Probabilità e Processi Stocastici (455AA)

Lezione 3

Dario Trevisan

30/09/2024

Section 1

Indipendenza probabilistica

Regole principali

- $0 \leq P(A|I) \leq 1$

$$\frac{1}{P(A)} \in [1, +\infty]$$

$$\text{odds}(A|I) = \frac{P(A|I)}{1-P(A|I)} = \frac{P(A|I)}{P(A^c|I)}$$

$$\text{odds}(A|I) = 1:5 = \frac{1}{5}$$

$$\text{odds}(B|I) = \frac{1}{1} = 1:1$$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

- Negazione $P(\text{ non } A) = 1 - P(A)$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)
- Formula di **Bayes**

$$P(B|A_i)$$

$$P(A_i|B) \propto P(A_i)\underbrace{L(A_i; B)}_{=}$$

$$\text{odds}(A|B) = \frac{P(A|B)}{P(A^c|B)} = \frac{P(A)\,L(A;B)}{P(A^c)\,L(A^c;B)} = \text{odds}(A)\,\frac{L(A;B)}{L(A^c;B)}$$

Regole principali

- $0 \leq P(A|I) \leq 1$
- **Somma** $P(A \text{ oppure } B) = P(A) + P(B)$ purché A, B *incompatibili*,

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

- Negazione $P(\text{ non } A) = 1 - P(A)$
- **Prodotto** $P(A \text{ e } B) = P(A)P(B|A)$.
- Sistema di alternative $(A_i)_{i=1}^n$ se *esattamente una* delle A_i si realizza (ma non si sa quale)
- Formula di **Bayes**

$$P(A_i|B) \propto P(A_i)L(A_i; B)$$

- **Verosimiglianza** $L(A; B) = P(B|A)$

Definizione di indipendenza

- Ricordiamo la regola del prodotto $P(A \text{ e } B) = P(A)P(B|\cancel{A})$.

In tante situazioni non sappiamo come "usare"

A per calcolare $P(B|\cancel{A}) = P(B)$

Definizione di indipendenza probabilistica

- Ricordiamo la regola del prodotto $P(A \text{ e } B) = P(A)P(B|A)$.
- Due affermazioni A, B si dicono **indipendenti** (condizionatamente all'informazione nota I) se

$$P(A|B, I) = P(A|I) \quad \text{oppure} \quad P(B|A, I) = P(B|I),$$

oppure ancora

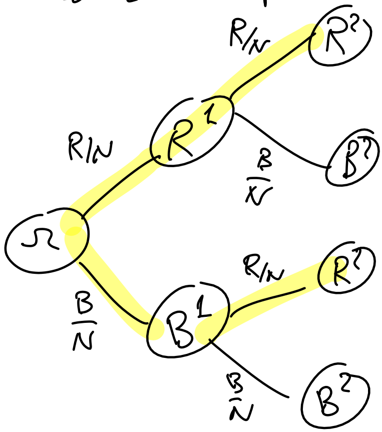
$$P(A, B|I) = P(A|I)P(B|I).$$

· sono equivalenti purché $P(B|I) > 0$ e $P(A|I) > 0$

es $\dfrac{P(A \text{ e } B|I)}{P(A|I)} = \dfrac{P(\cancel{A}|I)\,P(B|I)}{P(\cancel{A}|I)} = P(B|A, I)$

Estrazioni con rimpiazzo: seconda estrazione

Urna con N palline R rosse B blu



gli eventi R^1, R^2 sono indipendenti

$$P(R^2 | R^1) = \frac{R}{N}$$

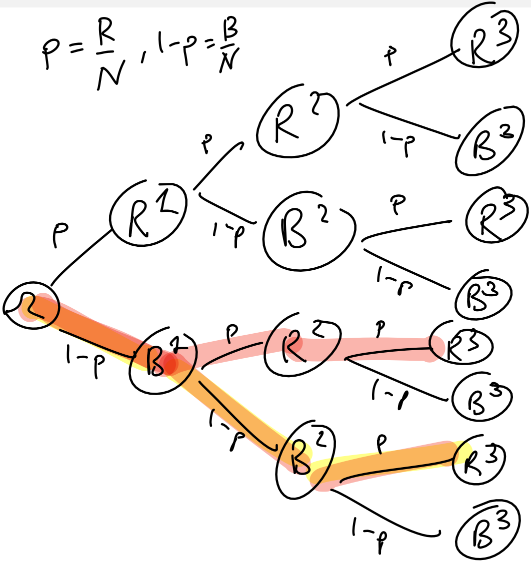
$$P(R^2 | \Omega) = \frac{R}{N} \cdot \frac{R}{N} + \frac{B}{N} \cdot \frac{R}{N} = \frac{R}{N}$$

