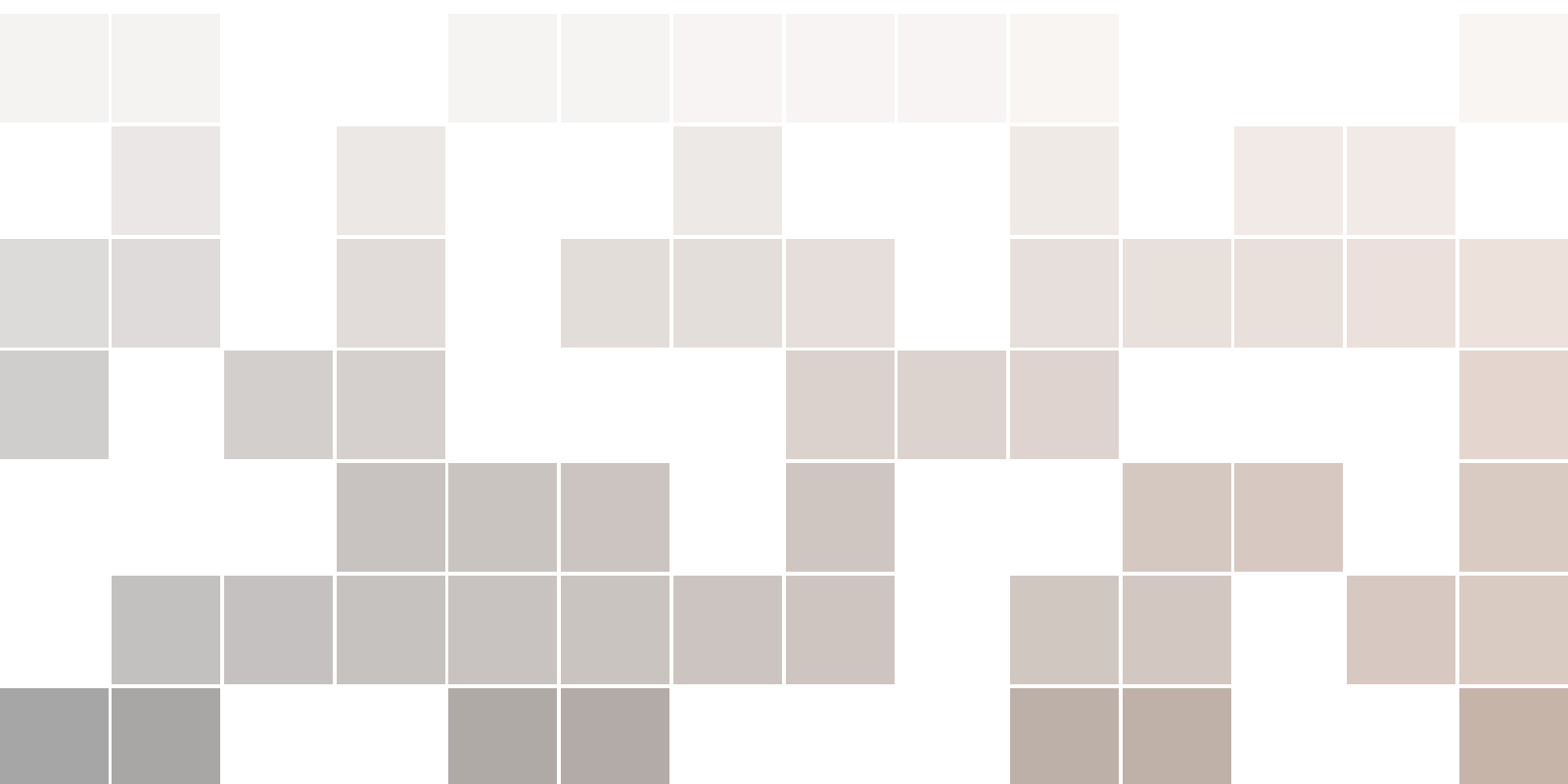




Scuole superiori

# Statistica

**OSB**



Copyright © 2023 OSB

PUBLISHED BY OSB

Licensed under the Creative Commons Attribution-NonCommercial 4.0 License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <https://creativecommons.org/licenses/by-nc-sa/4.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

*Latest version 22 settembre 2023*

# Indice

<b>1</b>	<b>Introduzione</b>	<b>9</b>
<b>2</b>	<b>Variabili casuali</b>	<b>11</b>
<b>2.1</b>	<b>Statistica univariata</b>	<b>11</b>
<b>2.2</b>	<b>Statistica multivariata</b>	<b>11</b>
2.2.1	Distribuzioni di probabilità (con variabili discrete)	11
2.2.2	Teorema di Bayes	12
2.2.3	Esempi nell'uso del teorema di Bayes	13
2.2.4	Indicatori sintetici	13
<b>3</b>	<b>Processi casuali</b>	<b>15</b>
<b>4</b>	<b>Introduzione all'intelligenza artificiale</b>	<b>17</b>
<b>4.1</b>	<b>Introduzione</b>	<b>17</b>
<b>4.2</b>	<b>Machine learning</b>	<b>17</b>
4.2.1	Supervised learning	17
4.2.2	Unsupervised learning	17
4.2.3	Reinforcement learning	18
<b>4.3</b>	<b>Deep learning</b>	<b>18</b>
	<b>Bibilografia</b>	<b>19</b>
	<b>Indice</b>	<b>21</b>
	<b>Appendices</b>	<b>23</b>
<b>A</b>	<b>Prima appendice</b>	<b>23</b>



