

Probabilità e Processi Stocastici (455AA)

Lezione 1

Dario Trevisan

25/09/2023

Section 1

Introduzione al corso

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
- 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
- 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)
- 4 Variabili aleatorie Gaussiane e applicazioni (reali e vettoriali, PCA, Regressione lineare)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
- 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)
- 4 Variabili aleatorie Gaussiane e applicazioni (reali e vettoriali, PCA, Regressione lineare)
- 5 Processi stocastici a stati discreti (catene di Markov, processi di Markov a salti, esempi dalla teoria delle code)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
- 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)
- 4 Variabili aleatorie Gaussiane e applicazioni (reali e vettoriali, PCA, Regressione lineare)
- 5 Processi stocastici a stati discreti (catene di Markov, processi di Markov a salti, esempi dalla teoria delle code)
- 6 Processi a stati continui (gaussiani, ARIMA)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
- 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
- 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)
- 4 Variabili aleatorie Gaussiane e applicazioni (reali e vettoriali, PCA, Regressione lineare)
- 5 Processi stocastici a stati discreti (catene di Markov, processi di Markov a salti, esempi dalla teoria delle code)
- 6 Processi a stati continui (gaussiani, ARIMA)
- 7 Teoremi limite (Legge dei grandi numeri, Teorema Ergodico, Teorema limite centrale)

Argomenti

- 1 Calcolo delle probabilità (eventi, formula di Bayes, indipendenza)
 - 2 Variabili aleatorie (densità continue e discrete, congiunte, marginali)
 - 3 Indicatori caratteristici (media, varianza, covarianza, correlazione)
 - 4 Variabili aleatorie Gaussiane e applicazioni (reali e vettoriali, PCA, Regressione lineare)
 - 5 Processi stocastici a stati discreti (catene di Markov, processi di Markov a salti, esempi dalla teoria delle code)
 - 6 Processi a stati continui (gaussiani, ARIMA)
 - 7 Teoremi limite (Legge dei grandi numeri, Teorema Ergodico, Teorema limite centrale)
- Introdurremo inoltre le basi del linguaggio R.

Materiale didattico

Alla pagina web del corso

<http://people.dm.unipi.it/trevisan/it/courses/455AA.html>

trovate

- **Appunti**

Materiale didattico

Alla pagina web del corso

<http://people.dm.unipi.it/trevisan/it/courses/455AA.html>

trovate

- **Appunti**
- Raccolta di prove scritte anni precedenti

Materiale didattico

Alla pagina web del corso

<http://people.dm.unipi.it/trevisan/it/courses/455AA.html>

trovate

- **Appunti**
- Raccolta di prove scritte anni precedenti
- Note delle lezioni (slides annotate)

Materiale didattico

Alla pagina web del corso

<http://people.dm.unipi.it/trevisan/it/courses/455AA.html>

trovate

- **Appunti**
- Raccolta di prove scritte anni precedenti
- Note delle lezioni (slides annotate)
- RegISTRAZIONI

Orario lezioni

- Lunedì 10:30 - 13:30

Orario lezioni

- Lunedì 10:30 - 13:30
- Giovedì 11:30 - 13:30

Ricevimento

Ogni ??? ore 18-20 (su appuntamento, anche su Teams, contattatemi sempre via mail oppure Teams).

Modalità di esame

- Prova **scritta** con problemi sugli argomenti da 1 a 5 (inclusi).
Risoluzione analitica oppure (a richiesta) indicando opportuni comandi R.

Modalità di esame

- Prova **scritta** con problemi sugli argomenti da 1 a 5 (inclusi).
Risoluzione analitica oppure (a richiesta) indicando opportuni comandi R.
- Prova **orale** con domande principalmente sulla teoria (tutto il programma svolto).

Modalità di esame

- Prova **scritta** con problemi sugli argomenti da 1 a 5 (inclusi).
Risoluzione analitica oppure (a richiesta) indicando opportuni comandi R.
- Prova **orale** con domande principalmente sulla teoria (tutto il programma svolto).
- Ogni prova scritta superata permette di accedere ad una (e una sola) prova orale: non necessariamente quella immediatamente successiva, **purché nella medesima sessione** (il compitino vale per la sessione invernale).

