Probabilità e Processi Stocastici (455AA)

Lezione 3

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

29/09/2025

Cenni agli assiomi di Kolmogorov

Richiami: probabilità elementare

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

Richiami: probabilità elementare

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

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

Richiami: probabilità elementare

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

Richiami: probabilità elementare

- $0 \leq P(A|I) \leq 1$
- **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** $P(A \text{ e } B) = P(A)P(B|A)$.

Richiami: probabilità elementare

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

Richiami: probabilità elementare

- $0 \leq P(A|I) \leq 1$
- **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**
$$P(A_i|B) \propto P(A_i)L(A_i; B)$$

Richiami: probabilità elementare

- $0 \leq P(A|I) \leq 1$
- **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** $L(A; B) = P(B|A)$

Problemi dell'approccio elmentare

1. Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?

→ Principio di Laplace

Problemi dell'approccio elmentare

1. Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?
2. Come garantire la **consistenza** del calcolo, ossia che $P(A|I)$ sia ben definita?

$$\begin{tikzcd} I \arrow[d, squiggly, bend right] \arrow[d, squiggly, bend left] \\ P(A|I) \end{tikzcd}$$

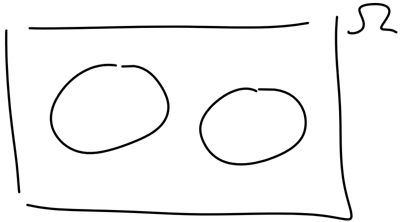
Problemi dell’approccio elmentare

1. Come attribuire le **probabilità iniziali** (quelle che abbiamo chiamato *a priori*)?
2. Come garantire la **consistenza** del calcolo, ossia che $P(A|I)$ sia ben definita?
3. Come trattare i **passaggi al limite**, in particolare, nel caso di infinite affermazioni?

Alcune risposte sono fornite dalla **descrizione assiomatica** della probabilità proposta da Kolmogorov.

L'idea: formalizzare i diagrammi di Eulero-Venn, identificando

- le affermazioni (eventi) A, I ecc. con *sottoinsiemi* di un insieme "universo" Ω (l'informazione iniziale)



Alcune risposte sono fornite dalla **descrizione assiomatica** della probabilità proposta da Kolmogorov.

L'idea: formalizzare i diagrammi di Eulero-Venn, identificando

- le affermazioni (eventi) A, I ecc. con *sottoinsiemi* di un insieme "universo" Ω (l'informazione iniziale)
- la probabilità $P(A|I)$ con una nozione astratta di *area* (relativa), definendo prima la probabilità rispetto all'informazione inizale Ω.

e poi quelle condizionate

a) insieme Ω

Si fissa un insieme "universo" Ω che rappresenta tutte le possibili situazioni (scenari) che si potrebbero presentare nel problema che si sta considerando.

Ad esempio, nel caso di un lancio di dado,

$$\Omega = \{1, 2, 3, 4, 5, 6\},$$

che corrisponde ai possibili esiti (ovviamente altre scelte sono ragionevoli).

b) Eventi

"esce pari" = $\{2, 4, 6\}$

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme $\mathcal{A}$ i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω, detto la *σ-algebra degli eventi*:

- $\Omega \in \mathcal{A}$, $\quad \Omega \rightarrow$ VERA

b) Eventi

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme $\mathcal{A}$ i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω, detto la *σ-algebra degli eventi*:

- $\Omega \in \mathcal{A}$,
- se A, $B \in \mathcal{A}$, anche A^c ("non A"), $A \cap B$ ("A e B") e $A \cup B$ ("A oppure B") sono eventi in $\mathcal{A}$.

b) Eventi

Si identificano quali affermazioni $A \subseteq \Omega$ sono d'interesse per il problema che si vuole affrontare.

Insieme $\mathcal{A}$ i cui elementi $A \in \mathcal{A}$ sono sottoinsiemi di Ω, detto la *σ-algebra degli eventi*:

- $\Omega \in \mathcal{A}$,
- se A, $B \in \mathcal{A}$, anche A^c ("non A"), $A \cap B$ ("A e B") e $A \cup B$ ("A oppure B") sono eventi in $\mathcal{A}$.
- Inoltre, per permettere di passare al limite, si richede che valga lo stesso per l'unione infinita di eventi: dati $A_n \in \mathcal{A}$, pure $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

passaggio al limite

c) funzione di probabilità (iniziale)

$P(\emptyset) = 0$

Si introduce una $P : \mathcal{A} \to [0, 1]$ tale che $P(\Omega) = 1$ e, per ogni A, $B \in \mathcal{A}$ con $A \cap B = \emptyset$ valga

$$P(A \cup B) = P(A) + P(B).$$

- $P(A)$ corrisponde alla probabilità $P(A|\Omega)$ rispetto all'informazione iniziale, di cui si richiede valga la regola della somma (per eventi incompatibili).