## **Elenco delle figure**



## **Elenco delle tabelle**





# 1. Introduzione

- motivazione della necessità della gestione delle incertezze
- approcci alla statistica
- ...



## 2. Variabili casuali

### 2.1 Statistica univariata

- definizione variabili casuali; variabili a valori discreti o continui
- distribuzioni di probabilità: proprietà (non-negatività, unitarietà); densità e proprietà cumulata
- esempi di distribuzione di probabilità
- teoremi: teorema del limite centrale e teorema dei grandi numeri

### 2.2 Statistica multivariata

- distribuzioni congiunte, marginali e condizionali
- formula di Bayes
- indicatori sintetici: media, correlazione

#### 2.2.1 Distribuzioni di probabilità (con variabili discrete)

Siano date due variabili casuali  $X, Y$ , che possono assumere rispettivamente i valori  $x \in \{x_1, x_2, \dots, x_{N_x}\}$ ,  $y \in \{y_1, y_2, \dots, y_{N_y}\}$ .

**Definition 2.1 — Densità di probabilità congiunta**  $p_{X,Y}(x,y)$ . É una funzione che ha come argomenti i valori che possono assumere le variabili casuali e rappresenta la probabilità che si verifichino insieme gli eventi  $X = x, Y = y$ .

**Unitarietà della probabilità congiunta.** La somma delle probabilità di tutte le combinazioni degli eventi è uguale a 1, cioè

$$1 = \sum_{i_x=1}^{N_x} \sum_{i_y=1}^{N_y} p(x_{i_x}, y_{i_y}) \quad (2.1)$$

**Definition 2.2 — Densità di probabilità marginale**  $p_X(x)$ . É una funzione che ha come argomento i valori che può assumere la variabile casuale  $x$  e rappresenta la probabilità che si verifichi l'evento  $X = x$ , indipendentemente dal valore di  $Y$ .

**Proprietà marginale come somma parziale.** La proprietà marginale dell'osservazione  $X = x_i$ ,  $p(x_i)$ , è la somma delle probabilità congiunte con tutte le possibili osservazioni dell'altra

variabile  $Y$ , cioè

$$p(x_{i_x}) = \sum_{i_y=1}^{N_y} p(x_{i_x}, y_{i_y}) , \quad (2.2)$$

così che l'equazione 2.1 può essere riscritta

$$1 = \sum_{i_x=1}^{N_x} \sum_{i_y=1}^{N_y} p(x_{i_x}, y_{i_y}) = \sum_{i_x=1}^{N_x} p(x_{i_x}) , \quad (2.3)$$

dimostrando l'unitarietà della probabilità marginale, come densità di probabilità su  $x$ .

**Definition 2.3 — Densità di probabilità condizionale**  $p_{Y|X}(y|x)$ . É una funzione che ha come argomenti il valori che possono assumere la variabile casuale  $y$  e come “parametro” la variabile casuale  $x$  e rappresenta la probabilità che si verifichi l'evento  $Y = y$ , dato l'evento  $X = x$ .

Per ottenere la probabilità di osservare  $Y = y_{i_y}$ , condizionato all'osservazione di  $X = x_{i_x}$  bisogna “scalare” la probabilità congiunta  $p(x_{i_x}, y_{i_y})$  per la probabilità marginale di aver ottenuto  $X = x_{i_x}$ , cioè

$$p(y_{i_y}|x_{i_x}) = \frac{p(x_{i_x}, y_{i_y})}{p(x_{i_x})} . \quad (2.4)$$

■ **Osservazione 2.1** La probabilità condizionale  $p(y|x)$  è ben definita quando la proprietà marginale  $p(x) > 0$ , ossia la probabilità di osservare il valore  $x$  non è nullo. ■

■ **Osservazione 2.2** La condizione di normalizzazione per la probabilità condizionale si ottiene dividendo l'espressione 2.2 per  $p(x_{i_x})$ , ■

$$1 = \sum_{i_y=1}^{N_y} \frac{p(x_{i_x}, y_{i_y})}{p(x_{i_x})} = \sum_{i_y=1}^{N_y} p(y_{i_y}|x_{i_x}) . \quad (2.5)$$

