Advanced statistics and modelling

2020. február 12.

Technical information
Contacts of the teachers

- **Gergely Palla** / pallag@hal.elte.hu
  Dept. of Biological Physics, Room 3.90.
- **Péter Pollner** / pollner@elte.hu
  Dept. of Biological Physics, Room 3.90.
- **László Oroszlány** / oroszlanyl@gmail.com
  Dept. of Physics of Complex Systems, Room 5.79
- **Zoltán Kaufmann** / kaufmann@complex.elte.hu
  Dept. of Physics of Complex Systems, Room 5.53
- **Illés J. Farkas** / fij@hal.elte.hu
  Quant AVP at Citi MQA,
  Dept. of Biological Physics, Room 3.90

Grading:

- 2 written tests,
- activity during the lectures affects the mark positively.

Course material:

- all slides, jupyter notebooks, etc. are going to be uploaded to the ELTE moodle site of the course.

Schedule

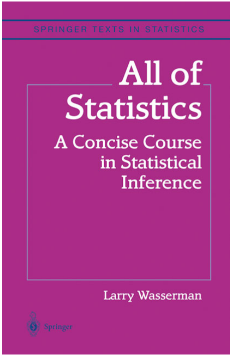
- **Sample space, Probability, Variable and Distribution**
- **Expectation, Inequalities, Convergence**
- **Models, Inference, Learning**
- **Empirical PDF, CDF, Smoothing, Binning**
- **Bootstrap, Maximum Likelihood, Hypothesis testing**
- **Regression, Inference About Independence**
- **Extreme Statistics, Posthoc Analysis**
- **Hierarchical Bayesian models**
- **Classical chaos, Ljapunov exponents, Frobenius Perron operators**
- **Random Matrices and Eigenvalue Statistics**
- **Statistics in fincance**
- **Stochastic Processes with applications in finance**

Schedule

Advanced statistics and modelling 2020 spring		
1. week	II.. 11. kedd Gergely Palla	Probability, variable, distribution
2. week	II.. 18. kedd Gergely Palla	Expectation, inequalities, convergence
3. week	II.. 25. kedd Gergely Palla	Statistical inference, estimating the CDF, PDF, smoothing
4. week	III.. 3. kedd Gergely Palla	Bootstrap, Maximum Likelihood, Hypothesis testing
5. week	III.. 10. kedd Péter Pollner	Regression, inference about independence
6. week	III.. 17. kedd Péter Pollner	Extreme statistics, posthoc analysis
7. week	III.. 24. kedd	!! MIDTERM TEST !!
8. week	III.. 31. kedd Illés Farkas	Statistics in finance
9. week	IV.. 7. kedd Illés Farkas	Stochastic processes with applications in finance
10. week	IV.. 14. kedd Holiday	
11. week	IV.. 21. kedd Péter Pollner	Bayesian models
12. week	IV.. 28. kedd Zoltán Kaufmann	Classical chaos, Ljapunov exponent, Frobenius-Perron
13. week	V.. 5. kedd László Oroszlány	Random matrices and eigenvalue statistics
14. week	V.. 12. kedd	!! FINAL TEST !!

Recommended literature

Larry Wasserman: **All of Statistics**
A Concise Course in Statistical Inference



Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

SAMPLE SPACE, PROBABILITY, VARIABLE AND DISTRIBUTION

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

SAMPLE SPACE

Sample space and events

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Sample space

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ is the set of possible outcomes of an experiment. The points ω_i in Ω are called **sample outcomes** or **realisations**.

Events

A subset $A \subset \Omega$ is called an **event**.

- True event: $A = \Omega$, (always true).
- Null event: $A = \emptyset$, (always false).

Examples

- Coin tossing: $\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \ldots,), \ \ \omega_i \in \{H, T\}\}$
  Event A can be e.g., that the first head appears on the third toss:
  $A = \{(\omega_1, \omega_2, \omega_3, \ldots,) : \omega_1 = T, \omega_2 = T, \omega_3 = H, \omega_i \in \{H, T\}\}$.
- Let ω be the outcome of a measurement of some physical quantity, e.g., temperature. Then $\Omega = [-\infty, \infty]$.
  Event A can be e.g., that he measurement is between 10 and 20: $A = [10, 20]$.

Sample space and events

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Sample space

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ is the set of possible outcomes of an experiment. The points ω_i in Ω are called **sample outcomes** or **realisations**.

Events

A subset $A \subset \Omega$ is called an **event**.

- True event: $A = \Omega$, (always true).
- Null event: $A = \emptyset$, (always false).

Examples

- Coin tossing: $\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \ldots,), \ \ \omega_i \in \{H, T\}\}$
  Event A can be e.g., that the first head appears on the third toss:
  $A = \{(\omega_1, \omega_2, \omega_3, \ldots,) : \omega_1 = T, \omega_2 = T, \omega_3 = H, \omega_i \in \{H, T\}\}$.
- Let ω be the outcome of a measurement of some physical quantity, e.g., temperature. Then $\Omega = [-\infty, \infty]$.
  Event A can be e.g., that he measurement is between 10 and 20: $A = [10, 20]$.

Sample space and events

Sample space

$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ is the set of possible outcomes of an experiment. The points ω_i in Ω are called **sample outcomes** or **realisations**.

Events

A subset $A \subset \Omega$ is called an **event**.

