([a, b]\)

Probability function: \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b-a}\)

Random variable:
\[
X : \Omega \rightarrow \mathbb{R},
\omega \mapsto \omega
\]

Cumulative distribution function:
What is \(P(X \leq t)\)?
\[
P(X \leq a + \frac{b-a}{2})? \\
P(X \leq a + \frac{b-a}{4})? \\
P(X \leq a + 3\frac{b-a}{4})?
\]
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\). We suppose that each number appears with the same “likelihood”.

**Outcomes:** \(\Omega = [a, b]\) (it is infinite and not countable)

**Events:** \(\mathcal{U}\) : the borel set over \([a, b]\)

**Probability function:** \(\forall I \in \mathcal{U}, \ P(I) = \frac{\text{width of } I}{b-a}\)

**Random variable:**

\[X : \Omega \rightarrow \mathbb{R} \]
\[\omega \mapsto \omega\]

**Cumulative distribution function:**

What is \(P(X \leq t)\)?

- \(P(X \leq a + \frac{b-a}{2})\)?
- \(P(X \leq a + \frac{b-a}{4})\)?
- \(P(X \leq a + 3\frac{b-a}{4})\)?
- …
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = [a, b]\)  (it is infinite and not countable)

Events: \(\mathcal{U}\) : the borel set over \([a, b]\)

Probability function: \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I}{b-a}\)

Random variable:

\[
X: \Omega \rightarrow \mathbb{R} \quad \omega \mapsto \omega
\]

Cumulative distribution function:

What is \(P(X \leq t)\)?

\[
P(X \leq t) = F_X(t) = \frac{t-a}{b-a} \quad \text{if } a \leq t \leq b
\]
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Outcomes: \(\Omega = \mathbb{R}\) (it is infinite and not countable)

Events: \(\mathcal{U}\) : the borel set over \(\mathbb{R}\)

Probability function: \(\forall I \in \mathcal{U}, P(I) = \frac{\text{width of } I \cap [a, b]}{b-a}\)

Random variable: \(X\) (the value of the number)

Cumulative distribution function:
What is \(P(X \leq t)\)?

\[
P(X \leq t) = F_X(t) = \begin{cases} 
0 & \text{if } t \leq a \\
\frac{t-a}{b-a} & \text{if } a \leq t \leq b \\
1 & \text{otherwise}
\end{cases}
\]
Cumulative distribution function

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$, and consider the function:

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto P(X \leq t)$$

$F_X$ is the *Cumulative Distribution Function (CDF)* of $X$.

**Remarks:** Any r.v. has a CDF.

In particular, if $X$ is a discrete r.v., then

$$[X \leq t] = \bigcup_{i \in X(\Omega) | i \leq t} [X = i]$$

and

$$F_X(t) = P(X \leq t) = \sum_{i \in X(\Omega) | i \leq t} P(X = i)$$
Cumulative distribution function

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$, and consider the function:

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto P(X \leq t)$$

$F_X$ is the **Cumulative Distribution Function (CDF)** of $X$.

**Example:** We roll a fair die. $X$ takes the value of the number obtained. The CDF of $X$ is:
Cumulative distribution function: properties

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a r.v. on \((\Omega, \mathcal{U}, P)\).

Proposition:

- \(\forall t_1 \leq t_2, \; P(X \leq t_1) \leq P(X \leq t_2)\) the CDF is monotonically non-decreasing
- \(\lim_{t \to -\infty} P(X \leq t) = 0, \; \lim_{t \to +\infty} P(X \leq t) = 1\)
- \(\forall t \in \mathbb{R}, \; \lim_{\epsilon \to 0^+} P(X \leq t + \epsilon) = P(X \leq t)\) the CDF is continuous on the right in any \(t \in \mathbb{R}\)
Continuous Random Variable

Let $(\Omega, \mathcal{U}, P)$ be a probability space. Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$.

**Definition:** If the CDF of $X$ is continuous, i.e.

$$\forall t \in \mathbb{R}, \lim_{\epsilon \to 0} P(X \leq t + \epsilon) = P(X \leq t)$$

we say that $X$ is a *continuous r.v.*
Continuous Random Variable

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a r.v. on \((\Omega, \mathcal{U}, P)\).

**Definition:** If the CDF of \(X\) is continuous, *i.e.*

\[
\forall t \in \mathbb{R}, \lim_{\epsilon \to 0} P(X \leq t + \epsilon) = P(X \leq t)
\]

we say that \(X\) is a *continuous r.v.*

**Remark:** Don’t be afraid of the mathematical formalities. We will only consider r.v. that are either discrete (see the previous lecture) or continuous (*i.e.* having a CDF that is continuous).
CDF of a continuous random variable

Let $(\Omega, \mathcal{U}, P)$ be a probability space. Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$.

Proposition: Let $X$ be continuous (i.e. its CDF is continuous), then:

(i) $\forall t \in \mathbb{R}, \ P(X < t) = P(X \leq t)$

(ii) $\forall 0 < v < 1, \ \exists t \in \mathbb{R} \text{ s.t. } P(X \leq t) = v.$

i.e. the image of $\mathbb{R}$ by the CDF contains $]0, 1[$

\[
\begin{align*}
\text{CDF Graph} & \\
\text{CDF Graph} & \\
\end{align*}
\]
CDF of a continuous random variable

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$.

**Proposition:** Let $X$ be continuous (i.e. its CDF is continuous), then:

(i) \( \forall t \in \mathbb{R}, P(X < t) = P(X \leq t) \)

(ii) \( \forall 0 < v < 1, \exists t \in \mathbb{R} \text{ s.t. } P(X \leq t) = v. \)

**Consequences of (i):** Let $a < b$ be two real numbers.

- \( P(X \in [a, b]) = P(X \leq b) - P(X \leq a) \)

  because: \( [X \leq b] = [X < a] \cup [X \in [a, b]] \)

  then \( P(X \leq b) = P(X < a) + P(X \in [a, b]) = P(X \leq a) + P(X \in [a, b]) \)

  and \( P(X \in [a, b]) = P(X \leq b) - P(X \leq a) \)
CDF of a continuous random variable

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$.

**Proposition:** Let $X$ be continuous (i.e. its CDF is continuous), then:

(i) $\forall t \in \mathbb{R}, \ P(X < t) = P(X \leq t)$

(ii) $\forall 0 < v < 1, \exists t \in \mathbb{R}$ s.t. $P(X \leq t) = v$.

**Consequences of (i):** Let $a < b$ be two real numbers.

- $P(X \in [a, b]) = P(X \leq b) - P(X \leq a)$
- $P(X \in ]a, b[) = P(X \in [a, b[) = P(X \in ]a, b]) = P(X \in [a, b])$
CDF of a continuous random variable

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. on $(\Omega, \mathcal{U}, P)$.