c) funzione di probabilità (iniziale)

Si introduce una $P : \mathcal{A} \to [0,1]$ tale che $P(\Omega) = 1$ e, per ogni A, $B \in \mathcal{A}$ con $A \cap B = \emptyset$ valga

$$P(A \cup B) = P(A) + P(B).$$

- $P(A)$ corrisponde alla probabilità $P(A|\Omega)$ rispetto all'informazione iniziale, di cui si richiede valga la regola della somma (per eventi incompatibili).
- Per passare al limite, si richiede in più che la regola della somma si estenda ad infiniti $A_n \in \mathcal{A}$ tali che $A_n \cap A_m = \emptyset$ per $n \neq m$:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \in [0,1]$$

d) Probabilità condizionata ad I

Per ogni $A, I \in \mathcal{A}$ tale che $P(I) > 0$, si pone

$$P(A|I) = \frac{P(A \cap I)}{P(I)}.$$

$(\Omega, \mathcal{A}, P)$ è uno **spazio di probabilità** secondo Kolmogorov.

2 ESTRAZIONI SENZA RIMPIAZZO

$N = 10$

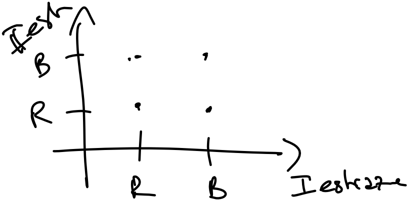
$R = 3$

$B = 7$

• $\Omega = \{(R,R), (R,B), (B,R), (B,B)\} \rightarrow$

$\{R1, B1\}$

$\{R2, B2\}$



Se $\#\Omega = d \rightarrow \#\mathcal{P}(\Omega) = 2^d$

• $\mathcal{A} = \mathcal{P}(\Omega)$ parti di Ω

$= \{\emptyset, \Omega, \{(R,R)\}, \{(R,B)\}, \{(B,R)\}, \{(B,B)\},$

$\underbrace{\{(R,R), (R,B)\}}_{\text{“}R1\text{”}}$

$P(A)$ $\quad A \subseteq \Omega$

$\overset{?}{=}$

3 rosse 7 blu

10 palline

$P(\{(R,R)\}) \overset{?}{=} \frac{3}{10} \cdot \frac{2}{9}$

$P(\{(R,B)\}) = \frac{3}{10} \cdot \frac{7}{9}$

$\vdots \qquad \vdots$

$$P(R2 \mid R1) = \frac{P(R2 \cap R1)}{P(R1)} = \frac{\frac{3}{10} \cdot \frac{2}{9}}{P((R,R)) + P((R,B))}$$

Indipendenza probabilistica

Definizione di indipendenza

- Regola del prodotto $P(A \text{ e } B) = P(A)P(B|A)$.

Definizione di indipendenza

- 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

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

oppure ancora

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

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

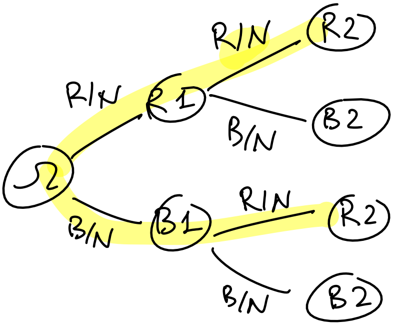
OSS La ↑ non richiede che

$P(A|I) > 0$

$P(B|I) > 0$

Estrazioni con rimpiazzo: seconda estrazione

N, R, B 2 estrazioni con rimpiazzo



$$P(R2 \mid R1) = \frac{R}{N} = P(R2)$$

$$P(R2) = \frac{R}{N}\cdot\frac{R}{N} + \frac{B}{N}\cdot\frac{R}{N} = \frac{R}{N^2}\underbrace{(R+B)}_{= N} = \frac{R}{N}$$

$$P(R_1 \mid R_2) = P(R_1) = \frac{R}{N}$$

Estrazioni con rimpiazzo: seconda estrazione

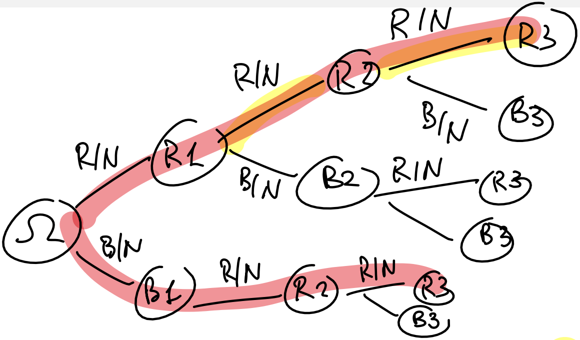
OSS $R1, B1$ NON sono indipendenti

$$P(R1 \mid B1) = 0$$

$$P(R1) = \frac{R}{N}$$

Esercizio Se A, B incompatibili $\Rightarrow$ determinare quando sono indipendenti