Section 2

Capitolo 1: probabilità elementare

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete
- regola del **prodotto** (o della probabilità composta) e conseguenze

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete
- regola del **prodotto** (o della probabilità composta) e conseguenze
- problemi elementari tramite **diagrammi ad albero** e modello delle estrazioni da un'urna (senza rimpiazzo)

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete
- regola del **prodotto** (o della probabilità composta) e conseguenze
- problemi elementari tramite **diagrammi ad albero** e modello delle estrazioni da un'urna (senza rimpiazzo)
- **formula di Bayes**

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete
- regola del **prodotto** (o della probabilità composta) e conseguenze
- problemi elementari tramite **diagrammi ad albero** e modello delle estrazioni da un'urna (senza rimpiazzo)
- **formula di Bayes**
- stima bayesiana di un'ipotesi sulla base di dati osservati e metodo di **massima verosimiglianza**, cenni ai test statistici

Cosa vedremo in questo capitolo:

- la probabilità dal punto di vista **soggettivo**
- regola della **somma** (o additività) e conseguenze
- sistemi di **alternative** (finiti) e densità discrete
- regola del **prodotto** (o della probabilità composta) e conseguenze
- problemi elementari tramite **diagrammi ad albero** e modello delle estrazioni da un'urna (senza rimpiazzo)
- **formula di Bayes**
- stima bayesiana di un'ipotesi sulla base di dati osservati e metodo di **massima verosimiglianza**, cenni ai test statistici
- Introduzione ad **R** ed RStudio.

Cos'è la probabilità?

La probabilità misura il grado di fiducia che un soggetto attribuisce alla validità di una affermazione, avendo a disposizione una informazione parziale (che in generale non permette di dedurre la verità o la falsità dell'affermazione).

Quale soggetto?

Assegnate

- 1 una informazione, che indichiamo con I , nota e ritenuta vera,

è richiesto di misurare il grado di incertezza circa la validità di A , *sulla base di tutta e sola l'informazione I* , nel modo più razionale possibile.

Tale misura, detta la **probabilità di A sapendo I** si indica

$$P(A|I).$$

Assegnate

- 1 una informazione, che indichiamo con I , nota e ritenuta vera,
- 2 una affermazione, che indichiamo con A , che nella realtà può essere solo vera oppure falsa (quindi senza ambiguità),

è richiesto di misurare il grado di incertezza circa la validità di A , *sulla base di tutta e sola l'informazione I* , nel modo più razionale possibile.

Tale misura, detta la **probabilità di A sapendo I** si indica

$$P(A|I).$$

Proprietà elementari

Operazioni logiche tra affermazioni

Monotonia

Date due affermazioni A e B e l'informazione nota I , se A è **vera in qualsiasi situazione in cui B sia vera** (supponendo sempre vera I), allora

$$P(B|I) \leq P(A|I).$$

Regola della somma

- Date A , B e l'informazione nota I , se A e B non possono in nessun caso essere entrambe vere (supponendo I vera), allora

$$P(A \text{ oppure } B|I) = P(A|I) + P(B|I).$$

Regola della somma

- Date A , B e l'informazione nota I , se A e B non possono in nessun caso essere entrambe vere (supponendo I vera), allora

$$P(A \text{ oppure } B|I) = P(A|I) + P(B|I).$$

- A e B si dicono **incompatibili** (o mutuamente esclusivi) se non possono essere entrambe vere (rispetto ad una informazione I), ossia

$$P(A \text{ e } B|I) = 0.$$

Regola della somma

- Date A , B e l'informazione nota I , se A e B non possono in nessun caso essere entrambe vere (supponendo I vera), allora

$$P(A \text{ oppure } B|I) = P(A|I) + P(B|I).$$

- A e B si dicono **incompatibili** (o mutuamente esclusivi) se non possono essere entrambe vere (rispetto ad una informazione I), ossia

$$P(A \text{ e } B|I) = 0.$$

- Come calcolare $P(A \text{ oppure } B)$ in generale?

Alternativa semplice

Sistemi di alternative

Un **sistema di alternative** (rispetto ad una informazione I) è una famiglia $(A_i)_{i=1}^n$ di affermazioni (dette alternative)

- 1 a due a due incompatibili (o mutuamente esclusive) e

Sistemi di alternative

Un **sistema di alternative** (rispetto ad una informazione I) è una famiglia $(A_i)_{i=1}^n$ di affermazioni (dette alternative)

- 1 a due a due incompatibili (o mutuamente esclusive) e
- 2 tali che almeno una tra loro è sicuramente vera.

Sistemi di alternative

Un **sistema di alternative** (rispetto ad una informazione I) è una famiglia $(A_i)_{i=1}^n$ di affermazioni (dette alternative)

- 1 a due a due incompatibili (o mutuamente esclusive) e
 - 2 tali che almeno una tra loro è sicuramente vera.
- In breve, **una e una sola** tra le alternative è sicuramente vera (nota I).

Formula di decomposizione

Sia $(A_i)_{i=1}^n$ un sistema di alternative (rispetto all'informazione I) e sia B una (qualsiasi) affermazione.