**Proposition:** Let $X$ be continuous (i.e. its CDF is continuous), then:

(i) \( \forall t \in \mathbb{R}, \) \( P(X < t) = P(X \leq t) \)

(ii) \( \forall 0 < v < 1, \) \( \exists t \in \mathbb{R} \) s.t. \( P(X \leq t) = v. \)

**Consequences of (i):** Let $a < b$ be two real numbers.

- \( P(X \in [a, b]) = P(X \leq b) - P(X \leq a) \)
- \( P(X \in]a, b[) = P(X \in [a, b[ = P(X \in]a, b]) = P(X \in [a, b]) \)
- \( P(X = a) = P(X \in [a, a]) = 0 \)
Course 10 - Continuous Random Variables

I Random variables
   1 Example
   2 Definitions
   3 Density of probability of a continuous r.v.
   4 Expected value, variance, standard deviation

II Two continuous distributions

III Joint distribution
Probability Density Function

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X$ be a continuous r.v. so that the CDF $F_X$ of $X$ is differentiable. The function

$$f_X : \mathbb{R} \rightarrow \mathbb{R}$$

$$t \mapsto F'_X(t)$$

is called the *Probability Density Function (PDF)* of $X$. 

**Remark:** when the CDF of $X$ is piecewise differentiable or differentiable, we say that $X$ has a PDF. Notation: when $X$ has a PDF $f_X$, we note $\tilde{P}(X = t)$ for $f_X(t)$. 

**Remark:** We will never try to prove that an $X$ has a PDF: we will assume it, and use the properties it implies.
Probability Density Function

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X\) be a continuous r.v. so that the CDF \(F_X\) of \(X\) is differentiable. The function

\[
f_X : \mathbb{R} \rightarrow \mathbb{R} \\
t \mapsto F_X'(t)
\]

is called the *Probability Density Function (PDF)* of \(X\).

**Remark:** When \(F_X\) is only piecewise differentiable, one can define the PDF of \(X\) as:

\[
f_X : \mathbb{R} \rightarrow \mathbb{R} \\
t \mapsto \begin{cases} 
F_X'(t) & \text{if } F_X \text{ is differentiable in } t \\
g(t) & \text{otherwise}
\end{cases}
\]

where \(g(t) \in \mathbb{R}\) (for instance the left or the right derivative in \(t\)).


**Probability Density Function**

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X$ be a continuous r.v. so that the CDF $F_X$ of $X$ is differentiable. The function

$$f_X : \mathbb{R} \rightarrow \mathbb{R} \quad t \mapsto F'_X(t)$$

is called the *Probability Density Function (PDF)* of $X$.

**Remark:** when the CDF of $X$ is piecewise differentiable or differentiable, we say that $X$ has a PDF.

**Notation:** when $X$ has a PDF $f_X$, we note $\tilde{P}(X = t)$ for $f_X(t)$. 
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\). We suppose that each number appears with the same “likelihood”.

Random variable: \(X\) (the value of the number)

Cumulative distribution function:

\[
P(X \leq t) = F_X(t) = \begin{cases} 
0 & \text{if } t \leq a \\ 
\frac{t-a}{b-a} & \text{if } a \leq t \leq b \\ 
1 & \text{otherwise}
\end{cases}
\]

Probability density function:

\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\ 
\frac{1}{b-a} & \text{if } a \leq t \leq b \\ 
0 & \text{otherwise}
\end{cases}
\]
Probability Density Function

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X\) be a continuous r.v. so that the CDF \(F_X\) of \(X\) is differentiable. The function

\[
f_X : \mathbb{R} \rightarrow \mathbb{R}
\]

\[
t \mapsto F_X'(t)
\]

is called the *Probability Density Function (PDF)* of \(X\).

**Remark:** when the CDF of \(X\) is piecewise differentiable or differentiable, we say that \(X\) has a PDF.

**Notation:** when \(X\) has a PDF \(f_X\), we note \(\tilde{P}(X = t)\) for \(f_X(t)\).

**Remark:** We will never try to prove that an \(X\) has a PDF: we will assume it, and use the properties it implies.
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

where $\sigma = 0.5$, $\mu = 2$. 
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{p}(X = t) : \mathbb{R} \rightarrow [0, 1]
$$

$$
t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}
$$

where $\sigma = 0.5$, $\mu = 2$.

**What does it mean?** Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

- $r$ intervals,
- to each interval $I$ we associate the ratio of squirrels having weight in $I$

$$r = 10$$
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1] \\
t \mapsto \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}
$$

where $\sigma = 0.5$, $\mu = 2$.

What does it mean? Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

$r$ intervals,
to each interval $I$
we associate the ratio
of squirrels having weight in $I$
each bar is a $P(u \leq X \leq u + \epsilon)$

$r = 10$
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]
$$

$$
t \mapsto \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(t - \mu)^2}{2\sigma^2}}
$$

where $\sigma = 0.5$, $\mu = 2$.

**What does it mean?** Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

$r$ intervals,
to each interval $I$
we associate the ratio
of squirrels having weight in $I$
each bar is a
$$
P(u \leq X \leq u + \epsilon)
$$
r = 20
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$\tilde{P}(X = t) : \mathbb{R} \to [0, 1]$$

$$t \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

where $\sigma = 0.5$, $\mu = 2$.

**What does it mean?** Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

- $r$ intervals,
- to each interval $I$ we associate the ratio of squirrels having weight in $I$
- each bar is a $P(u \leq X \leq u + \epsilon)$

$r = 50$
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

Probability density function: We suppose that $X$ has a PDF given by

$$\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

where $\sigma = 0.5$, $\mu = 2$.

What does it mean? Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

$r$ intervals, to each interval $I$ we associate the ratio of squirrels having weight in $I$ divided by width of intervals

each bar is a $\frac{P(u \leq X \leq u + \epsilon)}{\epsilon}$

$r = 50$
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

Probability density function: We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]
$$

$$
t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}
$$

where $\sigma = 0.5$, $\mu = 2$.

What does it mean? Suppose we catch $n$ squirrels, and weight them. We construct the histogram:

$r$ intervals, to each interval $I$ we associate the ratio of squirrels having weight in $I$ divided by width of intervals each bar is a

$$
\lim_{u \to t, \epsilon \to 0} \frac{P(u \leq X \leq u + \epsilon)}{\epsilon}
$$

$$
r = 50 = (P(X \leq t))'
$$
Probability Density Function: properties

Let $(\Omega, \mathcal{U}, P)$ be a probability space. Let $X$ be a continuous r.v.

Remark: if $X$ has a PDF then for any $t \in \mathbb{R}$

\[
P(X \leq t) = \int_{-\infty}^{t} \tilde{P}(X = u) du
\]
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]
$$

$$
t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}
$$

where $\sigma = 0.5$, $\mu = 2$.

**Cumulative distribution function:**

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du
$$
**Probability Density Function: properties**

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v.

**Remark:** if \(X\) has a PDF then for any \(t \in \mathbb{R}\)

\[
P(X \leq t) = \int_{-\infty}^{t} \tilde{P}(X = u)du
\]

if \(I = [a, b]\) is an interval in \(\mathbb{R}\),

\[
P(X \in I) = \int_{a}^{b} \tilde{P}(X = u)du
\]

\[
P(X \in [a, b]) = P(X \leq b) - P(X \leq a)
\]

\[
= \int_{-\infty}^{b} \tilde{P}(X = u)du - \int_{-\infty}^{a} \tilde{P}(X = u)du
\]

\[
= \int_{-\infty}^{a} \tilde{P}(X = u)du + \int_{a}^{b} \tilde{P}(X = u)du - \int_{-\infty}^{a} \tilde{P}(X = u)du
\]
Probability Density Function: properties

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v.