Esercizio. Nel caso senza rimpiazzo R_1, R_2 NON sono indipendenti

- B^1, R^2 indip.
- B^1, B^2 indip. ...

Estrazioni con rimpiazzo: terza estrazione

$p = \frac{R}{N}, \quad 1-p = \frac{B}{N}$



$B^1 \cap R^3 = A$

$B^2 = B$

$P(A \text{ e } B|\Omega) =$

$= P(B^1, B^2, R^3|\Omega) = (1-p)^2 p$

$P(B^2|\Omega) = p(1-p) + (1-p)^2 = 1-p$

$P(B^1 \text{ e } R^3|\Omega) = (1-p)p^2 +$

$+ (1-p)^2 p =$

$= p(1-p)$

Sistemi di alternative indipendenti

Definizione:

Dati $k \geq 2$ sistemi di alternative $\mathcal{S}_1$, $\mathcal{S}_2$, ... $\mathcal{S}_k$, essi si dicono indipendenti tra loro (rispetto all'informazione I) se

$$P(A^1, A^2, \dots, A^k | I) = \prod_{i=1}^{n} P(A^i | I)$$

per ogni scelta di $A^1 \in \mathcal{S}_1$, $A^2 \in \mathcal{S}_2$, ... $A^k \in \mathcal{S}_k$.

Esercizio Verificare che $S_1 = \{R^1, B^1\}$, $S_2 = \{R^2, B^2\}$ ... $S_k = \{R^k, B^k\}$ (con Rimpiazzo) sono indipendenti.

Densità binomiale

$N, R, B \quad p = \frac{R}{N} \quad 1-p = \frac{B}{N}$

$$P(R^1, B^2, R^3, B^4, B^5 \mid \Omega) = p^2 (1-p)^3 = \frac{R \cdot R}{N \cdot N} \cdot \frac{B \cdot B \cdot B}{N \cdot N \cdot N}$$

P(estrarre con ripiazzo una specifica sequenza lunga n con r rosse)

$$= p^r (1-p)^{n-r}.$$

$$\overbrace{P(\text{estrarre con ripiazzo una qualsiasi sequenza lunga } n \text{ con } r \text{ rosse})}^{A_r}$$

$$= \binom{n}{r} p^r (1-p)^{n-r}.$$

$(A_r)_{r=0}^n$ sistema di alternative $\Longrightarrow 1 = \sum_{r=0}^n P(A_r) = \sum_{r=0}^n \binom{n}{r} p^r (1-p)^{n-r}$

$$\boxed{(x+y)^n = \sum \binom{n}{r} x^r y^{n-r}}$$

$$P(A_r) = \binom{n}{r} p^r (1-p)^{n-r} \qquad r \in \{0, 1, 2, \ldots, n\}$$

è detta densità binomiale di parametri n, p

$$P(\text{estrarre } r \text{ rosse in } n \text{ estraz.} \mid p) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$\parallel$$