- True event: $A = \Omega$, (always true).
- Null event: $A = \varnothing$, (always false).

Examples

- Coin tossing: $\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \dots,), \;\; \omega_i \in \{H, T\}\}$
  Event A can be e.g., that the first head appears on the third toss:
  $A = \{(\omega_1, \omega_2, \omega_3, \dots,) : \omega_1 = T, \omega_2 = T, \omega_3 = H, \omega_i \in \{H, T\}\}$.
- Let ω be the outcome of a measurement of some physical quantity, e.g., temperature. Then $\Omega = [-\infty, \infty]$.
  Event A can be e.g., that he measurement is between 10 and 20: $A = [10, 20]$.

Sample space and events

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Sample space

$\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ is the set of possible outcomes of an experiment. The points ω_i in Ω are called **sample outcomes** or **realisations**.

Events

A subset $A \subset \Omega$ is called an **event**.

- True event: $A = \Omega$, (always true).
- Null event: $A = \emptyset$, (always false).

Examples

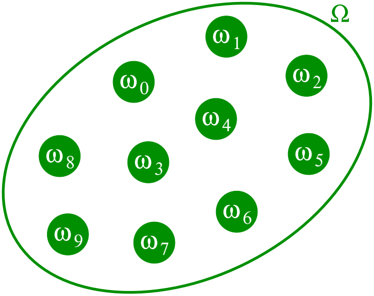
- Coin tossing: $\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \ldots,), \;\; \omega_i \in \{H, T\}\}$
  Event A can be e.g., that the first head appears on the third toss:
  $A = \{(\omega_1, \omega_2, \omega_3, \ldots,) : \omega_1 = T, \omega_2 = T, \omega_3 = H, \omega_i \in \{H, T\}\}$.
- Let ω be the outcome of a measurement of some physical quantity, e.g., temperature. Then $\Omega = [-\infty, \infty]$.
  Event A can be e.g., that he measurement is between 10 and 20: $A = [10, 20]$.

Operations between events

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Operations between events

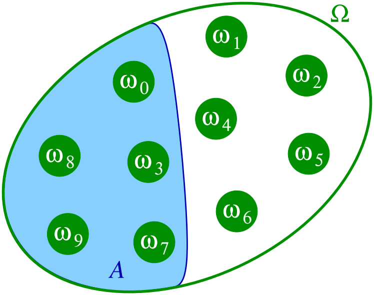
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

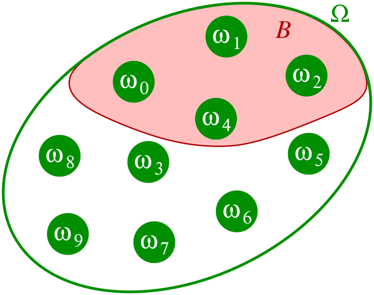
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

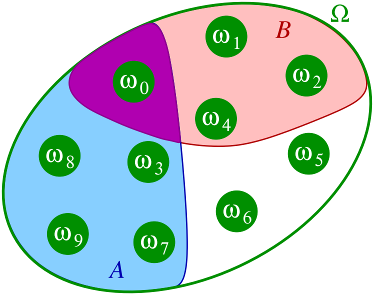
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

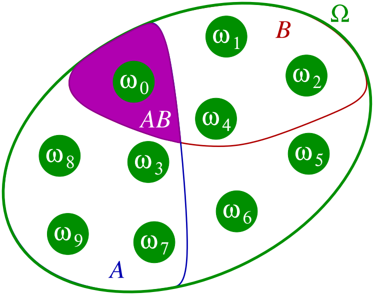
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Sample space, Probability, Variable and Distribution

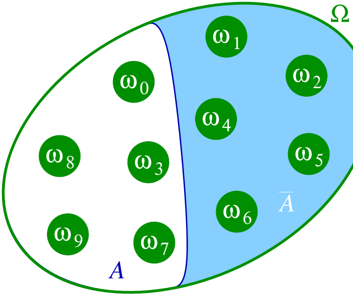
Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Operations between events

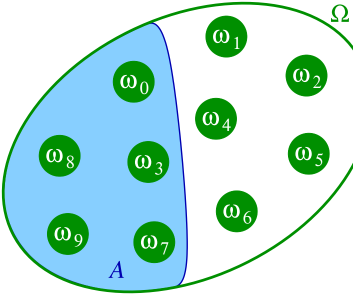
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^{c}: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

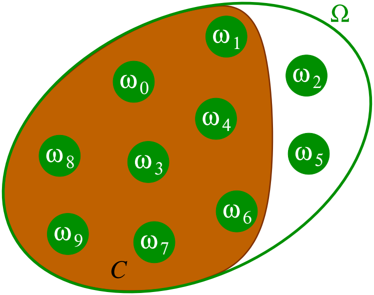
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

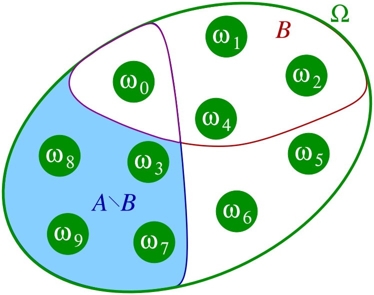
- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^c: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operations between events

Operations between events

- Union, $A \cup B$: at least one of them is true (OR): $A + B$.
- Intersection, $A \cap B$: both of them are true (AND): AB.
- Complement, A^{c}: it is false (NEGATION): $\overline{A}$.
- Inclusion: $A \subset C$.
- Set difference, $A \setminus B$: A is true and B is false: $A \setminus B = A \cap \overline{B}$.