Remark: if \(X\) has a PDF then for any \(t \in \mathbb{R}\)

\[
P(X \leq t) = \int_{-\infty}^{t} \tilde{P}(X = u)du
\]

if \(I = [a, b]\) is an interval in \(\mathbb{R}\),

\[
P(X \in I) = \int_{a}^{b} \tilde{P}(X = u)du
\]

\[
P(X \in [a, b]) = P(X \leq b) - P(X \leq a)
\]

\[
= \int_{-\infty}^{b} \tilde{P}(X = u)du - \int_{-\infty}^{a} \tilde{P}(X = u)du
\]

\[
= \int_{-\infty}^{a} \tilde{P}(X = u)du + \int_{a}^{b} \tilde{P}(X = u)du - \int_{-\infty}^{a} \tilde{P}(X = u)du
\]

Or you can just see the \(\int\) as the equivalent of the \(\sum\) on infinite non-countable sets.
Probability Density Function: properties

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v.

Remark: if \(X\) has a PDF then for any \(t \in \mathbb{R}\)

\[
P(X \leq t) = \int_{-\infty}^{t} \tilde{P}(X = u)\,du
\]

if \(I = [a, b]\) is an interval in \(\mathbb{R}\),

\[
P(X \in I) = \int_{a}^{b} \tilde{P}(X = u)\,du
\]

if \(I\) is any finite union or intersection of intervals in \(\mathbb{R}\),

\[
P(X \in I) = \int_{I} \tilde{P}(X = u)\,du
\]
Course 10 - Continuous Random Variables

I Random variables
   1 Example
   2 Definitions
   3 Density of probability of a continuous r.v.
   4 Expected value, variance, standard deviation

II Two continuous distributions

III Joint distribution
Expected value

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v. having a PDF.

**Definition:** The *expected value* of \(X\) is the number in \(\mathbb{R}\),

\[
\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) du
\]
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Random variable: $X$ (the value of the number)

Probability density function:

$$\tilde{P}(X = t) = \begin{cases} 0 & \text{if } t < a \\ \frac{1}{b-a} & \text{if } a \leq t \leq b \\ 0 & \text{otherwise} \end{cases}$$

Expected value:

$$\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) du$$
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Random variable: \(X\) (the value of the number)

Probability density function:
\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

Expected value:
\[
\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) du = \int_{-\infty}^{a} 0 du + \int_{a}^{b} u \frac{1}{b-a} du + \int_{b}^{+\infty} 0 du = \frac{1}{b-a} \int_{a}^{b} u du
\]
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Random variable: $X$ (the value of the number)

Probability density function:

$$\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}$$

Expected value:

$$\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) du = \int_{-\infty}^{a} 0 du + \int_{a}^{b} u \frac{1}{b-a} du + \int_{b}^{+\infty} 0 du = \frac{1}{b-a} \int_{a}^{b} u du$$

$$= \frac{1}{b-a} \left[ \frac{u^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(a+b)}{2(b-a)} = \frac{a+b}{2}$$
Expected value

Let $(\Omega, \mathcal{U}, P)$ be a probability space. Let $X$ be a continuous r.v. having a PDF.

**Definition:** The *expected value* of $X$ is the number in $\mathbb{R}$,

$$\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) \, du$$

**Remark:** The expected value of $X$ is not necessarily finite.
Expected value

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v. having a PDF.

**Definition:** The *expected value* of \(X\) is the number in \(\mathbb{R}\),

\[
\text{Exp}(X) = \int_{u \in \mathbb{R}} u \tilde{P}(X = u) \, du
\]

**Remark:** The expected value of \(X\) is not necessarily finite.

When \(\text{Exp}(X)\) is finite, we say that \(X\) has an expected value.
Expected value: properties (I)

Let \((\Omega, \mathcal{U}, P)\) be a probability space.
Let \(X\) be a continuous r.v. on \((\Omega, \mathcal{U}, P)\).

**Theorem**: If \(X\) has an expected value and \(a \in \mathbb{R}\),

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

(ii) \(\text{Exp}(a + X) = a + \text{Exp}(X)\)
Variance and standard deviation

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a continuous r.v. having a PDF.

**Definition:** If $X$ has an expected value $\mu$, the *variance* of $X$ is the number

$$Var(X) = \text{Exp}((X - \mu)^2) = \int_{u \in \mathbb{R}} P(X = u)(u - \mu)^2 du$$
Variance and standard deviation

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a continuous r.v. having a PDF.

Definition: If $X$ has an expected value $\mu$, the variance of $X$ is the number

$$Var(X) = \text{Exp}((X - \mu)^2) = \int_{u \in \mathbb{R}} P(X = u)(u - \mu)^2 \, du$$

Remark: Even if $X$ has an expected value, $Var(X)$ is not necessarily finite. If $Var(X)$ is finite, we say that $X$ has a variance.
Variance and standard deviation

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a continuous r.v. having a PDF.

**Definition:** If \(X\) has an expected value \(\mu\), the **variance** of \(X\) is the number

\[
\text{Var}(X) = \text{Exp}((X - \mu)^2) = \int_{u \in \mathbb{R}} P(X = u)(u - \mu)^2 \, du
\]

**Remark:** Even if \(X\) has an expected value, \(\text{Var}(X)\) is not necessarily finite. If \(\text{Var}(X)\) is finite, we say that \(X\) has a variance.

**Definition:** If \(X\) has an expected value and a variance, the **standard deviation** of \(X\) is the number

\[
\text{Std}(X) = \sqrt{\text{Var}(X)}
\]

\(\text{Std}(X)\) is the spread of the distribution around \(\text{Exp}(X)\)
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Random variable: $X$ (the value of the number)

Probability density function:

$$
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
$$

Expected value: $\text{Exp}(X) = \frac{a+b}{2}$

Variance:

$$
\text{Var}(X) = \int_{u \in \mathbb{R}} (u - \text{Exp}(X))^2 \tilde{P}(X = u) du
$$
Random number

We pick up a number in the interval $[a, b]$, where $a < b$. We suppose that each number appears with the same “likelihood”.

