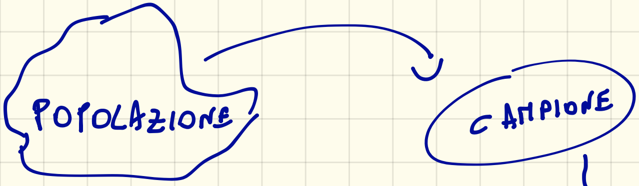
Calcolo delle probabilità

	Task Description	Due Date	Status
☐	Introduzione al calcolo delle probabilità		
☐	Gli elementi di base della teoria delle probabilità: esperimento, eventi e probabilità		
☐	Eventi elementari ed eventi composti		
☐	Due eventi particolari: l'evento certo e l'evento impossibile		
☐	Rappresentazione grafica degli eventi: i diagrammi di Eulero-Venn		
☐	Operazioni su eventi: negazione, intersezione ed unione		
☐	Relazioni tra eventi: incompatibilità, necessarietà e inclusione.		
☐	Partizione dello spazio campione		
☐	Leggi di De Morgan		
☐			

CALCOLO DELLE PROBABILITA'

Obiettivo: governare l'incertezza

- oggettiva: nel senso che il risultato del fenomeno di interesse non si è ancora verificato
- soggettiva: ho un'informazione parziale (il risultato del fenomeno si è già verificato)



effettuo le mie analisi (univariate o bivariate)

generalizzazione (dal particolare osservato al generale)

INFERENZA

errore legato al campionamento

→ valutare l'incertezza dovuta a questo passaggio

Elementi del calcolo delle probabilità

1. esperimento ⇝ fenomeno cui sono interessato il cui esito è caratterizzato da incertezza

↳ a scopi illustrativi lavoreremo inizialmente con schemi molto semplici: lancio moneta, lancio dado, estrazione di una carta, estrazione di palline da un'urna

2. evento (o esito) ⇝ uno dei possibili risultati dell'esperimento

elementare (atomo)
utilizzo la descrizione più fine possibile dell'esperimento

composto (o semplicemente evento)
utilizzo una descrizione più grossolana

3. probabilità ⟶ misura per quantificare l'incertezza

Esperimento lancio dado → { 1, 2, 3, 4, 5, 6 }

insieme eventi elementari

→ pari o dispari

eventi composti (semplicemente eventi)

{2, 4, 6} {1, 3, 5}

l'evento PARI si verifica quando come risultato dell'esperimento osservo 2 o 4 o 6

→ un evento si verifica quando si verifica uno qualunque degli eventi elementari in esso contenuto

→ uscita del 6

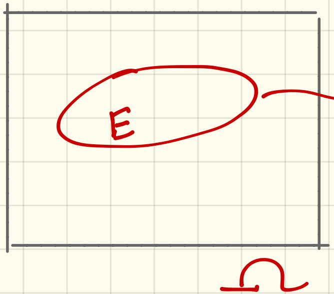
non sei { 1, 2, 3, 4, 5 }

sei { 6 }

Si indica con Ω o S lo spazio campione (insieme dei possibili esiti di un esperimento)

viene anche denominato evento certo

poiché contiene tutti i possibili risultati sono certo che si verificherà



generico evento E

a volte si fa riferimento ai diagrammi di Venn (o di Eulero - Venn)

E' possibile definire tre operazioni di base
- negazione
- intersezione
- unione

qualunque operazione definisce un evento: si parte da Ω e si finisce in Ω

operazione unaria: lavora con un solo operando

operazioni binarie: lavorano con due operandi

Due eventi particolari
- evento certo Ω
  l'evento che sicuramente si verifica
- evento impossibile $\emptyset$
  l'evento che non si può mai verificare

MEMO (ultima pagina lezione precedente)

E' possibile definire tre operazioni di base
- negazione
- intersezione
- unione

qualunque operazione definisca un evento: si parte da Ω e si finisce in Ω

operazione unaria: lavora con un solo operando

operazioni binarie: lavorano con due operandi

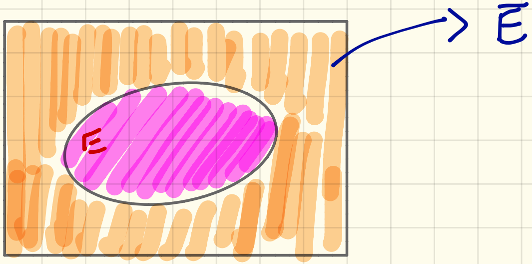
Due eventi particolari
- evento certo Ω (o S)
  l'evento che sicuramente si verifica
- evento impossibile ϕ
  l'evento che non si può mai verificare