$$P(\text{estrarre } n-r \text{ blu in } n \text{ estrazioni} \mid p)$$

$$\parallel$$

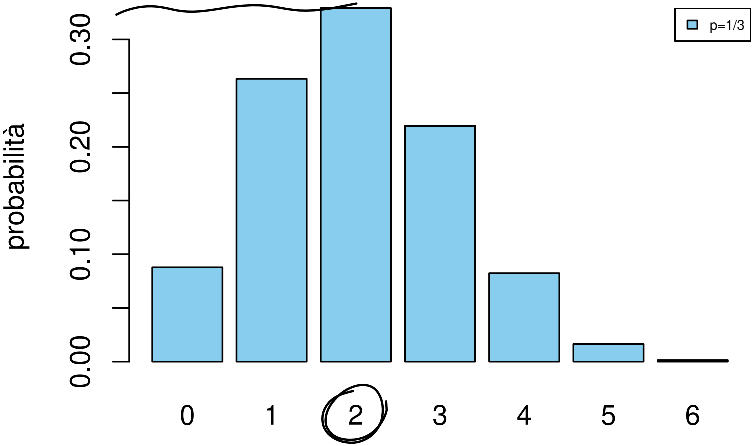
$$P(n-r \text{ "successi" in } n \text{ estrazioni} \mid 1-p) = \binom{n}{n-r} (1-p)^{n-r} p^r$$

$P($ estrarre tutte le n rosse $) = p^n$

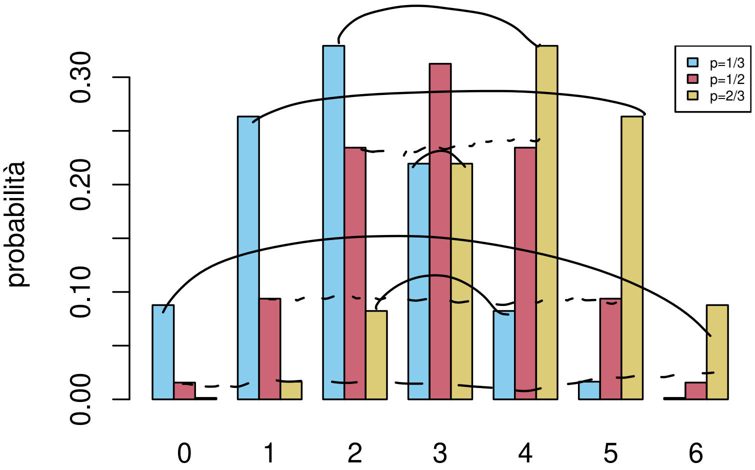
$P($ estrarre tutte le n blu $) = (1-p)^n$

Densità binomiale $p = 1/3$, $n = 6$

$p \cdot n = \frac{1}{3} \cdot 6 = 2$



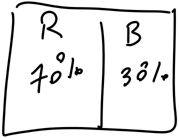
Densità binomiali a confronto



Section 2

Problemi

Problema 1



$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k(k-1)\cdot\ldots\cdot 1}$$

Supponiamo che il 70% degli studenti di una scuola abbia superato un esame di matematica. Se selezioniamo casualmente 8 studenti di questa scuola, qual è la probabilità che esattamente 5 di essi abbiano superato l'esame?

Senza rimpiazzo $\quad P(\text{" "}) = \dfrac{\binom{R}{c}\binom{B}{b}}{\binom{N}{n}}$ $\qquad N = 1000$, $R \approx 70\% \cdot N$

Con rimpiazzo $\longrightarrow$ $P(\text{"esattamente 5 su 8 superano"}) =$

$$\binom{8}{5} = \frac{8\cdot 7\cdot 6\cdot}{\cdot 3\cdot 2\cdot 1} \qquad = \binom{8}{5}(70\%)^5(30\%)^3 = \text{dbinom}(\ldots)$$

Problema 2

Si sa che una malattia rara colpisce circa l'1 per mille della popolazione di una regione. Si prendono a caso rispettivamente un campione di

a. 10 persone,

Calcolare le probabilità nei tre casi che nessuno del campione sia affetto dalla malattia.

Problema 2

Si sa che una malattia rara colpisce circa l'1 per mille della popolazione di una regione. Si prendono a caso rispettivamente un campione di

a. 10 persone,
b. 100 persone,

Calcolare le probabilità nei tre casi che nessuno del campione sia affetto dalla malattia.

Problema 2

Si sa che una malattia rara colpisce circa l'1 per mille della popolazione di una regione. Si prendono a caso rispettivamente un campione di

a. 10 persone,
b. 100 persone,
c. 1000 persone.

Calcolare le probabilità nei tre casi che nessuno del campione sia affetto dalla malattia.