Random variable: $X$ (the value of the number)

Probability density function:

$$\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}$$

Expected value: $\text{Exp}(X) = \frac{a+b}{2}$

Variance:

$$\text{Var}(X) = \int_{u \in \mathbb{R}} (u - \text{Exp}(X))^2 \tilde{P}(X = u) du = \int_a^b (u - \text{Exp}(X))^2 \frac{1}{b-a} du$$
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Random variable: \(X\) (the value of the number)

Probability density function:

\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

Expected value: \(\text{Exp}(X) = \frac{a+b}{2}\)

Variance:

\[
\text{Var}(X) = \int_{u \in \mathbb{R}} (u - \text{Exp}(X))^2 \tilde{P}(X = u) du = \int_a^b (u - \text{Exp}(X))^2 \frac{1}{b-a} du
\]

\[
= \int_a^b u^2 \frac{1}{b-a} du - 2\text{Exp}(X) \int_a^b u \frac{1}{b-a} du + \text{Exp}(X)^2 \int_a^b \frac{1}{b-a} du
\]
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Random variable: \(X\) (the value of the number)

Probability density function:

\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

Expected value: \(\text{Exp}(X) = \frac{a+b}{2}\)

Variance:

\[
\text{Var}(X) = \int_{u \in \mathbb{R}} (u - \text{Exp}(X))^2 \tilde{P}(X = u) \, du = \int_{a}^{b} (u - \text{Exp}(X))^2 \frac{1}{b-a} \, du \\
= \int_{a}^{b} u^2 \frac{1}{b-a} \, du - 2\text{Exp}(X) \int_{a}^{b} u \frac{1}{b-a} \, du + \text{Exp}(X)^2 \int_{a}^{b} \frac{1}{b-a} \, du \\
= \frac{b^3-a^3}{3(b-a)} - 2\left(\frac{a+b}{2}\right)^2 + \left(\frac{a+b}{2}\right)^2 \frac{b-a}{b-a}
\]
Random number

We pick up a number in the interval \([a, b]\), where \(a < b\).
We suppose that each number appears with the same “likelihood”.

Random variable: \(X\) (the value of the number)

Probability density function:
\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t < a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

Expected value: \(\text{Exp}(X) = \frac{a+b}{2}\)

Variance:
\[
\text{Var}(X) = \int_{u \in \mathbb{R}} (u - \text{Exp}(X))^2 \tilde{P}(X = u) \, du = \int_a^b (u - \text{Exp}(X))^2 \frac{1}{b-a} \, du
\]
\[
= \int_a^b u^2 \frac{1}{b-a} \, du - 2\text{Exp}(X) \int_a^b u \frac{1}{b-a} \, du + \text{Exp}(X)^2 \int_a^b \frac{1}{b-a} \, du
\]
\[
= \frac{b^3-a^3}{3(b-a)} - 2\left(\frac{a+b}{2}\right)^2 + \left(\frac{a+b}{2}\right)^2 \frac{b-a}{b-a} = \ldots = \frac{(b-a)^2}{12}
\]
Variance: properties (I)

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a continuous r.v. having a PDF.

**Theorem:**

(i) if $X$ has a variance, then $X$ has an expected value

(ii) if $X$ has a variance, then

$$\text{Var}(X) = \text{Exp}(X^2) - (\text{Exp}(X))^2$$
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions
   1 The continuous uniform distribution
   2 The Gaussian, or normal, distribution

III Joint distribution
The continuous uniform distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X\) be a r.v. having a PDF.

**Definition:** If \(X(\Omega) = [a, b]\) and \(\forall t \in \mathbb{R}\)

\[
P(X \leq t) = \begin{cases} 
0 & \text{if } t \leq a \\
\frac{t-a}{b-a} & \text{if } a \leq t \leq b \\
1 & \text{otherwise}
\end{cases}
\]

and

\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t \leq a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

we say that \(X\) follows the continuous uniform distribution.
The continuous uniform distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.
Let \(X\) be a r.v. having a PDF.

**Definition:** If \(X(\Omega) = [a, b]\) and \(\forall t \in \mathbb{R}\)

\[
P(X \leq t) = \begin{cases} 
0 & \text{if } t \leq a \\
\frac{t-a}{b-a} & \text{if } a \leq t \leq b \\
1 & \text{otherwise}
\end{cases}
\]

and

\[
\tilde{P}(X = t) = \begin{cases} 
0 & \text{if } t \leq a \\
\frac{1}{b-a} & \text{if } a \leq t \leq b \\
0 & \text{otherwise}
\end{cases}
\]

we say that \(X\) follows the continuous uniform distribution.

**Expected value:** \(\text{Exp}(X) = \frac{a+b}{2}\)

**Variance:** \(\text{Var}(X) = \frac{(b-a)^2}{12}\)
Course 10 - Continuous Random Variables

I  Random variables
II  Two continuous distributions
   1  The continuous uniform distribution
   2  The Gaussian, or normal, distribution
III Joint distribution
The Gaussian, or normal distribution

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. having a PDF.

**Definition:** If $X(\Omega) = \mathbb{R}$ and $\exists \mu \in \mathbb{R}, \sigma \in \mathbb{R}$ s.t. $\forall t \in \mathbb{R}$

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

we say that $X$ follows the Gaussian, or Normal, distribution.
The Gaussian, or normal distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.
Let \(X\) be a r.v. having a PDF.

**Definition:** If \(X(\Omega) = \mathbb{R}\) and \(\exists \mu \in \mathbb{R}, \sigma \in \mathbb{R}\) s.t. \(\forall t \in \mathbb{R}\)

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

we say that \(X\) follows the Gaussian, or Normal, distribution.

**Cumulative distr. func.:**

\[
P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du
\]
The Gaussian, or normal distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.
Let \(X\) be a r.v. having a PDF.

**Definition:** If \(X(\Omega) = \mathbb{R}\) and \(\exists \mu \in \mathbb{R}, \sigma \in \mathbb{R}\) s.t. \(\forall t \in \mathbb{R}\)

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

we say that \(X\) **follows the Gaussian, or Normal, distribution.**

**Cumulative distr. func.:** \(P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du\)

**Expected value:** \(\text{Exp}(X) = \mu\)

**Variance:** \(\text{Var}(X) = \sigma^2\)
The Gaussian, or normal distribution

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. having a PDF.

**Definition:** If $X(\Omega) = \mathbb{R}$ and $\exists \mu \in \mathbb{R}, \sigma \in \mathbb{R}$ s.t. $\forall t \in \mathbb{R}$

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

we say that $X$ follows the Gaussian, or Normal, distribution.

Cumulative distr. func.: $P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^{t} N_{\mu,\sigma}(u) du$

Expected value: $\text{Exp}(X) = \mu$

Variance: $\text{Var}(X) = \sigma^2$

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

$N_{0,1}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

called standard normal distribution
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1] \\
\quad t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{2,0.5}(t)
$$

where $\sigma = 0.5$, $\mu = 2$.

**Cumulative distribution function:**

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, du = \int_{-\infty}^{t} N_{2,0.5}(u) \, du
$$
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1] \\
t \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{2,0.5}(t)
$$

where $\sigma = 0.5$, $\mu = 2$.

**Cumulative distribution function:**

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^{t} N_{2,0.5}(u) du
$$

**Question:** What is $\tilde{P}(X = 2.2)$?
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

Probability density function: We suppose that $X$ has a PDF given by

$$\tilde{P}(X = t) : \mathbb{R} \rightarrow [0, 1]$$

$$t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{2,0.5}(t)$$

where $\sigma = 0.5$, $\mu = 2$.