NEGAZIONE

Dato un generico evento E, si definisce $\bar{E}$ (oppure E^c) l'evento che si verifica quando non si verifica E

$\bar{E}$ → E negato

E^c → E complementare



Due relazioni scontate:

$$\bar{\Omega} = \emptyset$$

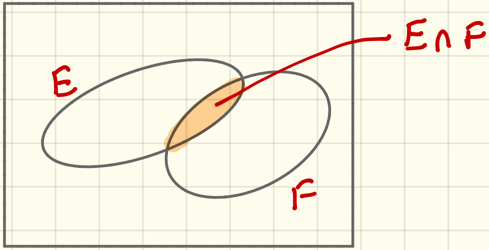
$$\bar{\emptyset} = \Omega$$

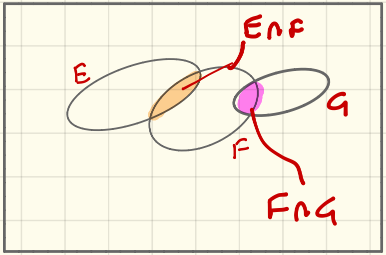
Esempio: esperimento lancio dado, $E = \{6\}$

$$\bar{E} = \{1, 2, 3, 4, 5\}$$

INTERSEZIONE

Dati due generici eventi E ed F, si definisce intersezione (o prodotto logico) $E \cap F$ (oppure $E \cdot F$) è l'evento che si verifica quando si verificano contemporaneamente sia E che F





esperimento estrazione carte da un mazzo di carte napoletane

$E = \{\text{uscita asso}\} = \{D1, S1, C1, B1\}$

$F = \{\text{uscita carta denari}\} = \{D1, D2, D3, D4, D5, D6, D7, D8, D9, D10\}$

$G = \{\text{uscita figura}\} = \{D8, D9, D10, S8, S9, S10, C8, C9, C10, B8, B9, B10\}$

$E \cap G = \emptyset$ la definizione dell'evento impossibile permette di chiudere l'operazione di intersezione su Ω

E e G sono due eventi incompatibili $\rightsquigarrow$ non possono verificarsi contemporaneamente

E ed F sono compatibili $\rightsquigarrow$ $E \cap F \neq \emptyset$

Ritornando alla negazione:

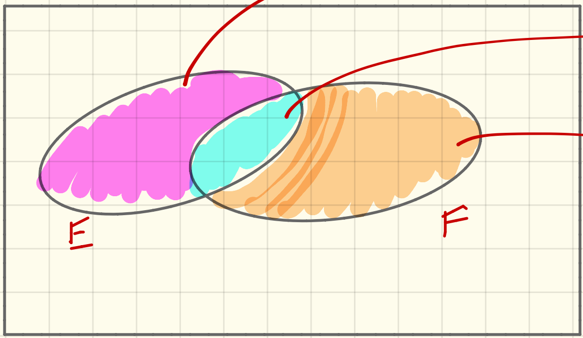
$$E \cap \bar{E} = \emptyset$$

UNIONE

Dati due generici eventi E ed F, si definisce unione $E \cup F$ (o somma logica $(E+F)$) l'evento che si verifica se:

- si verifica E ma non F
- si verifica F ma non E
- si verificano entrambi

$\left.\right\}$ ovvero se si verifica E o F



$E \cap \bar{F}$

$E \cap F$ nel caso di eventi incompatibili questa parte sarebbe assente ($\emptyset$)

$\bar{E} \cap F$

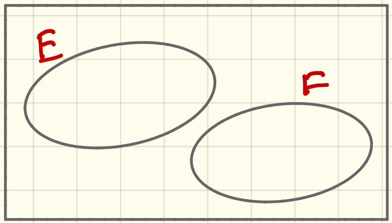
Ritornando di nuovo alla negazione:

$$E \cup \bar{E} = \Omega$$

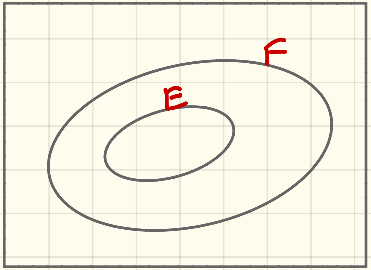
E' possibile definire, oltre alle tre operazioni, anche tre relazioni rilevanti tra eventi:

- incompatibilità $\longrightarrow$ due eventi si dicono incompatibili se non possono verificarsi <u>insieme</u>

$$E \cap F = \phi$$



- inclusione ⇝ un evento si dice incluso in F ($E \subset F$) se tutti gli eventi elementari di E sono in F
(si dice anche che F contiene E, ovvero $F \supset E$)



Estraz. carta da un mazzo di carte napoletane

$E = \{$ figure denaro $\} = \{ D8, D9, D10 \}$

$F \{$ carta denaro $\}$

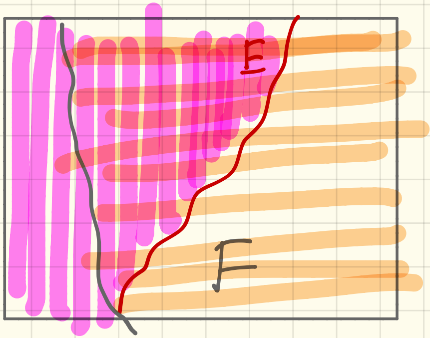
Se $E \subset F$ e contemporaneamente $F \supset E$
allora $E \equiv F$

• necessarietà ⇝ dati due eventi E ed F, questi si dicono necessari se la loro unione dà l'evento certo

$$E \cup F = \Omega$$

↙

dovrò necessariamente osservare E o F come risultato dell'esperimento



Attenzione: due eventi necessari non devono essere "per forza" incompatibili

Esperimento lancio dado

$E = \{ \text{numero} \geq 3 \} = \{3, 4, 5, 6\}$

$F = \{ \text{non sei} \} = \{1, 2, 3, 4, 5\}$

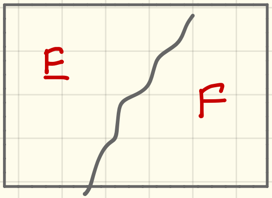
Dati n eventi $E_1, E_2, ..., E_n$ essi si dicono necessari se la loro unione dà l'evento certo:

$$E_1 \cup E_2 \cup .... \cup E_n = \bigcup_{i=1}^{n} E_i = \Omega$$

Un caso particolare si ha quando gli eventi sono insieme necessari ed incompatibili $\longrightarrow$ partizione di Ω

Due eventi E ed F definiscono una partizione di Ω se:

$$\begin{cases} E \cup F = \Omega \\ E \cap F = \emptyset \end{cases}$$



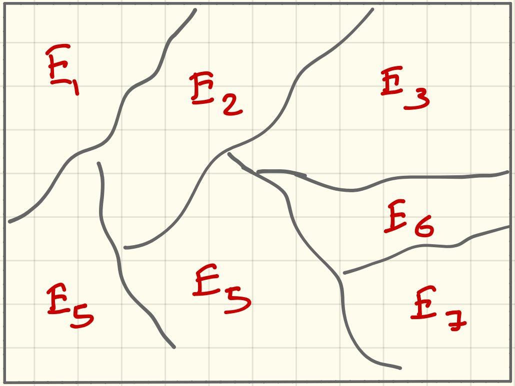
Lancio dado $\longrightarrow$ $E = \{$ numero pari $\}$

$F = \{$ numero dispari $\}$

Nel caso di n eventi: $E_1, E_2, \dots, E_n$, questi definiscono una partizione di Ω se:

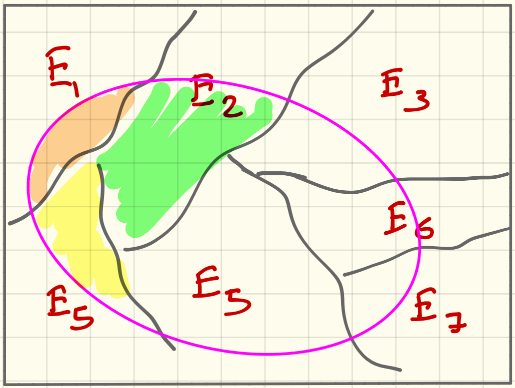
$$\bigcup_{i=1}^{n} E_i = \Omega \qquad \text{(sono necessari)}$$

$$E_i \cap E_j = \phi \ , \ \forall i \neq j \qquad \text{(ed incompatibili)}$$



NOTA: l'insieme degli eventi elementari costituisce una partizione di Ω

Può essere talvolta utile esprimere un generico evento F in termini di una partizione di Ω



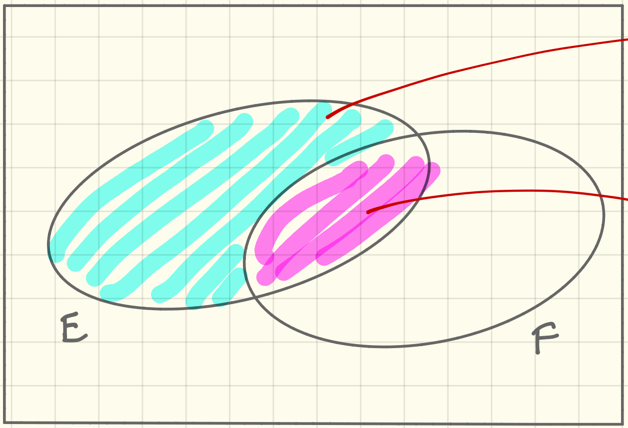
F può essere espresso come unione di n eventi incompatibili (le parti di Ω)

$$F = (F \cap E_1) \cup (F \cap E_2) \cup (F \cap E_3) \cup \dots \cup (F \cap E_n) =$$

$$= \bigcup_{i=1}^{n} (F \cap E_i)$$

PROMEMORIA: vedi teorema delle probabilità totali (prossima settimana)

Può essere talvolta utile esprimere un evento E in funzione di un altro evento F e della sua negazione:



E ma non F

$E \cap \overline{F}$

E ed F

$E \cap F$

$$E = \underbrace{(E \cap F)} \cup \underbrace{(E \cap \overline{F})}$$

Le due operazioni di unione ed intersezione possono essere collegate tra loro attraverso la negazione

Leggi De Morgan

$$\overline{E \cup F} = \overline{E} \cap \overline{F}$$

$$\overline{E \cap F} = \overline{E} \cup \overline{F}$$

esercizio: verifica grafica di queste due proprietà (da svolgere nel compito a casa)

	Task Description	Due Date	Status
☐	Definizione classica, frequentista e soggettivista di probabilità		
☐	Approccio assiomatico: i tre assiomi e i principali teoremi derivanti		
☐	Probabilità condizionata: definizione ed interpretazione		
☐	Probabilità dell'intersezione di due eventi: derivazione a partire dalla probabilità condizionata		
☐	Indipendenza stocastica: definizioni alternative per la verifica della condizione di indipendenza		
☐	Teoremi legati all'indipendenza		
☐			
☐			
☐			
☐			

La prima definizione di probabilità è quella classica:

$$P(E) = \frac{\text{n. casi favorevoli ad } E}{\text{n. casi possibili}}$$

Es. $\begin{cases} E = \{\text{uscita faccia testa moneta}\} \\ P(E) = \frac{1}{2} \end{cases}$

$E = \{\text{uscita 6 dado}\}$

$P(E) = \frac{1}{6}$

$F = \{\text{uscita numero pari sul dado}\}$

$F = \frac{3}{6} = \frac{1}{2}$

$E = \{\text{uscita carta denaro}\} \rightarrow P(E) = 10/40 = 1/4$

La definizione classica suppone che sia possibile enumerare tutti gli esiti (e che siano classificabili in favorevoli e non favorevoli)

Inoltre il problema principale è che tratta tutti gli esiti allo stesso → suppone implicitamente l'equiprobabilità di tutti gli esiti

? come gestico ad esempio il caso di un dado truccato

Definizione frequentista → usa la frequenza (del passato) per stimare la probabilità