Operational identities

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Operational identities

„AND"

- $A \cap B = B \cap A$
- $A \cap A = A$
- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cap \overline{A} = \varnothing$
- $A \cap \varnothing = \varnothing$
- $A \cap \Omega = A$

„OR"

- $A \cup B = B \cup A$
- $A \cup A = A$
- $A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cup \overline{A} = \Omega$
- $A \cup \varnothing = A$
- $A \cup \Omega = \Omega$

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (A \cap B) = A$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Partition

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Mutually exclusive events

A and B are mutually exclusive events if $A \cap B = \emptyset$

Partition

$A_1, A_2, \dots, A_n$ is a partition of Ω, if for all $k = 1, \dots, n$

a) $A_k \neq \emptyset$,

b) $A_j \cap A_k = \emptyset$ if $j \neq k$,

c) $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$.

Partition

Mutually exclusive events

A and B are mutually exclusive events if $A \cap B = \varnothing$

Partition

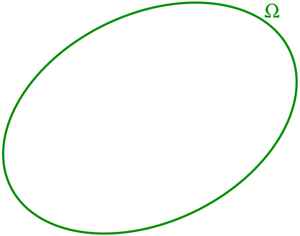
$A_1, A_2, \ldots, A_n$ is a partition of Ω, if for all $k = 1, \ldots, n$

a) $A_k \neq \varnothing$,

b) $A_j \cap A_k = \varnothing$ if $j \neq k$,

c) $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$.

Illustration:



Partition

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Mutually exclusive events

A and B are mutually exclusive events if $A \cap B = \emptyset$

Partition

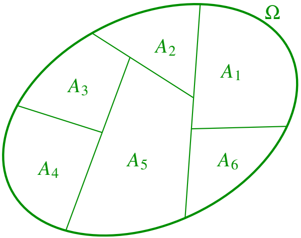
$A_1, A_2, \ldots, A_n$ is a partition of Ω, if for all $k = 1, \ldots, n$

a) $A_k \neq \emptyset$,

b) $A_j \cap A_k = \emptyset$ if $j \neq k$,

c) $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$.

Illustration:



Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

PROBABILITY

Probability

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Definition of probability (Kolmogorov)

Given a sample space Ω, a function $P(A)$ over the subsets of Ω is a **probability distribution** or a **probability measure** if

K1 $0 \leq P(A) \leq 1 \; \forall A \subset \Omega$,

K2 $P(\Omega) = 1$,

K3 If $A_1, A_2, \ldots$ are mutually exclusive, then

$$P\left(\bigcup_{i=1}^{\infty} A_k\right) = \sum_{i=1}^{\infty} P(A_i).$$

Probability

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Definition of probability (Kolmogorov)

Given a sample space Ω, a function $P(A)$ over the subsets of Ω is a **probability distribution** or a **probability measure** if

K1 $0 \leq P(A) \leq 1 \ \forall A \subset \Omega$,

K2 $P(\Omega) = 1$,

K3 If $A_1, A_2, \ldots$ are mutually exclusive, then

$$P\left(\bigcup_{i=1}^{\infty} A_k\right) = \sum_{i=1}^{\infty} P(A_i).$$

Probability

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Definition of probability (Kolmogorov)

Given a sample space Ω, a function $P(A)$ over the subsets of Ω is a **probability distribution** or a **probability measure** if

K1 $0 \leq P(A) \leq 1 \ \forall A \subset \Omega$,

K2 $P(\Omega) = 1$,

K3 If $A_1, A_2, \ldots$ are mutually exclusive, then

$$P\left(\bigcup_{i=1}^{\infty} A_k\right) = \sum_{i=1}^{\infty} P(A_i).$$

Probability identities

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

The probability of the false event $\varnothing$ is 0.

$$\begin{aligned} A \cup \varnothing = A, &\qquad A \cap \varnothing = \varnothing \\ P(A) = P(A \cup \varnothing) &\underset{K3}{=} P(A) + P(\varnothing) \\ P(\varnothing) &= 0 \end{aligned}$$

For a partition $A_1, A_2, \ldots, A_n$ the sum of the probabilities is $\sum_i P(A_i) = 1$.

- $A_i \cap A_j = \varnothing$ and $\bigcup_{i=1}^{n} A_i = \Omega$.
- Since $P(\Omega) = 1$, according to (K3) we obtain $\sum_i P(A_i) = 1$.

The probability of the complement $\overline{A}$ is $P(\overline{A}) = 1 - P(A)$.

(A and $\overline{A}$ together form a partition.)

Probability identities

The probability of the false event $\varnothing$ is 0.

$$A \cup \varnothing = A, \qquad A \cap \varnothing = \varnothing$$

$$\begin{aligned} P(A) = P(A \cup \varnothing) &\underset{K3}{=} P(A) + P(\varnothing) \\ P(\varnothing) &= 0 \end{aligned}$$

For a partition $A_1, A_2, \ldots, A_n$ the sum of the probabilities is $\sum_i P(A_i) = 1$.