Problema 3

Si sa che una malattia rara colpisce una piccola frazione della popolazione di una regione. Si suppongono inizialmente tre possibili scenari, ciascuno con probabilità uniforme: essa colpisce una persona ogni

a. 10

Problema 3

Si sa che una malattia rara colpisce una piccola frazione della popolazione di una regione. Si suppongono inizialmente tre possibili scenari, ciascuno con probabilità uniforme: essa colpisce una persona ogni

a. 10

b. 100

Problema 3

Si sa che una malattia rara colpisce una piccola frazione della popolazione di una regione. Si suppongono inizialmente tre possibili scenari, ciascuno con probabilità uniforme: essa colpisce una persona ogni

a. 10
b. 100
c. 1000.

Problema 3

Si sa che una malattia rara colpisce una piccola frazione della popolazione di una regione. Si suppongono inizialmente tre possibili scenari, ciascuno con probabilità uniforme: essa colpisce una persona ogni

a. 10
b. 100
c. 1000.

1. Si considera un campione di 70 persone e si osserva che nessuno è affetto dalla malattia. Dire quale delle tre ipotesi è più probabile.

Problema 3

Si sa che una malattia rara colpisce una piccola frazione della popolazione di una regione. Si suppongono inizialmente tre possibili scenari, ciascuno con probabilità uniforme: essa colpisce una persona ogni

a. 10
b. 100
c. 1000.

1. Si considera un campione di 70 persone e si osserva che nessuno è affetto dalla malattia. Dire quale delle tre ipotesi è più probabile.
2. Si completa successivamente la ricerca analizzando altre 30 persone e si osserva che una persona è affatta dalla malattia. Dire quale delle tre ipotesi è più probabile.

Problema 4

Un'urna contenente palline di due colori: rosse e blu. L'urna è stata preparata da un amico in uno dei due modi seguenti:

a. L'urna contiene esclusivamente 5 palline rosse.

con probabilità 10% per la prima opzione, 90% per la seconda. Cominci ad estrarre casualmente dall'urna una pallina alla volta (con rimpiazzo) osservandone il colore.

Problema 4

Un'urna contenente palline di due colori: rosse e blu. L'urna è stata preparata da un amico in uno dei due modi seguenti:

a. L'urna contiene esclusivamente 5 palline rosse.

b. L'urna contiene 5 palline rosse e 5 palline blu.

con probabilità 10% per la prima opzione, 90% per la seconda. Cominci ad estrarre casualmente dall'urna una pallina alla volta (con rimpiazzo) osservandone il colore.

Problema 4

Un'urna contenente palline di due colori: rosse e blu. L'urna è stata preparata da un amico in uno dei due modi seguenti:

a. L'urna contiene esclusivamente 5 palline rosse.

b. L'urna contiene 5 palline rosse e 5 palline blu.

con probabilità 10% per la prima opzione, 90% per la seconda. Cominci ad estrarre casualmente dall'urna una pallina alla volta (con rimpiazzo) osservandone il colore.

1. Dire se le alternative relative alle varie estrazioni sono tra loro indipendenti.

Problema 4

Un'urna contenente palline di due colori: rosse e blu. L'urna è stata preparata da un amico in uno dei due modi seguenti:

a. L'urna contiene esclusivamente 5 palline rosse.
b. L'urna contiene 5 palline rosse e 5 palline blu.

con probabilità 10% per la prima opzione, 90% per la seconda. Cominci ad estrarre casualmente dall'urna una pallina alla volta (con rimpiazzo) osservandone il colore.

1. Dire se le alternative relative alle varie estrazioni sono tra loro indipendenti.
2. Supponendo di avere estratto $k \geq 1$ palline e di avere osservato solo palline rosse, dire qual è il più probabile contenuto dell'urna (come funzione di k)