Si ripete l'esperimento per n volte ($n \to +\infty$) e usa la frequenza relativa dell'evento per stimare la corrispondente probabilità (legame tra frequenza e probabilità)

$$P(E) = \lim_{n \to +\infty} \frac{fr(E)}{n}$$

limite cui converge la frequenza relativa

non va inteso in senso analitico

suppone che sia possibile ripetere l'esperimento nelle stesse condizioni

quando non è possibile si può fare ricorso al parere dell'esperto → definiz. soggettivista

Approccio soggettivista ⟶ parallelo probabilità — scommessa

La probab. è definita come la somma che un individuo è disposto a scommettere in un gioco equo in cui ricave 1 se l'evento si verifica e 0 se l'evento non si verifica

si basa sul meccanismo della coerenza ⟶ due soggetti diversi, a parità di informazioni devono arrivare alla stessa valutazione

il ruolo dello scommettitore può essere scambiato con quello del banco

I tre approcci sono unificati nel quadro teorico offerto dall'approccio assiomatico ⟶ a partire da 3 assiomi (non dimostrabili) è possibile costruire una teoria completa della probabilità

APPROCCIO ASSIOMATICO (Kolmogorov)

$\downarrow$

principio non dimostrabile
(ragionevole)

① Dato un generico evento E :

$P(E) \geq 0$ $\longrightarrow$ definisce un limite inferiore per la probabilità

② $P(\Omega) = 1$ $\longrightarrow$ definisce un limite superiore

i due assiomi portano a dire che $P(E) \in [0,1]$ (teorema)

③ Dati due eventi E ed F tali che $E \cap F = \emptyset$ (incompatibili):

$$P(E \cup F) = P(E) + P(F)$$

Il terzo esteso può essere esteso al caso di n eventi incompatib.

$E_1, E_2, \ldots, E_n$ tali che $E_i \cap E_j = \emptyset$
$\forall i \neq j$

l'assioma può essere scritto:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

A partire da questi tre assiomi è possibile dimostrare una serie di teoremi:

1. $P(\bar{E}) = 1 - P(E)$
2. $P(\phi) = 0$
3. $P(E) \leq 1$
4. $E \subset F \Rightarrow P(E) \leq P(F)$
5. Dati due generici eventi E ed F

$$P(E \cup F) = P(E) + P(F) - P(\underbrace{E \cap F})$$

regola generale per $P(E \cup F)$

rimane il problema del calcolo della prob. dell'intersezione (vedi prob. condizionata)

Se E ed F sono incompatibili

$E \cap F = \phi$

$\Downarrow$

$P(E \cap F) = 0$

TEOREMA 1

Dato un evento E ed il suo complemento $\bar{E}$, si ha:

$$E \cup \bar{E} = \Omega$$

Per definizione di negazione, E ed $\bar{E}$ sono incompatibili

$$E \cap \bar{E} = \emptyset$$

$$P(E \cup \bar{E}) = P(\Omega) = 1$$ dal secondo assioma

$\Downarrow$ III assioma

$$P(E \cup \bar{E}) = \underbrace{P(E) + P(\bar{E}) = 1}$$

$$\Downarrow$$

$$P(E) = 1 - P(\bar{E})$$

$$P(\bar{E}) = 1 - P(E)$$

TEOREMA 2

Se nel teorema 1 considerate al posto dell'evento E l'evento certo, avrete che $\bar{E} = \bar{\Omega} = \phi$, da cui:

$$P(\phi) = 1 - P(\Omega) = 1 - 1 = 0$$

Attenzione:

Dato un evento E, se $P(E) = 1 \not\Rightarrow E = \Omega$

evento quasi certo

Lo stesso se $P(E) = 0 \not\Rightarrow E = \phi$

evento quasi impossibile

TEOREMA 3

Se consideriamo un generico evento E, dal teorema 1 sappiamo:

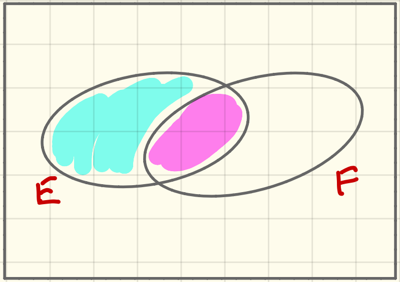
$$P(E) = 1 - \underbrace{P(\bar{E})}_{\text{dal I assioma}}$$

$$P(\bar{E}) \geq 0$$

$$P(\bar{E}) = 1 - P(E)$$

$$\boxed{\leq 1}$$

MEMO: $E = (E \cap F) \cup (E \cap \bar{F})$ un evento E può essere espresso come unione di due eventi incompatibili, costruiti considerando l'intersezione di E con F e con $\bar{F}$