- $A_i \cap A_j = \varnothing$ and $\bigcup_{i=1}^{n} A_i = \Omega$.
- Since $P(\Omega) = 1$, according to (K3) we obtain $\sum_i P(A_i) = 1$.

The probability of the complement $\overline{A}$ is $P(\overline{A}) = 1 - P(A)$.

(A and $\overline{A}$ together form a partition.)

Probability identities

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Probability of a subset

If A is included in B, i.e., $A \subset B$, then $P(B \setminus A) = P(B) - P(A)$.

- Since $A \subset B$, we can write B as $B = A \cup (B \setminus A)$.
- Also $A(B \setminus A) = \varnothing$, thus according to (K3) $P(B) = P(A) + P(B \setminus A)$.

Probability of the union

For any events A and B the probability of their union is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- Since $A \cup B = A \cup (B \setminus A \cap B)$ and $A(B \setminus A \cap B) = \varnothing$, according to (K3) $P(A \cup B) = P(A) + P(B \setminus A \cap B)$.
- Since $A \cap B \subset B$, according to the previous result $P(B \setminus A \cap B) = P(B) - P(A \cap B)$, which substituted back yields the end result.

Probability identities

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Limit theorems

L I. If $A_1, A_2, \ldots$ is an infinite series of events where $A_n \supset A_{n+1}$ for all n, and $\bigcap_{i=1}^{\infty} A_i = A$, then $\lim_{n\to\infty} P(A_n) = P(A)$.

L II. If $A_1, A_2, \ldots$ is an infinite series of events where $A_n \subset A_{n+1}$ for all n, and $\bigcup_{i=1}^{\infty} A_i = A$, then $\lim_{n\to\infty} P(A_n) = P(A)$.

Conditional probability

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Conditional probability

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Conditional probability

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Conditional probability

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Conditional probability

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Conditional probability

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Conditional probability

The conditional probability of event A given event B is defined as

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)},$$

(where we assumed that $P(B) > 0$).

Properties:

- $0 \leq P(A \mid B) \leq 1$,
- $P(B \mid B) = 1$,
- If $A_1, A_2, \ldots,$ are pairwise disjoint, then $P(\bigcup_i A_i \mid B) = \sum_i P(A_i \mid B)$.
- The $P(A \mid B)$ function over the subsets A in Ω is also a probability distribution.
- $P(A \mid B)P(B) = P(A \cap B) = P(B \mid A)P(A)$.

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Multiplication rule

Multiplication rule

For any events $A_1, A_2, \ldots, A_n$

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(A_1 \mid A_2 \cap \cdots \cap A_n) P(A_2 \mid A_3 \cap \cdots \cap A_n) \cdots \\ & \cdots P(A_{n-1} \mid A_n) P(A_n) = \\ & P(A_n) \prod_{i=1}^{n-1} P(A_i \mid A_{i+1} \cap \cdots \cap A_n) \end{aligned}$$

Proof: by applying the multiplication rule for two events recursively, e.g.,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \mid A_2 \cap A_3) P(A_2 \cap A_3) = P(A_1 \mid A_2 \cap A_3) P(A_2 \mid A_3) P(A_3),$$

etc.

Multiplication rule

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Multiplication rule

For any events $A_1, A_2, \ldots, A_n$

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(A_1 \mid A_2 \cap \cdots \cap A_n) P(A_2 \mid A_3 \cap \cdots \cap A_n) \cdots \\ &\cdots P(A_{n-1} \mid A_n) P(A_n) = \\ &P(A_n) \prod_{i=1}^{n-1} P(A_i \mid A_{i+1} \cap \cdots \cap A_n) \end{aligned}$$

Proof: by applying the multiplication rule for two events recursively, e.g.,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \mid A_2 \cap A_3) P(A_2 \cap A_3) = P(A_1 \mid A_2 \cap A_3) P(A_2 \mid A_3) P(A_3),$$

etc.

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Law of total probability

Law of total probability

If $A_1, A_2, \ldots, A_n$ provide a partition and $P(A_i) > 0\ \forall i$, then for any event B

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i).$$

- Since the A_i are pairwise disjoint, so are $B \cap A_i$.
- Based on $\bigcup_i A_i = \Omega$, the union of the events $B \cap A_i$ is

$$(B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = B \cap \Omega = B.$$

- According to (K3) $P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i)$.
- Multiplication rule for two events $P(B \cap A_i) = P(B \mid A_i) P(A_i)$.

$$\rightarrow P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Law of total probability

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Law of total probability

If $A_1, A_2, \ldots, A_n$ provide a partition and $P(A_i) > 0\ \forall i$, then for any event B

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i).$$

- Since the A_i are pairwise disjoint, so are $B \cap A_i$.
- Based on $\bigcup_i A_i = \Omega$, the union of the events $B \cap A_i$ is

$$(B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = B \cap \Omega = B.$$

- According to (K3) $P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i)$.
- Multiplication rule for two events $P(B \cap A_i) = P(B \mid A_i) P(A_i)$.

$$\rightarrow P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Law of total probability

Law of total probability

If $A_1, A_2, \ldots, A_n$ provide a partition and $P(A_i) > 0\ \forall i$, then for any event B

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i).$$

- Since the A_i are pairwise disjoint, so are $B \cap A_i$.
- Based on $\bigcup_i A_i = \Omega$, the union of the events $B \cap A_i$ is

$$(B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = B \cap \Omega = B.$$

- According to (K3) $P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i)$.
- Multiplication rule for two events $P(B \cap A_i) = P(B \mid A_i) P(A_i)$.