Cumulative distribution function:

$$P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^{t} N_{2,0.5}(u) du$$

Question: What is $\tilde{P}(X = 2.2)$? evaluate $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2.2-\mu)^2}{2\sigma^2}} \approx 0.7365$ or normpdf(2.2,2,0.5) in Matlab.
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \to [0, 1] \\

t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{2,0.5}(t)
$$

where $\sigma = 0.5$, $\mu = 2$.

**Cumulative distribution function:**

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^{t} N_{2,0.5}(u) du
$$

**Question:** What is $\tilde{P}(X = 2.2)$? evaluate $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2.2-\mu)^2}{2\sigma^2}} \approx 0.7365$
or `normpdf(2.2,2,0.5)` in Matlab

**Question:** What is $P(X \leq 2.2)$?
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \to [0, 1]
$$

$$
t \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = \mathcal{N}_{2,0.5}(t)
$$

where $\sigma = 0.5$, $\mu = 2$.

**Cumulative distribution function:**

$$
P(X \leq t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{t} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^{t} \mathcal{N}_{2,0.5}(u) du
$$

**Question:** What is $\tilde{P}(X = 2.2)$? evaluate

$$
\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2.2-\mu)^2}{2\sigma^2}} \approx 0.7365
$$

or `normpdf(2.2,2,0.5)` in Matlab

**Question:** What is $P(X \leq 2.2)$? it is

$$
\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du
$$

How to compute this?
Computing the CDF of a Gaussian

**Question:** What is $P(X \leq 2.2)$? It is

$$
\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, du
$$

How to compute this?

In Matlab:

- $E(y)$ is obtained with `(1/2)*(1+erf(y/sqrt(2)))`.
- $P(X \leq t)$ is obtained with `(1/2)*(1+erf(((t-\mu)/\sigma)/sqrt(2)))`.
- Or: $P(X \leq t)$ is obtained with `normcdf(t,\mu,\sigma)`.
Computing the CDF of a Gaussian

Question: What is $P(X \leq 2.2)$? It is $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$

How to compute this?

Proposition: If $X$ follows $N_{\mu,\sigma}$ then $Y = \frac{X-\mu}{\sigma}$ follows $N_{0,1}$.
Computing the CDF of a Gaussian

Question: What is $P(X \leq 2.2)$? It is $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$

How to compute this?

Proposition: If $X$ follows $N(\mu, \sigma)$ then $Y = \frac{X-\mu}{\sigma}$ follows $N(0,1)$.

Consequences: $P(X \leq t) = P\left(\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}\right) = P(Y \leq \frac{t-\mu}{\sigma}) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$
Computing the CDF of a Gaussian

Question: What is $P(X \leq 2.2)$? It is
\[\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, du\]

How to compute this?

Proposition: If $X$ follows $N(\mu, \sigma)$ then $Y = \frac{X-\mu}{\sigma}$ follows $N(0,1)$.

Consequences: $P(X \leq t) = P\left(\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}\right) = P(Y \leq \frac{t-\mu}{\sigma}) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du$

If we know how to compute

\[E(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du = \int_{-\infty}^{y} N_{0,1}(u) \, du\]

for any $y$, we are all set
Computing the CDF of a Gaussian

**Question:** What is \( P(X \leq 2.2) \)? It is \( \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, du \)

How to compute this?

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

**Consequences:** \( P(X \leq t) = P\left( \frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma} \right) = P(Y \leq \frac{t-\mu}{\sigma}) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du \)

If we know how to compute

\[
E(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du = \int_{-\infty}^{y} N_{0,1}(u) \, du
\]

for any \( y \), we are all set

In Matlab: \( E(y) \) is obtained with \( (1/2)*(1+\text{erf}(y/\text{sqrt}(2))) \).
Computing the CDF of a Gaussian

Question: What is $P(X \leq 2.2)$? It is $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$

How to compute this?

Proposition: If $X$ follows $N_{\mu,\sigma}$ then $Y = \frac{X-\mu}{\sigma}$ follows $N_{0,1}$.

Consequences: $P(X \leq t) = P(\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}) = P(Y \leq \frac{t-\mu}{\sigma}) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$

If we know how to compute

$$E(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \int_{-\infty}^{y} N_{0,1}(u)du$$

for any $y$, we are all set

In Matlab: $E(y)$ is obtained with $(1/2)*(1+\text{erf}(y/\text{sqrt}(2)))$.

then $P(X \leq t)$ is obtained with $(1/2)*(1+\text{erf}(((t-\mu)/\sigma)/\text{sqrt}(2)))$. 
Computing the CDF of a Gaussian

**Question:** What is $P(X \leq 2.2)$? It is $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{2.2} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$

**How to compute this?**

**Proposition:** If $X$ follows $N_\mu,\sigma$ then $Y = \frac{X-\mu}{\sigma}$ follows $N_{0,1}$.

**Consequences:** $P(X \leq t) = P\left(\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}\right) = P(Y \leq \frac{t-\mu}{\sigma}) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$

If we know how to compute

$$E(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \int_{-\infty}^{y} N_{0,1}(u)du$$

for any $y$, we are all set

In Matlab: $E(y)$ is obtained with $(1/2)*(1+erf(y/sqrt(2)))$.

then $P(X \leq t)$ is obtained with $(1/2)*(1+erf(((t-\mu)/\sigma)/sqrt(2)))$.

or: $P(X \leq t)$ is obtained with `normcdf(t,mu,sigma)`. 
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions

III Joint distribution
   1 Definitions
   2 Independence
   3 Conditional probabilities
   4 Expected value, variance
   5 Sum of Gaussian distribution
Joint distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X, Y\) be two continuous r.v.

The **joint distribution of \(X, Y\)** is the CDF:

\[
F : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1] \\
t_x, t_y \mapsto P(X \leq t_x, Y \leq t_y)
\]
Joint distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X, Y\) be two continuous r.v.

The *joint distribution of \(X, Y\)* is the CDF:

\[
F : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1] \\
t_x, t_y \mapsto P(X \leq t_x, Y \leq t_y)
\]

**Definition:** If the joint distribution of \(X, Y\) is continuous,
we call *joint probability density of \(X, Y\)* the function

\[
\tilde{P}(X = t_x, Y = t_y) = \lim_{u \rightarrow t_x, v \rightarrow t_y, \epsilon \rightarrow 0^+} \frac{P((u \leq X \leq u + \epsilon) \cap (v \leq Y \leq v + \epsilon))}{\epsilon^2}
\]

if this limit exists.
Joint distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

Definition: Let \(X, Y\) be two continuous r.v.

The **joint distribution of \(X, Y\)** is the CDF:

\[
F : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1] \\
t_x, t_y \mapsto P(X \leq t_x, Y \leq t_y)
\]

Definition: If the joint distribution of \(X, Y\) is continuous, we call **joint probability density of \(X, Y\)** the function

\[
\tilde{P}(X = t_x, Y = t_y) = \lim_{u \rightarrow t_x, v \rightarrow t_y, \epsilon \rightarrow 0^+} \frac{P((u \leq X \leq u + \epsilon) \cap (v \leq Y \leq v + \epsilon))}{\epsilon^2}
\]

if this limit exists.