## 2.2.2 Teorema di Bayes

É possibile riscrivere l'equazione 2.4 nella forma

$$p(x, y) = p(y|x)p(x) \quad (2.6)$$

o alternativamente

$$p(x, y) = p(x|y)p(y) . \quad (2.7)$$

**Theorem 2.1 — Teorema di Bayes.** Usando le equazioni 2.6, 2.7,

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y) , \quad (2.8)$$

è possibile scrivere

$$p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(y|x)p(x)}{p(y)} . \quad (2.9)$$

### 2.2.3 Esempi nell'uso del teorema di Bayes

### 2.2.4 Indicatori sintetici

Dopo aver raccolto le variabili casuali scalari  $X_i$  in una variabile casuale vettoriale  $\mathbf{X}$ , usando il formalismo matriciale

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \dots \\ X_n \end{bmatrix} \quad (2.10)$$

possiamo definire alcuni indicatori sintetici.

#### 2.2.4.1 Valore atteso (volgarmente chiamato media)

La media di una variabile casuale multivariata (o multidimensionale) viene definita come la media pesata di tutti i possibili valori  $\mathbf{x}_I$  della variabile casuale  $\mathbf{X}_I$ , pesati per il valore corrispondente della densità di probabilità

$$\mathbb{E}[\mathbf{x}] := \bar{\mathbf{X}} := \mu_X = \sum_I f(\mathbf{x}_I) \mathbf{x}_I . \quad (2.11)$$

#### 2.2.4.2 Covarianza

La covarianza viene definita come il valore atteso del “prodotto tensoriale” della deviazione della media con sé stesso, cioè

$$\mathbf{C}_{XX} := \mathbb{E}[(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T] . \quad (2.12)$$

Usando le proprietà della media, si può riscrivere la covarianza come

$$\begin{aligned} \mathbf{C}_{XX} &= \mathbb{E}[(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T] = \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \mathbb{E}[\mathbf{X}\bar{\mathbf{X}}^T] - \mathbb{E}[\bar{\mathbf{X}}\mathbf{X}^T] + \bar{\mathbf{X}}\bar{\mathbf{X}}^T = \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \mathbb{E}[\mathbf{X}]\bar{\mathbf{X}}^T - \bar{\mathbf{X}}\mathbb{E}[\mathbf{X}^T] + \bar{\mathbf{X}}\bar{\mathbf{X}}^T = \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \mathbb{E}[\mathbf{X}]\bar{\mathbf{X}}^T - \bar{\mathbf{X}}\mathbb{E}[\mathbf{X}^T] + \bar{\mathbf{X}}\bar{\mathbf{X}}^T = \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \bar{\mathbf{X}}\bar{\mathbf{X}}^T \end{aligned} \quad (2.13)$$



### 3. Processi casuali

- definizione dei processi casuali
- esempi di processi casuali
  - catene di Markov
  - random walk





## 4. Introduzione all'intelligenza artificiale

- introduzione intelligenza artificiale
- machine learning:
  - supervised learning: regressione, classificazione
  - unsupervised learning: clustering, riduzione della dimensionalità
  - reinforcement learning
- deep learning: neural networks

### 4.1 Introduzione

**Cosa si intende per intelligenza artificiale.**

**Nessun pasto è gratis – compromesso bias-varianza.** [https://it.wikipedia.org/wiki/Compromesso\\_bias-varianza](https://it.wikipedia.org/wiki/Compromesso_bias-varianza)

**La maledizione della dimensionalità.**

**Deep learning.**

### 4.2 Machine learning

#### 4.2.1 Supervised learning

Il supervised learning (o apprendimento supervisionato) è un paradigma che permette di allenare un modello.

##### 4.2.1.1 Regressione

La regressione consiste nell'approssimazione di funzioni continue.

**Regressione lineare.**

**Regressione lineare generalizzata.**

##### 4.2.1.2 Classificazione

La classificazione consiste nell'identificazione di una categoria alla quale appartiene un oggetto.

#### 4.2.2 Unsupervised learning

##### 4.2.2.1 Dimensionality reduction

La riduzione delle dimensioni di un problema consente di:

- ridurre la complessità del problema
- mantenendo solo le informazioni principali

- **Example 4.1 — Compressione immagini.**

■

#### **4.2.2.2 Clustering**

Il clustering è il raggruppamento di oggetti in insiemi che dimostrano caratteristiche simili.

#### **4.2.3 Reinforcement learning**

#### **4.3 Deep learning**

...

## Bibiliografia



# Indice analitico

## C

Clustering ..... 18

## D

Deep learning ..... 18

Dimensionality reduction.....17

## I

Intelligenza Artificiale ..... 17

## M

Machine learning ..... 17

## R

Reinforcement learning ..... 18

## S

Statistica

    Introduzione ..... 9

    Multivariata ..... 11

    Processi ..... 15

    Univariata ..... 11

    Variabili ..... 11

Supervised learning.....17

## T

Teorema

    di Bayes.....12

## U

Unsupervised learning ..... 17



## A. Prima appendice

...