$$\rightarrow P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Law of total probability

Law of total probability

If $A_1, A_2, \ldots, A_n$ provide a partition and $P(A_i) > 0\ \forall i$, then for any event B

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i).$$

- Since the A_i are pairwise disjoint, so are $B \cap A_i$.
- Based on $\bigcup_i A_i = \Omega$, the union of the events $B \cap A_i$ is

$$(B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = B \cap \Omega = B.$$

- According to (K3) $P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i)$.
- Multiplication rule for two events $P(B \cap A_i) = P(B \mid A_i) P(A_i)$.

$$\rightarrow P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Law of total probability

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Law of total probability

If $A_1, A_2, \ldots, A_n$ provide a partition and $P(A_i) > 0\ \forall i$, then for any event B

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i).$$

- Since the A_i are pairwise disjoint, so are $B \cap A_i$.
- Based on $\bigcup_i A_i = \Omega$, the union of the events $B \cap A_i$ is

$$(B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) = B \cap (A_1 \cup A_2 \cup \cdots \cup A_n) = B \cap \Omega = B.$$

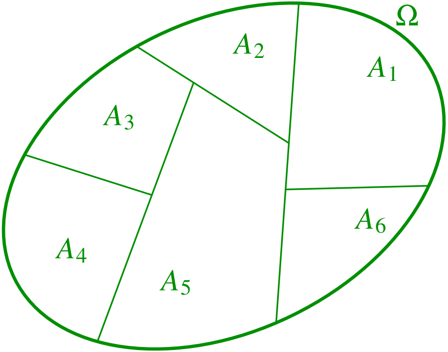
- According to (K3) $P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i)$.
- Multiplication rule for two events $P(B \cap A_i) = P(B \mid A_i) P(A_i)$.

$$\rightarrow P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

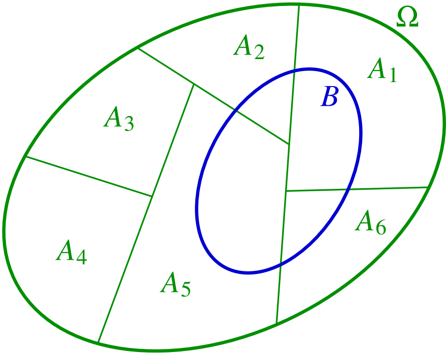
Law of total probability

Illustration:



Law of total probability

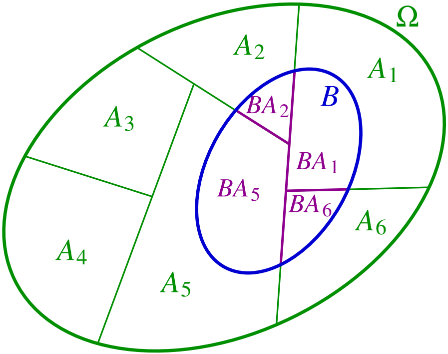
Illustration:



Law of total probability

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Illustration:



Bayes' theorem

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Bayes' theorem

Let $A_1, A_2, \ldots, A_n$ be a partition where $P(A_i) > 0 \; \forall i$. If $P(B) > 0$ for any event B, then

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Proof:

- According to the definition of the conditional probability $P(A_k \mid B)P(B) = P(B \mid A_k)P(A_k)$.
- We divide by $P(B)$ and use the law of total probability,

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{P(B)} = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Bayes' theorem

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Bayes' theorem

Let $A_1, A_2, ..., A_n$ be a partition where $P(A_i) > 0 \; \forall i$. If $P(B) > 0$ for any event B, then

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Proof:

- According to the definition of the conditional probability $P(A_k \mid B)P(B) = P(B \mid A_k)P(A_k)$.
- We divide by $P(B)$ and use the law of total probability,

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{P(B)} = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Bayes' theorem

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Bayes' theorem

Let $A_1, A_2, \ldots, A_n$ be a partition where $P(A_i) > 0 \; \forall i$. If $P(B) > 0$ for any event B, then

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Proof:

- According to the definition of the conditional probability $P(A_k \mid B)P(B) = P(B \mid A_k)P(A_k)$.
- We divide by $P(B)$ and use the law of total probability,

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{P(B)} = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Independence of events

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Independent events

Events A and B are independent if

$$P(A \cap B) = P(A)P(B).$$

- According to the definition the conditional probability of A given B in case of independence is

$$P(A \mid B) = P(A \cap B)/P(B) = \frac{P(A)P(B)}{P(B)} = P(A).$$

→ This can also be the alternative definition of independence.

Independence of events

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Independent events

Events A and B are independent if

$$P(A \cap B) = P(A)P(B).$$

- According to the definition the conditional probability of A given B in case of independence is

$$P(A \mid B) = P(A \cap B)/P(B) = \frac{P(A)P(B)}{P(B)} = P(A).$$

$\rightarrow$ This can also be the alternative definition of independence.