We extend these definitions to any vector of \(n\) r.v.
Joint probability density

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X, Y\) be two continuous r.v.

Remark: if \(X, Y\) have a joint probability density, then for any \(t_x, t_y\),

\[
P(X \leq t_x, Y \leq t_y) = \int_{-\infty}^{t_x} \int_{-\infty}^{t_y} \tilde{P}(X = u, Y = v) dv du
\]

\[
= \int_{-\infty}^{t_y} \int_{-\infty}^{t_x} \tilde{P}(X = u, Y = v) dudv
\]
**Joint probability density**

Let \((\Omega, \mathcal{U}, P)\) be a probability space. Let \(X, Y\) be two continuous r.v.

**Remark:** if \(X, Y\) have a joint probability density, then for any \(t_x, t_y\),

\[
P(X \leq t_x, Y \leq t_y) = \int_{-\infty}^{t_x} \int_{-\infty}^{t_y} \tilde{P}(X = u, Y = v) dvdu
\]

for any intervals \(I, J\) of \(\mathbb{R}^2\):

\[
P(X \in I, Y \leq t_y) = \int_{u \in I} \int_{-\infty}^{t_y} \tilde{P}(X = u, Y = v) dvdu
\]

\[
P(X \leq t_x, Y \in J) = \int_{-\infty}^{t_y} \int_{v \in J} \tilde{P}(X = u, Y = v) dvdu
\]

\[
P(X \in I, Y \in J) = \int_{u \in I} \int_{v \in J} \tilde{P}(X = u, Y = v) dvdu
\]
Marginal distributions

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Proposition:** Let $X, Y$ be two continuous r.v. having a joint probability density. Then

(i) $X$ and $Y$ have a PDF

(ii) the PDF of $X$ is:

$$
\tilde{P}(X = t_x) = \int_{v \in \mathbb{R}} \tilde{P}(X = t_x, Y = v) \, dv
$$
Marginal distributions

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Proposition:** Let \(X, Y\) be two continuous r.v. having a joint probability density. Then

1. \(X\) and \(Y\) have a PDF
2. the PDF of \(X\) is:
   \[
   \tilde{P}(X = t_x) = \int_{v \in \mathbb{R}} \tilde{P}(X = t_x, Y = v) dv
   \]
3. the PDF of \(Y\) is:
   \[
   \tilde{P}(Y = t_y) = \int_{u \in \mathbb{R}} \tilde{P}(X = u, Y = t_y) du
   \]
Operations on continuous random variables

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Proposition:** Let $\lambda \in \mathbb{R}$, and $X$, $Y$ be two r.v. having a joint probability density.

- $\lambda X$ is a continuous r.v. having a PDF
- $X + Y$ is a continuous r.v. having a PDF
- $XY$ is a continuous r.v. having a PDF
- In particular: $X_1 + \ldots + X_n$ is a continuous r.v. having a PDF.
Operations on continuous random variables

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Proposition:** Let \(\lambda \in \mathbb{R}\), and \(X, Y\) be two r.v. having a joint probability density.

- \(\lambda X\) is a continuous r.v. having a PDF
- \(X + Y\) is a continuous r.v. having a PDF
Operations on continuous random variables

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Proposition:** Let $\lambda \in \mathbb{R}$, and $X, Y$ be two r.v. having a joint probability density.

- $\lambda X$ is a continuous r.v. having a PDF
- $X + Y$ is a continuous r.v. having a PDF
- $XY$ is a continuous r.v. having a PDF
Operations on continuous random variables

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Proposition:** Let $\lambda \in \mathbb{R}$, and $X, Y$ be two r.v. having a joint probability density.

- $\lambda X$ is a continuous r.v. having a PDF
- $X + Y$ is a continuous r.v. having a PDF
- $XY$ is a continuous r.v. having a PDF

In particular: $X_1 + \ldots + X_n$ is a continuous r.v. having a PDF.
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions

III Joint distribution
   1 Definitions
   2 Independence
   3 Conditional probabilities
   4 Expected value, variance
   5 Sum of Gaussian distribution
Independence

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X, Y\) be two r.v. having a joint probability density. \(X\) and \(Y\) are *independent* if \(\forall t_x \in \mathbb{R}, \forall t_y \in \mathbb{R},\)

\[
\tilde{P}(X = t_x, Y = t_y) = \tilde{P}(X = t_x) \times \tilde{P}(Y = t_y)
\]
Independence

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X, Y$ be two r.v. having a joint probability density. $X$ and $Y$ are *independent* if $\forall t_x \in \mathbb{R}, \forall t_y \in \mathbb{R},$

$$\tilde{P}(X = t_x, Y = t_y) = \tilde{P}(X = t_x) \times \tilde{P}(Y = t_y)$$

**Definition:** Let $X_1, \ldots, X_n$ be $n$ r.v. having a joint probability density. $X_1, \ldots, X_n$ are *mutually independent* if $\forall (t_1, \ldots, t_n) \in \mathbb{R}^n$, for any subset $\{Y_1 = y_1, \ldots, Y_n' = y_n'\}$ of $\{X_1 = t_1, \ldots, X_n = t_n\},$

$$\tilde{P}(Y_1 = y_1, \ldots, Y_n' = y_n') = \tilde{P}(Y_1 = y_1) \times \ldots \times \tilde{P}(Y_n' = y_n')$$

In particular,

$$\tilde{P}(X_1 = t_1, \ldots, X_n = t_n) = \tilde{P}(X_1 = t_1) \times \ldots \times \tilde{P}(X_n = t_n)$$
Independence

Let $(\Omega, \mathcal{U}, P)$ be a probability space.

**Definition:** Let $X, Y$ be two r.v. having a joint probability density. $X$ and $Y$ are *independent* if $\forall t_x \in \mathbb{R}, \forall t_y \in \mathbb{R},$

$$\tilde{P}(X = t_x, Y = t_y) = \tilde{P}(X = t_x) \times \tilde{P}(Y = t_y)$$

**Definition:** Let $X_1, \ldots, X_n$ be $n$ r.v. having a joint probability density. $X_1, \ldots, X_n$ are *mutually independent* if $\forall (t_1, \ldots, t_n) \in \mathbb{R}^n$, for any subset $\{Y_1 = y_1, \ldots, Y_n' = y_n'\}$ of $\{X_1 = t_1, \ldots, X_n = t_n\},$

$$\tilde{P}(Y_1 = y_1, \ldots, Y_n' = y_n') = \tilde{P}(Y_1 = y_1) \times \ldots \times \tilde{P}(Y_n' = y_n')$$

In particular,

$$\tilde{P}(X_1 = t_1, \ldots, X_n = t_n) = \tilde{P}(X_1 = t_1) \times \ldots \times \tilde{P}(X_n = t_n)$$

**Remark:** We will never prove that r.v. are independent; we will suppose it!
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions

III Joint distribution
   1 Definitions
   2 Independence
   3 Conditional probabilities
   4 Expected value, variance
   5 Sum of Gaussian distribution
Conditional density of probability

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition**: Let \(X, Y\) be two r.v. having a joint probability density. Suppose that \(t_y\) is such that \(\tilde{P}(Y = t_y) \neq 0\). The *conditional density probability of \(X\) knowing \(Y\)* is defined as

\[
\tilde{P}(X = t_x | Y = t_y) = \frac{\tilde{P}(X = t_x, Y = t_y)}{\tilde{P}(Y = t_y)}
\]
Conditional density of probability

Let \((\Omega, \mathcal{U}, \mathbb{P})\) be a probability space.