Estrazioni con rimpiazzo: terza estrazione



$$P(R3 \mid R1, R2) = \frac{R}{N} = P(R3 \mid R1) = P(R3 \mid R2)$$

$$= P(R3)$$

$$P(R3 \mid R2) = \frac{P(R3 \text{ e } R2)}{P(R2)} = \frac{\left(\frac{R}{N}\right)^3 + \left(\frac{B}{N}\right)\left(\frac{R}{N}\right)^2}{R/N} = \frac{R}{N}$$

Sistemi di alternative indipendenti

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

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

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

Estrazioni senza rimpiazzo   K estrazioni

$S^1 = \{R1, B1\}$ $S^2 = \{R2, B2\}$ ---- $S^K = \{RK, BK\}$

Densità binomiale

$p = \frac{R}{N}$ = probabilità di successo $\qquad \frac{B}{N} = \frac{N-R}{N} = 1 - \frac{R}{N} = 1-p$

$$\underbrace{RRR \ldots R}_{r \text{ rosse}}\ \underbrace{BB \ldots B}_{n-r \text{ blu}}$$

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

$= p^r(1-p)^{n-r}$. $\qquad \left[\frac{R(R-1)(R-2)\ldots(R-r+1)\,B(B-1)\ldots(B-b+1)}{N(N-1)\ldots(N-n+1)} \right]$

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

$r \in \{0,1,2,3,\ldots,n\}$

$$= \underbrace{\binom{n}{r}}_{\text{numero dei casi}} \underbrace{p^r(1-p)^{n-r}}_{\text{peso di ciascun caso}}.$$

OSS $\quad 1 = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} = (p + (1-p))^n = 1^n = 1$

- Estrazioni con rimpiazzo $\longrightarrow$ <u>indipendenza</u>

$$P(R100 \mid \cancel{R1, R2, R3, \ldots, R99}, \cancel{R101, R122, \ldots, R1000}) = \frac{R}{N}$$

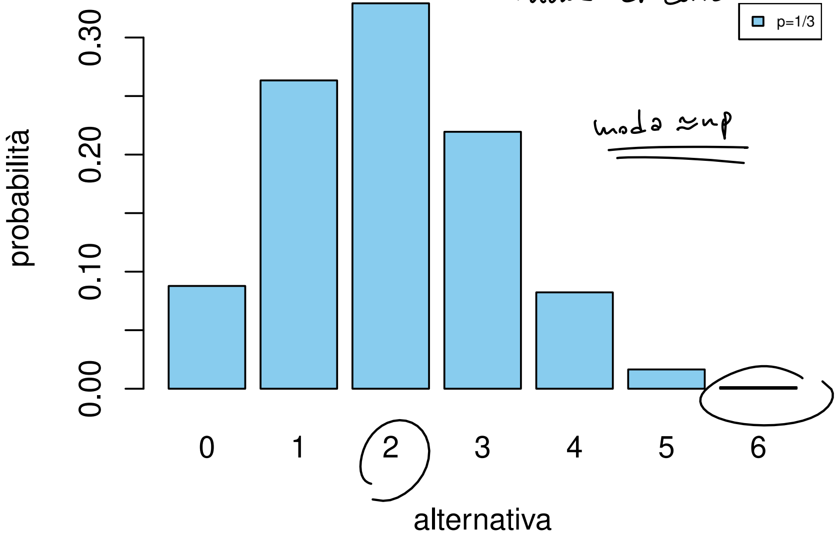
- Estrazioni senza rimpiazzo $\longrightarrow$ NON vale indipendenza