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

RANDOM VARIABLE, CDF and PDF

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Random variable

Random variable

A **random variable** is usually a mapping $X : \Omega \longrightarrow \mathbb{R}$ assigning a real number $X(\omega)$ to each event. However in general a random variable can map also as

$$X(\omega) : \Omega \longrightarrow \begin{cases} \mathbb{N} \\ \mathbb{R} \\ \mathbb{C} \\ \mathbb{R}^n \end{cases}$$

Cumulative distribution function

Cumulative distribution function (CDF)

The **cumulative distribution function** of a random variable X denoted by $F_X(x)$ is defined as

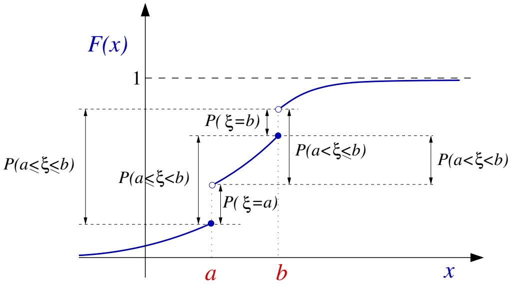
$$F_X(x) := P(X < x) = P\left(\{\omega \in \Omega\} : X(\omega) < x\right)$$

Properties:

- If $x_1 < x_2$ then $F(x_1) \leq F(x_2)$,
  (because event $X < x_1$ includes $X < x_2$).
- $\lim\limits_{x \to -\infty} F(x) = 0$ and $\lim\limits_{x \to \infty} F(x) = 1$.
  (These are the consequences of L I. and L II. limit theorems).
- $F(x)$ is continuous from the left:
  ha $x_1 < x_2 \cdots < x_i < \cdots$ and $\lim\limits_{n \to \infty} x_n = x$ then $\lim\limits_{n \to \infty} F(x_n) = F(x)$.
  (This comes from the limit theorem L II.)

Cumulative distribution function

Illustration:



Point mass function

Point mass function

- A random variable is **discrete** if its values $X(\omega) = x$ can take up finite or countable many different values.

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Point mass function

Point mass function

- A random variable is **discrete** if its values $X(\omega) = x$ can take up finite or countable many different values.
- Discrete random variables can be also characterised by their **point mass function** $f_X(x)$. By denoting the i-th discrete value $X(\omega)$ can take as x_i, the $f_X(x)$ can be defined as

$$f_X(x_i) := P(X(\omega) = x_i) = P(\{\omega_j\} : X(\omega_i) = x_i).$$

Point mass function

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Point mass function

- A random variable is **discrete** if its values $X(\omega) = x$ can take up finite or countable many different values.
- Discrete random variables can be also characterised by their **point mass function** $f_X(x)$. By denoting the i-th discrete value $X(\omega)$ can take as x_i, the $f_X(x)$ can be defined as

$$f_X(x_i) := P(X(\omega) = x_i) = P(\{\omega_j\} : X(\omega_i) = x_i).$$

- The relation between the CDF and the point mass function can be written as

$$F(x) = \sum_{i:\, x_i < x} f_X(x_i) = \sum_{i:\, x_i < x} P(\{\omega_j\} : X(\omega_i) = x_i).$$

Probability density function

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Probability density function

- A random variable is **continuous**, if there is a $\rho(x) \geq 0$ function fulfilling

$$F(b) - F(a) = P(a \leq X < b) = P(a < X < b) = \int_a^b \rho(x) dx.$$

Probability density function

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Probability density function

- A random variable is **continuous**, if there is a $\rho(x) \geq 0$ function fulfilling

$$F(b) - F(a) = P(a \leq X < b) = P(a < X < b) = \int_a^b \rho(x)dx.$$

- Using $F(-\infty) = 0$ we can express the CDF as

$$F(x) = \int_{-\infty}^{x} \rho(x')dx'.$$

Probability density function

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Probability density function

- A random variable is **continuous**, if there is a $\rho(x) \geq 0$ function fulfilling

$$F(b) - F(a) = P(a \leq X < b) = P(a < X < b) = \int_a^b \rho(x)dx.$$

- Using $F(-\infty) = 0$ we can express the CDF as

$$F(x) = \int_{-\infty}^{x} \rho(x')dx'.$$

- The function $\rho(x)$ is called the **probability density function**.

Probability density function

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Properties of the PDF:

Probability density function

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Properties of the PDF:

- It is normalised:

$$\int_{-\infty}^{\infty} \rho(x)dx = F(\infty) - F(-\infty) = 1.$$

Probability density function

Properties of the PDF:

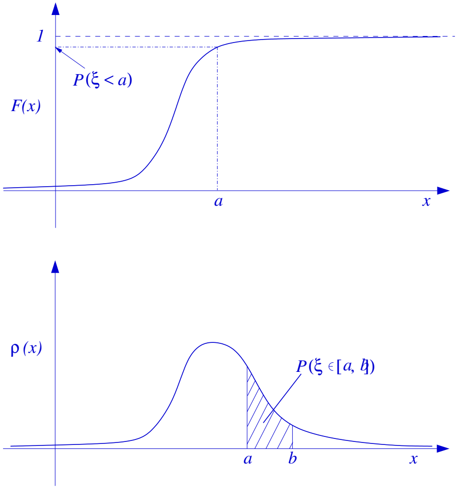
- It is normalised:

$$\int_{-\infty}^{\infty} \rho(x)dx = F(\infty) - F(-\infty) = 1.$$

- The probability that $X \in [a, b]$ can be expressed as

$$P(X \in [a, b]) = \int_{a}^{b} \rho(x)dx.$$

F(x) and $\rho(x)$
Illustration



Important distributions

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Discrete uniform distribution