Allora

$$P(B|I) = P(B \text{ e } A_1|I) + \dots + P(B \text{ e } A_n|I).$$

Densità discreta

Ad un sistema di alternative $(A_i)_{i=1}^n$ (rispetto all'informazione I) possiamo associare la collezione delle probabilità

$$(P(A_i|I))_{i=1}^n.$$

Densità Bernoulli

Densità Uniforme

Moda di una densità discreta

La moda indica l'alternativa **più probabile**, ossia

$$i_{\max} \in \arg \max \{P(A_i|I) : i \in \{1, \dots, n\}\}.$$

Regola del prodotto

Date affermazioni A , B e l'informazione nota I , vale

$$P(A \text{ e } B|I) = P(A|I)P(B|A, I).$$

formula di Kolmogorov per la probabilità condizionata:

$$P(B|A, I) = \frac{P(A \text{ e } B|I)}{P(A|I)}.$$

Formula di disintegrazione

Sia $(A_i)_{i=1}^n$ un sistema di alternative rispetto ad una informazione I . Allora, data una affermazione B (qualsiasi),

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

Notazione “proporzionale”

Data una densità discreta $(p_i)_{i=1}^n$ e una funzione $f(i)$ a valori positivi (non necessariamente $f(i) \leq 1$) scriviamo

$$p_i \propto f(i)$$

per dire che $p_i = cf(i)$, dove la costante c è data da

$$c = \left(\sum_i f(i) \right)^{-1}$$

per garantire che la somma delle p_i sia 1.

- Esempio: densità uniforme $p_i \propto 1$

Notazione “proporzionale”

Data una densità discreta $(p_i)_{i=1}^n$ e una funzione $f(i)$ a valori positivi (non necessariamente $f(i) \leq 1$) scriviamo

$$p_i \propto f(i)$$

per dire che $p_i = cf(i)$, dove la costante c è data da

$$c = \left(\sum_i f(i) \right)^{-1}$$

per garantire che la somma delle p_i sia 1.

- Esempio: densità uniforme $p_i \propto 1$
- Esercizio: determinare c se $p_i \propto i$ per $i = 1, 2, 3, 4$.

Diagrammi ad albero: Generalità

- Le regole di somma e prodotto (e le conseguenti formule di disintegrazione) forniscono strumenti utili per *analizzare* problemi di probabilità elementare, riducendoli a sotto-problemi più semplici tramite l'introduzione (anche ripetuta) di **sistemi di alternative**.

Diagrammi ad albero: Generalità

- Le regole di somma e prodotto (e le conseguenti formule di disintegrazione) forniscono strumenti utili per *analizzare* problemi di probabilità elementare, riducendoli a sotto-problemi più semplici tramite l'introduzione (anche ripetuta) di **sistemi di alternative**.
- È utile rappresentare l'analisi tramite **diagrammi ad albero** costruiti con il seguente algoritmo.

Costruzione dell'albero

- Si introduce un nodo **radice** (informazione iniziale, ad esempio indicata con Ω). Si iterano poi i seguenti passi:

Costruzione dell'albero

- Si introduce un nodo **radice** (informazione iniziale, ad esempio indicata con Ω). Si iterano poi i seguenti passi:
- ① si considera un nodo del grafo che sia una “foglia”, ossia senza archi uscenti, etichettato da una affermazione B ,

Costruzione dell'albero

- Si introduce un nodo **radice** (informazione iniziale, ad esempio indicata con Ω). Si iterano poi i seguenti passi:
 - 1 si considera un nodo del grafo che sia una “foglia”, ossia senza archi uscenti, etichettato da una affermazione B ,
 - 2 si sceglie un sistema di alternative $(A_i)_{i=1}^n$, e si introducono tanti nodi quante le alternative, etichettate appunto da esse,

Costruzione dell'albero

- Si introduce un nodo **radice** (informazione iniziale, ad esempio indicata con Ω). Si iterano poi i seguenti passi:
 - 1 si considera un nodo del grafo che sia una “foglia”, ossia senza archi uscenti, etichettato da una affermazione B ,
 - 2 si sceglie un sistema di alternative $(A_i)_{i=1}^n$, e si introducono tanti nodi quante le alternative, etichettate appunto da esse,
 - 3 si introducono archi uscenti dalla foglia (B) verso il nodo corrispondente a ciascuna alternativa (A_i),

Costruzione dell'albero

- Si introduce un nodo **radice** (informazione iniziale, ad esempio indicata con Ω). Si iterano poi i seguenti passi:
 - ① si considera un nodo del grafo che sia una “foglia”, ossia senza archi uscenti, etichettato da una affermazione B ,
 - ② si sceglie un sistema di alternative $(A_i)_{i=1}^n$, e si introducono tanti nodi quante le alternative, etichettate appunto da esse,
 - ③ si introducono archi uscenti dalla foglia (B) verso il nodo corrispondente a ciascuna alternativa (A_i),
 - ④ si pesa ciascun arco introdotto sopra con la probabilità

$$P(A_i|B, I),$$

dove I consiste della congiunzione di tutte le affermazioni nell'unico cammino (orientato) che collega l'informazione iniziale Ω a B .