Problema 4

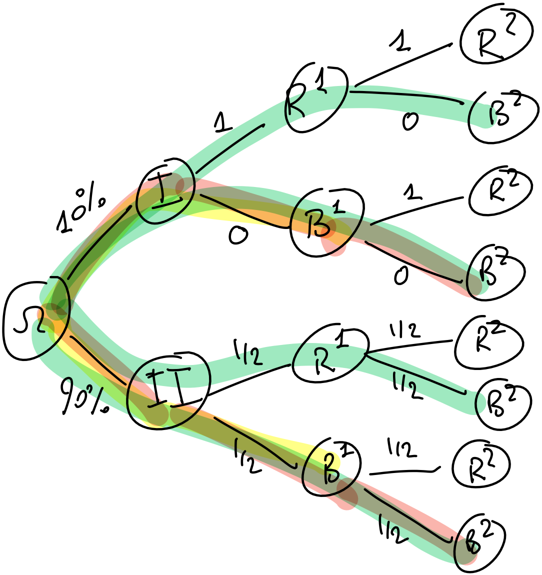
Un'urna contenente palline di due colori: rosse e blu. L'urna è stata preparata da un amico in uno dei due modi seguenti:

a. L'urna contiene esclusivamente 5 palline rosse.

b. L'urna contiene 5 palline rosse e 5 palline blu.

con probabilità 10% per la prima opzione, 90% per la seconda. Cominci ad estrarre casualmente dall'urna una pallina alla volta (con rimpiazzo) osservandone il colore.

1. Dire se le alternative relative alle varie estrazioni sono tra loro indipendenti.
2. Supponendo di avere estratto $k \geq 1$ palline e di avere osservato solo palline rosse, dire qual è il più probabile contenuto dell'urna (come funzione di k)
3. Come cambiano le risposte se le estrazioni sono senza rimpiazzo?



$>$

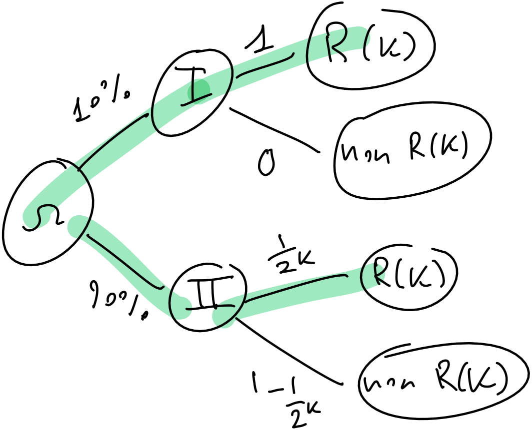
$$P(B^2|B^1) \overset{?}{=} P(B^2|\Omega)$$

$\shortparallel$

$$\frac{P(B^2 e B^1|\Omega)}{P(B^1|\Omega)} = \frac{90\% \cdot \frac{1}{4}}{90\% \cdot \frac{1}{2}} = \frac{1}{2}$$

$$P(B^2|\Omega) = 90\% \cdot \frac{1}{2}$$

$R(k) = $ "$R_1, R_2, \dots, R_k$"



$$P(I \mid R(k)) = \frac{P(I \text{ e } R(k) \mid \Omega)}{P(R(k) \mid \Omega)} = \frac{10\%}{10\% + 90\% \cdot \frac{1}{2^k}}$$

<u>Esercizio</u>: trova k per cui $\underline{P(I \mid R(k)) > \frac{1}{2}}$

Problema 5

Un'urna da cui un amico effettua estrazioni contiene inizialmente 40 palline di cui 10 rosse e 30 blu. Un amico ci informa che ha estratto la sequenza, nell'ordine

rosso, rosso, blu, rosso.

Sappiamo inoltre sicuramente che l'amico ha effettuato tutte estrazioni o con rimpiazzo o senza rimpiazzo, ma non sappiamo esattamente come (supponiamo a priori probabilità uniforme).

1. Dire se è più probabile che le estrazioni siano state effettuate con rimpiazzo o senza.

Problema 5

Un'urna da cui un amico effettua estrazioni contiene inizialmente 40 palline di cui 10 rosse e 30 blu. Un amico ci informa che ha estratto la sequenza, nell'ordine

rosso, rosso, blu, rosso.

Sappiamo inoltre sicuramente che l'amico ha effettuato tutte estrazioni o con rimpiazzo o senza rimpiazzo, ma non sappiamo esattamente come (supponiamo a priori probabilità uniforme).

1. Dire se è più probabile che le estrazioni siano state effettuate con rimpiazzo o senza.
2. Calcolare la probabilità che la pallina successiva estratta dall'amico sia rossa.