X has a uniform distribution on $\{x_1, x_2, \ldots, x_n\}$ if its point mass function is given by

$$f_X(x_i) = \frac{1}{n}.$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Discrete uniform distribution

X has a uniform distribution on $\{x_1, x_2, \ldots, x_n\}$ if its point mass function is given by

$$f_X(x_i) = \frac{1}{n}.$$

Bernoulli distribution

Let X represent an experiment with possibly two different outcomes (e.g., coin flip), where $P(X = 1) = p$ and $P(X = 0) = 1 - p$. Then X has a **Bernoulli distribution**, with a point mass function for $x \in 0, 1$ written as

$$f_X(x) = p^x(1-p)^{1-x} = \begin{cases} p, & \text{if } x = 1, \\ 1-p, & \text{if } x = 0. \end{cases}$$

This is usually denoted as $X \sim$Bernoulli(p).

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Binomial and Geometric distributions

- Assume a coin flip with probability p for heads and probability $q = 1 - p$ for tails, repeated N times independently.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Binomial and Geometric distributions

- Assume a coin flip with probability p for heads and probability $q = 1 - p$ for tails, repeated N times independently.
- Let X count to the number of heads. X has a **Binomial distribution** with a point mass function

$$f_X(x = k) = P(X = k) = \binom{N}{k} p^k (1-p)^{N-k} = \binom{N}{k} p^k q^{N-k}.$$

This is usually denoted as $X \sim \text{Binomial}(N, p)$.

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Binomial and Geometric distributions

- Assume a coin flip with probability p for heads and probability $q = 1 - p$ for tails, repeated N times independently.
- Let X count to the number of heads. X has a **Binomial distribution** with a point mass function

$$f_X(x = k) = P(X = k) = \binom{N}{k} p^k (1-p)^{N-k} = \binom{N}{k} p^k q^{N-k}.$$

  This is usually denoted as $X \sim \text{Binomial}(N, p)$.
- Let Y count the number of flips needed until the first heads. Y has a **Geometrical distribution** with a point mass function

$$f_Y(y = k) = P(Y = k) = (1-p)^{k-1} p.$$

  This is usually denoted as $Y \sim \text{Geom}(p)$.

Important distributions

Discrete variables

Poisson distribution

X taking up non-negative integer values has a Poisson distribution if its point mass function can be written as

$$f_X(x=k) = P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

This is usually denoted as $X \sim$Poisson(λ).

Simeon Poisson

If we take a binomial distribution $X \sim \mathrm{Binom}(N,p)$ in the following limit:

$$\lim N = \infty, \quad \lim p = 0 \quad \lim_{\substack{N\to\infty \\ p\to 0}} pN = \lambda,$$

then its point mass function is converging to a Poisson distribution

$$f_X(x=k) = P(X=k) = \binom{N}{k} p^k (1-p)^{N-k} \xrightarrow{N\to\infty} \frac{(Np)^k}{k!}e^{-Np} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Poisson distribution

X taking up non-negative integer values has a Poisson distribution if its point mass function can be written as

$$f_X(x=k) = P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

This is usually denoted as $X \sim$Poisson(λ).

Simeon Poisson

If we take a binomial distribution $X \sim \mathrm{Binom}(N,p)$ in the following limit:

$$\lim N = \infty, \quad \lim p = 0 \quad \lim_{\substack{N\to\infty \\ p\to 0}} pN = \lambda,$$

then its point mass function is converging to a Poisson distribution

$$f_X(x=k) = P(X=k) = \binom{N}{k} p^k (1-p)^{N-k} \xrightarrow{N\to\infty} \frac{(Np)^k}{k!} e^{-Np} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Important distributions
Continuous variables

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Continuous uniform distribution

The continuous random variable $X(\omega) \in [x_1, x_2]$ has a uniform distribution if

$$\rho_X(x) = \begin{cases} \frac{1}{x_2 - x_1} & \text{if } x_1 \leq x \leq x_2 \\ 0 & \text{otherwise} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{if } x < x_1 \\ \frac{x - x_1}{x_2 - x_1} & \text{if } x_1 \leq x \leq x_2 \\ 1 & \text{if } x > x_2 \end{cases}$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Continuous uniform distribution

The continuous random variable $X(\omega) \in [x_1, x_2]$ has a uniform distribution if

$$\rho_X(x) = \begin{cases} \frac{1}{x_2 - x_1} & \text{if } x_1 \leq x \leq x_2 \\ 0 & \text{otherwise} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{if } x < x_1 \\ \frac{x - x_1}{x_2 - x_1} & \text{if } x_1 \leq x \leq x_2 \\ 1 & \text{if } x > x_2 \end{cases}$$

Exponential distribution

A continuous random variable X over the non-negative real numbers has an **Exponential distribution** if

$$F_X(x) = \begin{cases} 1 - e^{-\frac{x}{\lambda}} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \qquad \rho_X(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

This is usually denoted as $X \sim \text{Exp}(\lambda)$.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Gamma distribution

- A continuous random variable X over the non-negative real numbers has a **Gamma distribution**, usually denoted as $X \sim$Gamma(q, λ) if

$$\rho_X(x) = \frac{1}{\lambda^q \Gamma(q)} x^{q-1} e^{-\frac{x}{\lambda}},$$

where the Gamma function $\Gamma(z)$ is defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx.$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Gamma distribution