Calcolo delle probabilità

Supponendo di aver costruito un albero, come calcolare $P(A|\Omega)$?

- Si effettuano ancora una volta i passaggi 1-4 aggiungendo ad ogni foglia l'alternativa semplice (A) (ed eventualmente la sua negazione) e la probabilità condizionata

Note:

Calcolo delle probabilità

Supponendo di aver costruito un albero, come calcolare $P(A|\Omega)$?

- Si effettuano ancora una volta i passaggi 1-4 aggiungendo ad ogni foglia l'alternativa semplice (A) (ed eventualmente la sua negazione) e la probabilità condizionata
- Si calcola il peso di ciascun cammino che porta da una foglia (A) verso la radice, **moltiplicando** le probabilità sugli archi (è una applicazione della regola del prodotto)

Note:

Calcolo delle probabilità

Supponendo di aver costruito un albero, come calcolare $P(A|\Omega)$?

- Si effettuano ancora una volta i passaggi 1-4 aggiungendo ad ogni foglia l'alternativa semplice (A) (ed eventualmente la sua negazione) e la probabilità condizionata
- Si calcola il peso di ciascun cammino che porta da una foglia (A) verso la radice, **moltiplicando** le probabilità sugli archi (è una applicazione della regola del prodotto)
- Si **sommano** i pesi di tutti i cammini così ottenuti (è una applicazione della regola della somma)

Note:

Calcolo delle probabilità

Supponendo di aver costruito un albero, come calcolare $P(A|\Omega)$?

- Si effettuano ancora una volta i passaggi 1-4 aggiungendo ad ogni foglia l'alternativa semplice (A) (ed eventualmente la sua negazione) e la probabilità condizionata
- Si calcola il peso di ciascun cammino che porta da una foglia (A) verso la radice, **moltiplicando** le probabilità sugli archi (è una applicazione della regola del prodotto)
- Si **sommano** i pesi di tutti i cammini così ottenuti (è una applicazione della regola della somma)

Note:

- 1 se un arco ha probabilità 0 allora i pesi dei cammini che lo percorrono sono 0, non serve contarli.

Calcolo delle probabilità

Supponendo di aver costruito un albero, come calcolare $P(A|\Omega)$?

- Si effettuano ancora una volta i passaggi 1-4 aggiungendo ad ogni foglia l'alternativa semplice (A) (ed eventualmente la sua negazione) e la probabilità condizionata
- Si calcola il peso di ciascun cammino che porta da una foglia (A) verso la radice, **moltiplicando** le probabilità sugli archi (è una applicazione della regola del prodotto)
- Si **sommano** i pesi di tutti i cammini così ottenuti (è una applicazione della regola della somma)

Note:

- 1 se un arco ha probabilità 0 allora i pesi dei cammini che lo percorrono sono 0, non serve contarli.
- 2 se è richiesta la probabilità $P(A|I)$ è l'informazione (cumulata) I è un nodo dell'albero, basta considerare solo l'albero con radice I

Estrazioni da un'urna (senza rimpiazzo)

- **Probabilità di estrarre una precisa sequenza ordinata** di $n \leq N$ palline colorate, di cui $r \leq R$ sono rosse e le rimanenti $b \leq B$ sono blu:

$$\frac{R(R-1) \cdot \dots \cdot (R-r+1) \cdot B(B-1) \cdot \dots \cdot (B-b+1)}{N(N-1) \cdot \dots \cdot (N-n+1)}.$$

- **Probabilità di estrarre una precisa sequenza ordinata** di $n \leq N$ palline colorate, di cui $r \leq R$ sono rosse e le rimanenti $b \leq B$ sono blu:

$$\frac{R(R-1) \cdot \dots \cdot (R-r+1) \cdot B(B-1) \cdot \dots \cdot (B-b+1)}{N(N-1) \cdot \dots \cdot (N-n+1)}.$$

- **Probabilità di estrarre una qualsiasi sequenza ordinata** di $n \leq N$ palline colorate, di cui $r \leq R$ sono rosse e le rimanenti $b \leq B$ sono blu:

$$\frac{\binom{R}{r} \binom{B}{b}}{\binom{N}{n}}.$$