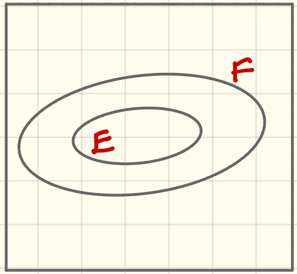
Da questo si ha:

$$P(E) = P\{\underbrace{(E \cap F) \cup (E \cap \bar{F})}_{\substack{\text{incompatibili} \\ \text{(III assioma)}}}\} = P(E \cap F) + P(E \cap \bar{F})$$

TEOREMA 4

Dati due eventi E ed F, tali che $E \subset F$, usando quanto visto nella pagina precedente:

$$P(F) = \underbrace{P(F \cap E)}_{P(E)} + P(F \cap \bar{E})$$



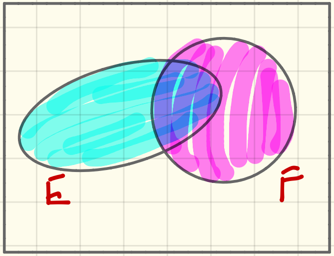
$$E \cap F = E$$

$$P(F) = P(E) + \underbrace{P(F \cap \bar{E})}_{\text{I assioma } \geq 0}$$

$$\Downarrow$$

$$P(F) \geq P(E)$$

TEOREMA 5

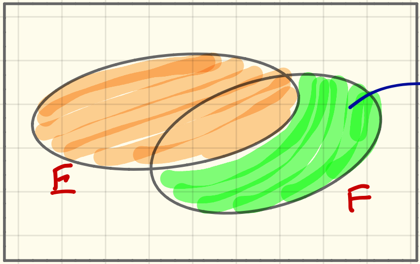


Se calcolassi $P(E \cup F)$ usando il terzo assioma

$$P(E \cup F) = P(E) + P(F)$$

considererei due volte gli eventi elementari che sono in $E \cap F$

$$P(E \cup F) = P(E) + P(F) \underbrace{- P(E \cap F)}_{\text{mi evita questo problema}}$$



F ma non E $\longrightarrow$ $\boxed{\bar{E} \cap F}$

$$E \cup F = \underbrace{E \cup (\bar{E} \cap F)}_{\text{eventi incompatibili}}$$

$$P(E \cup F) = P\{E \cup (\bar{E} \cap F)\} = P(E) + P(\bar{E} \cap F) =$$

Dalle pagine sopra sappiamo che:

$$P(F) = P(E \cap F) + P(\bar{E} \cap F)$$

$$\Downarrow$$

$$P(\bar{E} \cap F) = P(F) - P(E \cap F)$$

$$= P(E) + P(F) - P(E \cap F)$$

Rimane il modo del calcolo di $P(E \cap F)$

↓

è necessario definire il concetto di probabilità condizionata

↓

l'idea del condizionamento è di inserire un'informazione aggiuntiva circa il verificarsi di un evento

↓

si suppone che si sia verificato un evento F

↓

può essere una semplice ipotesi di lavoro o un'informazione reale circa F → diventa il nuovo Ω di riferimento

A partire da questo si vanno a considerare le probabilità degli altri eventi

La conoscenza di F richiede di aggiornare le probabilità degli altri eventi → meccanismo utilizzo per introdurre informazioni nel processo decisionale e aggiornare le valutazione numeriche

$P(E \mid F)$

E : evento condizionato

F : evento condizionante

prob. di E dato F

(condizionata dal F)

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

Consideriamo il seguente esperimento:

estrazione di 1 carta da un mazzo di carte napoletane

e i seguenti eventi:

$E = \{ \text{asso} \}$ $\qquad$ $G = \{ \text{denaro} \}$

$F = \{ \text{figura} \}$ $\qquad$ $H = \{ \text{dieci} \}$

Usando la definizione classica è facile calcolare le seguenti probabilità:

$$P(E) = \frac{4}{40} = \frac{1}{10}$$

$$P(F) = \frac{12}{40} = \frac{3}{10}$$

$$P(G) = \frac{10}{40} = \frac{1}{4}$$

$$P(H) = \frac{4}{40} = \frac{1}{10}$$

$$P(E \cap F) = P(\emptyset) = 0$$

$$P(E \cap G) = 1/40$$

$$P(E \cap H) = P(\emptyset) = 0$$

$$P(F \cap G) = 3/40$$

$$P(F \cap H) = 4/40$$

$$P(G \cap H) = 1/40$$

Immaginiamo ora che io estragga una carta e, senza mostrarvela, vi dico che si tratta di una carta di denaro ($\Omega = G$)

↳ aggiorniamo le probabilità dei vari eventi sfruttando questa informazione ⇝ prob. condizionata

$$P(E|G) = \frac{1}{10}$$

1 → 1 sola è un asso

10 → ci sono 10 carte possibili

si arriva allo stesso risultato applicando la definizione di probabilità condizionata

$$P(E|G) = \frac{P(E \cap G)}{P(G)} = \frac{1/40}{10/40} = \frac{1}{10}$$

NOTA: $P(E) = \frac{1}{10} = P(E|G)$

i due eventi E e G sono indipendenti

$$P(F|G) = \frac{3}{10}$$

3 → le 3 figure di denaro

10 → le 10 carte possibili (seme = denaro)

$$= \frac{P(F\cap G)}{P(G)} = \frac{3/40}{10/40}$$

$P(F|G) = P(F)$ → anche F e G sono indipendenti

Abbiamo visto per ora:

1. E ed F sono indipendenti
2. F e G sono indipendenti

? [cosa ci aspettiamo succeda per E e G]

$$P(E|G) = \frac{1}{10} \left[= \frac{P(E\cap G)}{P(G)} = \frac{1/40}{10/40} = \frac{1}{10} \right] = P(E)$$

anche E e G sono indipendenti

? vale una sorta di proprietà transitiva ?

$$P(H \mid G) = \frac{1}{10} \left[= \frac{P(H \cap G)}{P(G)} = \frac{1/40}{10/40} \right] = P(H)$$

anche H e G sono indipendenti

Per ora abbiamo visto che:

E ed F sono indipendenti

F e G sono indipendenti

E e G sono indipendenti

H e G sono indipendenti

~~se è vera la proprietà~~ transitiva di cui sopra, anche F e H dovrebbero essere indipendenti

verifichiamolo condizionando su F, ovvero supponiamo che la carta estratta sia una figura

$$P(H|F) = \frac{4}{12} = \frac{1}{3} \neq P(H) = \frac{1}{10}$$

i 4 dieci

le 12 carte possibili (figure)

Se $P(H|F) \neq P(H)$ H e F sono dipendenti

conoscere che si è verificato F cambia la prob. che possa verificarsi H

correlati positivamente se la prob. aumenta in seguito al condizionamento

Inoltre non vale alcuna proprietà transitiva:

se E_1 ed E_2 sono indipendenti
E_2 ed E_3 sono indipendenti

non è detto che anche E_1 e E_3 siano indipendenti

Abbiamo prima visto che E ed F sono incompatibili
$[E \cap F = \emptyset]$

non ci sono assi tra le figure

$$P(E \mid F) = \frac{0}{12} = 0 \neq P(E) = \frac{1}{10}$$

le 12 carte possibili (figure)

caso di massima dipendenza

se so che è uscita una figura sono certo che non può uscire un asso

due eventi incompatibili sono massimamente dipendenti

Una cosa analoga, anche se nel verso opposto, succede nel caso di un evento incluso in un altro