$$P(R2 \mid R1) = \frac{R-1}{N-1} \neq P(R2) = \frac{R}{N}$$

Densità binomiale $p = 1/3$, $n = 6$

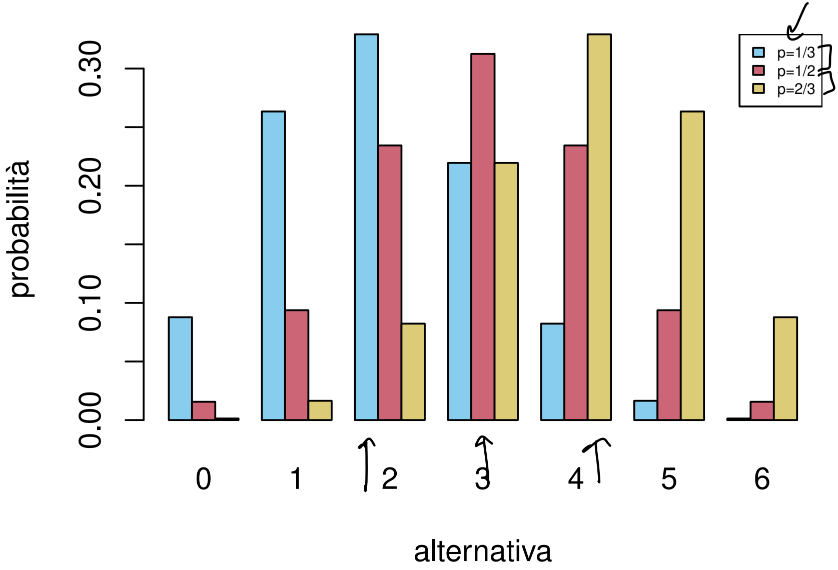
frazione di rosse

numero di estrazioni



moda $\approx np$

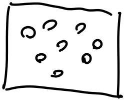
Densità binomiali a confronto $n = 6$ $p = \frac{R}{N}$



Problemi

Problema 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?



N = numero di studenti
R = studenti che hanno passato il test

$$\frac{R}{N} = 70\% = p$$

A_5 = "sequenza lunga 8 con $r=5$ palline rosse"

$$P(A_5) = \begin{cases} \text{ipergeometrica} & \text{senza rimpiazzo} \\ \text{binomiale} & \text{con rimpiazzo} \end{cases}$$

Ipergeometrica

$$P(A_r) = \frac{\binom{R}{r}\binom{B}{b}}{\binom{N}{n}} = \frac{\frac{R(R-1)(R-2)(R-3)(R-4)}{4\cdot 3\cdot 2\cdot 1}\,\frac{B(B-1)(B-2)}{3\cdot 2\cdot 1}}{\frac{N(N-1)(N-2)(N-3)(N-4)(N-5)\ldots(N-7)}{8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}}$$

$r=5$

$n=8$

Binomiale

$$P(A_r \mid \text{con rimpiazzo}) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$= \binom{8}{5}(70\%)^5(30\%)^3$$

$$= \frac{8\cdot 7\cdot 6}{3\cdot 2\cdot 1}(0{,}7)^5(0{,}3)^3 \overset{?}{\simeq} 25\%$$

$$\binom{n}{b} = \binom{n}{n-b} = \binom{n}{r}$$

Problema 3 $\quad A_{10}, A_{100}, A_{1000} \quad P(A_i) = 1/3 \quad B_{70}$

Bayes $\quad P(A_i \mid B_{70}) \propto P(A_i)\, P(B_{70} \mid A_i)$

$$P(B_{70} \mid A_{10}) = \binom{70}{0}\left(\frac{1}{10}\right)^{0}\left(\frac{9}{10}\right)^{70} = \left(\frac{9}{10}\right)^{70}$$

$$P(B_{70} \mid A_{100}) = \left(\frac{99}{100}\right)^{70} \qquad P(B_{70} \mid A_{1000}) = \left(\frac{999}{1000}\right)^{70}$$

$$P(A_{10} \mid B_{70}) = \frac{1/3 \cdot \left(\frac{9}{10}\right)^{70}}{\frac{1}{3}\left(\frac{9}{10}\right)^{70} + \frac{1}{3}\left(\frac{99}{100}\right)^{70} + \frac{1}{3}\left(\frac{999}{1000}\right)^{70}}$$

Nuova osservazione B_{30} = "una persona contatta su 30"

$$P(A_i \mid B_{30}, B_{70}) \propto \underbrace{P(A_i \mid B_{70})}_{\substack{\downarrow \\ \text{già calcolata}}} \underbrace{P(B_{30} \mid A_i, B_{70})}_{\substack{\downarrow \\ ?}}$$

$$P(B_{30} \mid A_{10}, \cancel{B_{70}}) = \binom{30}{1} \frac{1}{10} \cdot \left(\frac{9}{10}\right)^{29}$$

per indipendenza

$$P(B_{30} \mid A_{100}, \cancel{B_{70}}) = \binom{30}{1} \frac{1}{100} \cdot \left(\frac{99}{100}\right)^{29}$$

$$P(B_{30} \mid A_{1000}) = \binom{30}{1} \frac{1}{1000} \cdot \left(\frac{999}{1000}\right)^{29}$$

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.

$A_0 =$ nessun "successo"

$$P(A_0 \mid n) = \binom{n}{0} \left(\frac{1}{1000}\right)^0 \left(\frac{999}{1000}\right)^n \quad \begin{cases} n=10 \\ n=100 \\ n=1000 \end{cases}$$

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?

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.