**Definition:** Let \(X, Y\) be two r.v. having a joint probability density. Suppose that \(t_y\) is such that \(\tilde{P}(Y = t_y) \neq 0\). The *conditional density probability of \(X\) knowing \(Y\)* is defined as

\[
\tilde{P}(X = t_x | Y = t_y) = \frac{\tilde{P}(X = t_x, Y = t_y)}{\tilde{P}(Y = t_y)}
\]

**Remark:** and the Bayes’ law is

\[
\tilde{P}(X = t_x | Y = t_y) = \frac{\tilde{P}(Y = t_y | X = t_x) \times \tilde{P}(X = t_x)}{\tilde{P}(Y = t_y)}
\]
Conditional density of probability

Let \((\Omega, \mathcal{U}, P)\) be a probability space.

**Definition:** Let \(X, Y\) be two r.v. having a joint probability density. Suppose that \(t_y\) is such that \(\tilde{P}(Y = t_y) \neq 0\). The **conditional density probability of \(X\) knowing \(Y\)** is defined as

\[
\tilde{P}(X = t_x | Y = t_y) = \frac{\tilde{P}(X = t_x, Y = t_y)}{\tilde{P}(Y = t_y)}
\]

**Remark:** and the Bayes’ law is

\[
\tilde{P}(X = t_x | Y = t_y) = \frac{\tilde{P}(Y = t_y | X = t_x) \times \tilde{P}(X = t_x)}{\tilde{P}(Y = t_y)}
\]

**Remark:** the definition of conditional independence is similar to the case of discrete r.v.
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

Random variables: $X : \Omega \rightarrow \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \rightarrow \mathbb{R}$, its weight.

Assumptions: $P(X = 0) = (1 - p)$, $P(X = 1) = p$.

Assumptions: $\tilde{P}(Y = t | X = 0) = N(\mu_0, \sigma_0)(t)$.

Assumptions: $\tilde{P}(Y = t | X = 1) = N(\mu_1, \sigma_1)(t)$.

Classification step: we pick-up randomly a squirrel in WSP, its weight is $w$ kg.

Classification step: Decide if it is a male or a female.

Learning step: find the parameters $p$, $\mu_0$, $\sigma_0$, $\mu_1$, $\sigma_1$ from a learning set $L$.

- $p \leftarrow \text{the number of males in } L / \text{size of } L$.
- $\mu_0 \leftarrow \text{the average weight of males in } L$.
- $\sigma_0 \leftarrow \text{the std. dev. of weights of males in } L$, i.e. $\sqrt{\frac{\sum (w_i - \mu_0)^2}{|L|}}$. . .
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

Random variables: $X : \Omega \to \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \to \mathbb{R}$, its weight.

Assumptions: $P(X = 0) = (1 - p)$, $P(X = 1) = p$

$\tilde{P}(Y = t|X = 0) = N_{\mu_0,\sigma_0}(t)$

$\tilde{P}(Y = t|X = 1) = N_{\mu_1,\sigma_1}(t)$
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

Random variables: $X : \Omega \to \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \to \mathbb{R}$, its weight.

Assumptions: $P(X = 0) = (1 - p), P(X = 1) = p$

$$\tilde{P}(Y = t | X = 0) = N_{\mu_0, \sigma_0}(t)$$

$$\tilde{P}(Y = t | X = 1) = N_{\mu_1, \sigma_1}(t)$$

Classification step: we pick-up randomly a squirrel in WSP, its weight is $w$ kg. Decide if it is a male or a female.
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

**Random variables:** $X : \Omega \rightarrow \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \rightarrow \mathbb{R}$, its weight.

**Assumptions:** $P(X = 0) = (1 - p)$, $P(X = 1) = p$

$$\tilde{P}(Y = t|X = 0) = N_{\mu_0, \sigma_0}(t)$$

$$\tilde{P}(Y = t|X = 1) = N_{\mu_1, \sigma_1}(t)$$

**Classification step:** we pick-up randomly a squirrel in WSP, its weight is $w$ kg. Decide if it is a male or a female

**Learning step:** find the parameters $p, \mu_0, \sigma_0, \mu_1, \sigma_1$ from a learning set $L$

- $p \leftarrow$
- $\mu_0 \leftarrow$
- $\sigma_0 \leftarrow$
- $\ldots$
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

**Random variables:** $X : \Omega \to \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \to \mathbb{R}$, its weight.

**Assumptions:** $P(X = 0) = (1 - p)$, $P(X = 1) = p$

$$\tilde{P}(Y = t | X = 0) = N_{\mu_0, \sigma_0}(t)$$

$$\tilde{P}(Y = t | X = 1) = N_{\mu_1, \sigma_1}(t)$$

**Classification step:** we pick-up randomly a squirrel in WSP, its weight is $w$ kg. Decide if it is a male or a female

**Learning step:** find the parameters $p, \mu_0, \sigma_0, \mu_1, \sigma_1$ from a learning set $\mathcal{L}$

- $p \leftarrow$ the number of males in $\mathcal{L}$ / size of $\mathcal{L}$
- $\mu_0 \leftarrow$
- $\sigma_0 \leftarrow$
- $\ldots$
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

**Random variables:** $X : \Omega \to \{0, 1\}$, the sex of a squirrel picked up randomly,
$Y : \Omega \to \mathbb{R}$, its weight.

**Assumptions:** $P(X = 0) = (1 - p)$, $P(X = 1) = p$

$\tilde{P}(Y = t | X = 0) = N_{\mu_0, \sigma_0}(t)$

$\tilde{P}(Y = t | X = 1) = N_{\mu_1, \sigma_1}(t)$

**Classification step:** we pick-up randomly a squirrel in WSP, its weight is $w$ kg. Decide if it is a male or a female

**Learning step:** find the *parameters* $p, \mu_0, \sigma_0, \mu_1, \sigma_1$ from a *learning set* $\mathcal{L}$

- $p \leftarrow$ the number of males in $\mathcal{L}$ / size of $\mathcal{L}$
- $\mu_0 \leftarrow$ the average weight of males in $\mathcal{L}$
- $\sigma_0 \leftarrow$
Example: a bayesian classifier in the continuous case

Let $\Omega$ be the set of the squirrels in WSP.

**Random variables:** $X : \Omega \rightarrow \{0, 1\}$, the sex of a squirrel picked up randomly, $Y : \Omega \rightarrow \mathbb{R}$, its weight.