- A continuous random variable X over the non-negative real numbers has a **Gamma distribution**, usually denoted as X ~Gamma(q, λ) if

$$\rho_X(x) = \frac{1}{\lambda^q \Gamma(q)} x^{q-1} e^{-\frac{x}{\lambda}},$$

where the Gamma function $\Gamma(z)$ is defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx.$$

- Connections with the exponential distribution:
  - The exponential distribution corresponds to the special case of Gamma$(q = 1, \lambda)$.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Gamma distribution

- A continuous random variable X over the non-negative real numbers has a **Gamma distribution**, usually denoted as $X \sim$Gamma(q, λ) if

$$\rho_X(x) = \frac{1}{\lambda^q \Gamma(q)} x^{q-1} e^{-\frac{x}{\lambda}},$$

where the Gamma function $\Gamma(z)$ is defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx.$$

- Connections with the exponential distribution:
  - The exponential distribution corresponds to the special case of Gamma$(q = 1, \lambda)$.
  - The sum of n independent random variables $X_i \sim$Exp(λ) has a Gamma distribution Gamma(n, λ).

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Gamma distribution

- A continuous random variable X over the non-negative real numbers has a **Gamma distribution**, usually denoted as $X\sim$Gamma(q, λ) if

$$\rho_X(x) = \frac{1}{\lambda^q \Gamma(q)} x^{q-1} e^{-\frac{x}{\lambda}},$$

where the Gamma function $\Gamma(z)$ is defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx.$$

- Connections with the exponential distribution:
  - The exponential distribution corresponds to the special case of Gamma$(q = 1, \lambda)$.
  - The sum of n independent random variables $X_i\sim$Exp(λ) has a Gamma distribution Gamma(n, λ).

  → Thus, also the sum of independent $X_i\sim$Gamma(q_i, λ) has a distribution $\sum_{i=1}^n X_i\sim$Gamma$(\sum_{i=1}^n q_i, \lambda)$.

Sample space, Probability, Variable and Distribution

Sample space
- Events
- Operations between events
- Partition

Probability
- Kolmogorov's axioms
- Identities
- Conditional probability
- Bayes' theorem
- Independence

Random variable, CDF and PDF
- Random variable
- CDF
- Point mass function
- PDF
- Distributions

Normal distribution

A continuous random variable X has a **Normal** (Gaussian) distribution (usually denoted by $X \sim \mathrm{N}(\mu, \sigma)$) if

$$\rho_X(x) = \mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad F_X(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right],$$

A standard normal distribution is corresponding to $\mathrm{N}(\mu = 0, \sigma = 1)$.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition
Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence
Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

Normal distribution

A continuous random variable X has a **Normal** (Gaussian) distribution (usually denoted by $X \sim \mathrm{N}(\mu, \sigma)$) if

$$\rho_X(x) = \mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad F_X(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right],$$

A standard normal distribution is corresponding to $\mathrm{N}(\mu = 0, \sigma = 1)$.

χ^2 distribution

- A continuous random variable X over the non-negative real numbers has a χ^2 distribution with n degrees of freedom (usually denoted by $X \sim \chi^2_n$) if

$$\rho_X(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$$

- Connection with the Normal distribution:
  If $X_1, X_2, \ldots, X_n$ are independent standard Normal random variables, then $\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$.

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

t distribution and Cauchy distribution

- A continuous random variable X over the non-negative real numbers has a t **distribution** (also called as Student's t distribution) with n degrees of freedom (usually denoted by $X \sim t_n$) if

$$\rho_X(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}}.$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

t distribution and Cauchy distribution

- A continuous random variable X over the non-negative real numbers has a t **distribution** (also called as Student's t distribution) with n degrees of freedom (usually denoted by $X \sim t_n$) if

$$\rho_X(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}.$$

- Connection with the Normal distribution:
  If $X_1, X_2, \ldots, X_n$ and Y are independent standard Normal random variables, then the variable

$$Z := \frac{\sqrt{n}Y}{\sqrt{X_1^2 + X_2^2 + \cdots + X_n^2}} \quad \text{has a } t \text{ ditribution, } Z \sim t_n$$

Sample space, Probability, Variable and Distribution

Sample space
Events
Operations between events
Partition

Probability
Kolmogorov's axioms
Identities
Conditional probability
Bayes' theorem
Independence

Random variable, CDF and PDF
Random variable
CDF
Point mass function
PDF
Distributions

t distribution and Cauchy distribution

- A continuous random variable X over the non-negative real numbers has a t **distribution** (also called as Student's t distribution) with n degrees of freedom (usually denoted by $X \sim t_n$) if

$$\rho_X(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}.$$

- Connection with the Normal distribution:
  If $X_1, X_2, \ldots, X_n$ and Y are independent standard Normal random variables, then the variable

$$Z := \frac{\sqrt{n}Y}{\sqrt{X_1^2 + X_2^2 + \cdots + X_n^2}} \quad \text{has a } t \text{ ditribution, } Z \sim t_n$$

- The $n = 1$ special case of the t distribution corresponds to the **Cauchy distribution**, where

$$\rho_X(x) = \frac{1}{\pi} \frac{1}{1 + x^2}.$$