**Assumptions:** $P(X = 0) = (1 - p)$, $P(X = 1) = p$

\[
\tilde{P}(Y = t | X = 0) = N_{\mu_0, \sigma_0}(t)
\]

\[
\tilde{P}(Y = t | X = 1) = N_{\mu_1, \sigma_1}(t)
\]

**Classification step:** we pick-up randomly a squirrel in WSP, its weight is $w$ kg. Decide if it is a male or a female.

**Learning step:** find the parameters $p, \mu_0, \sigma_0, \mu_1, \sigma_1$ from a learning set $\mathcal{L}$

- $p \leftarrow$ the number of males in $\mathcal{L} / \text{size of } \mathcal{L}$
- $\mu_0 \leftarrow$ the average weight of males in $\mathcal{L}$
- $\sigma_0 \leftarrow$ the std. dev. of weights of males in $\mathcal{L}$, i.e. $\sqrt{\left(\sum (w_i - \mu_0)^2\right) / |\mathcal{L}|}$
- …
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions

III Joint distribution
   1 Definitions
   2 Independence
   3 Conditional probabilities
   4 Expected value, variance
   5 Sum of Gaussian distribution
Expected value: properties (II)

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X, Y$ be two r.v. on $(\Omega, \mathcal{U}, P)$, having a joint probability density.

**Theorem:** If $X$ and $Y$ have an expected value, and $a \in \mathbb{R}$,

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

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

(iii) $\text{Exp}(a + X) = a + \text{Exp}(X)$
Expected value: properties (II)

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X, Y$ be two r.v. on $(\Omega, \mathcal{U}, P)$. having a joint probability density.

**Theorem:** If $X$ and $Y$ have an expected value, and $a \in \mathbb{R}$,

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

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

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

**Theorem:** If $X$ and $Y$ have an expected value, and are independent,

$$\text{Exp}(XY) = \text{Exp}(X)\text{Exp}(Y)$$
Variance: properties (II)

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X, Y$ be two r.v. having a joint distribution.

**Theorem:**

(i) if $X$ has a variance, then $X$ has an expected value

(ii) if $X$ has a variance, then

$$Var(X) = Exp(X^2) - (Exp(X))^2$$

(iii) if $X$ and $Y$ have a variance and are independent, then

$$Var(X + Y) = Var(X) + Var(Y)$$
Course 10 - Continuous Random Variables

I Random variables

II Two continuous distributions

III Joint distribution
   1 Definitions
   2 Independence
   3 Conditional probabilities
   4 Expected value, variance
   5 Sum of Gaussian distribution
Sum of r.v. with Gaussian distribution

Let $(\Omega, \mathcal{U}, P)$ be a probability space.
Let $X$ be a r.v. following a Gaussian distribution with mean $\mu_X$ and variance $\sigma_X^2$.
Let $Y$ be a r.v. following a Gaussian distribution with mean $\mu_Y$ and variance $\sigma_Y^2$.

Proposition: if $X$ and $Y$ are independent, $X + Y$ follows a Gaussian distribution with mean $\mu_X + \mu_Y$ and variance $\sigma_X^2 + \sigma_Y^2$. 
Sum of r.v. with Gaussian distribution

Let \((\Omega, \mathcal{U}, P)\) be a probability space.
Let \(X\) be a r.v. following a Gaussian distribution with mean \(\mu_X\) and variance \(\sigma_X^2\).
Let \(Y\) be a r.v. following a Gaussian distribution with mean \(\mu_Y\) and variance \(\sigma_Y^2\).

**Proposition:** if \(X\) and \(Y\) are independent, \(X + Y\) follows a Gaussian distribution
with mean \(\mu_X + \mu_Y\)
and variance \(\sigma_X^2 + \sigma_Y^2\).

Let \(X_1, \ldots, X_n\) be \(n\) r.v. where \(\forall 1 \leq i \leq n\),
\(X_i\) follows a Gaussian distributions with mean \(\mu_i\) and variance \(\sigma_i^2\).

**Proposition:** if \(X_1, \ldots, X_n\) are mutually independent, \(X_1 + \ldots + X_n\) follows a Gaussian distribution
with mean \(\mu_1 + \ldots + \mu_n\)
and variance \(\sigma_1^2 + \ldots + \sigma_n^2\).
Example: squirrels in Washington Square Park

Let $X$ be the weight of a squirrel picked up randomly in WSP.

**Probability density function:** We suppose that $X$ has a PDF given by

$$
\tilde{P}(X = t) : \mathbb{R} \to [0, 1] \\
t \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = N_{2,0.5}(t)
$$

where $\sigma = 0.5$, $\mu = 2$.

We pick up randomly 2 squirrels in WSP; let $Z$ be the sum of their weight.

$Z$ follows a Gaussian distribution of mean $\mu_Z = 4$
and variance $\sigma^2_Z = 2 \times 0.5^2 = 0.5$
and standard deviation $\sigma_Z = \frac{\sqrt{2}}{2} \simeq 0.7$. 
Exercise

A firm constructs steel sheets meant to be stacked.

The width of a constructed sheet follows a normal (Gaussian) law of mean $\mu = 0.3 \, mm$ and standard deviation $\sigma = 0.1 \, mm$.

1. Give the probability that the width of a sheet is less than $0.35 \, mm$, and the probability that it is between $0.25$ and $0.35 \, mm$.

2. One chooses randomly $n$ sheets and stack them. Let $Z$ be the width of the obtained stack. Give the law of $Z$, its mean and its standard deviation.

3. Let $n = 20$. Give the probability for the width of the stack to be between $5.8 \, mm$ and $6.2 \, mm$. 
Exercise: solution

1. Let $X$ be the width of a sheet picked up randomly.

\[ P(X \leq 0.35) = \int_{-\infty}^{0.35} N_{0.3,0.1}(u)du = \text{normcdf}(0.35,0.3,0.1) \approx 0.6015 \]

\[ P(X \leq 0.25) = \int_{-\infty}^{0.25} N_{0.3,0.1}(u)du = \text{normcdf}(0.25,0.3,0.1) \approx 0.3085 \]

\[ P(0.25 \leq X \leq 0.35) = P(X \leq 0.35) - P(X \leq 0.25) \approx 0.3830 \]

2. We index the sheets of the stack from 1 to $n$, let $X_i$ be the width of the $i$-th sheet; it follows $N_{0.3,0.1}$.

Thus $Z = X_1 + \ldots + X_n$ follows a normal law with:

\[ \text{Exp}(Z) = n\text{Exp}(X_i) = 0.3n, \quad \text{Var}(Z) = n\text{Var}(X_i) = n(\text{Std}(X_i))^2 = 0.01n. \]

3. $Z$ follows a normal law of parameters:

\[ \text{Exp}(Z) = 0.6, \quad \text{Var}(Z) = 0.2. \]

Thus $P(5.8 \leq Z \leq 6.2) = P(Z \leq 6.2) - P(Z \leq 5.8)$

\[ = \text{normcdf}(6.2,6,\sqrt{0.2}) - \text{normcdf}(5.8,6,\sqrt{0.2}) \]