$F = \{\text{figura}\}$

$H = \{\text{dieci}\}$

$$\boxed{H \subset F}$$

Supponiamo che estraendo la carta vi dica che è uscito un dieci

$$P(F \mid H) = \frac{4}{4} = 1$$

$\parallel$

$P(\Omega)$

sono tutte figura

i 4 dieci possibili

NOTA: l'indipendenza è simmetrica

se E ed F sono indipendenti lo sono

sia se si condiziona E su F che viceversa

se E è indipendente da F, anche F è indipendente da E

provate come esercizio a verificarlo cambiando il

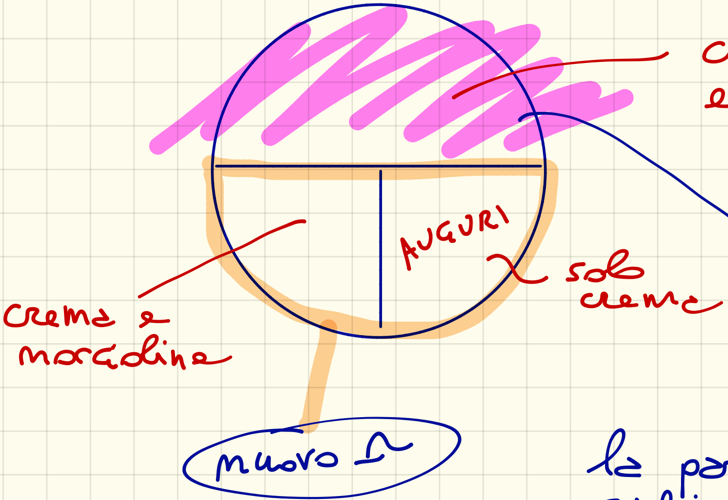
ruolo degli eventi negli esempi appena visti

ALTRO ESEMPIO → non trattato in aula

→ torta divisa in 4 parti: 2 cioccolato, 2 crema

$\frac{3}{4}$ nocciline

sul quarto alla crema non ci sono nocciline ma la scritta AUGURI



Inizialmente

$P(\text{nocciline}) = \frac{3}{4}$

la parte cioccolato è la più richiesta, per cui al nostro turno è finita

Al nostro turno:

$$P(\text{noccioline} \mid \text{metà al cioccolato è finita}) = \frac{1}{2}$$

Usando la formula:

$$P(\text{noccioline} \mid \overline{\text{cioccolato}}) = \frac{P(\text{noccioline} \cap \overline{\text{cioccolato}})}{P(\overline{\text{cioccolato}})} =$$

$$= \frac{1/4}{1/2} = \frac{1}{2}$$

$$P(\overline{\text{noccioline}} \mid \overline{\text{cioccolato}}) = \frac{1}{2}$$

$$P(\overline{\text{cioccolato}}) = 1$$

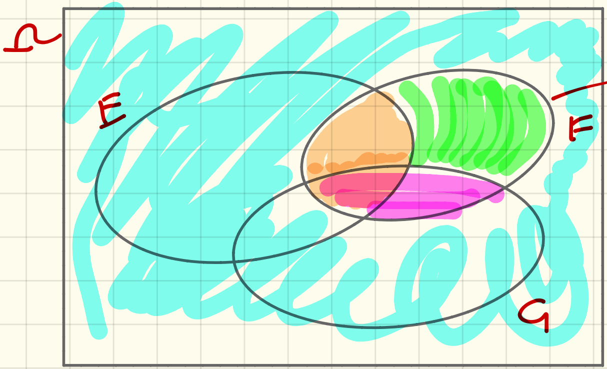
(quando so che il cioccolato è finito)

$$P(\text{nocciolina} \mid \overline{\text{cioccolato}}) + P(\overline{\text{nocciolina}} \mid \overline{\text{cioccolato}}) =$$

$$= P(\overline{\text{cioccolato}} \mid \overline{\text{cioccolato}}) = 1$$

nuovo Ω

In generale



ipotizziamo che F si sia verificato (condizionare su F

$$P(F|F) = P(\Omega) = 1$$

dividere tutto per $P(F)$, nel calcolare le probabil. condizionate,

fa sì che la somma di tutte le probabilità condizionate relative ad F sommino ad 1

Dati due eventi E ed F

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

è noto che F si è verificato oppure si tratta di un'ipotesi

$P(E|F)$

$\boxed{> P(E)}$ i due eventi sono correlati positivamente

$\boxed{< P(E)}$ i due eventi sono correlati negativamente

$\boxed{= P(E)}$ i due eventi sono indipendenti

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \qquad P(F|E) = \frac{P(E \cap F)}{P(E)}$$

regola generale per il calcolo della probab. dell'intersezione

$$P(E \cap F) = P(E|F)\,P(F)$$
$$= P(F|E)\,P(E)$$

la scelta tra le due dipende dall'informaz. disponibile

NOTA: Se E ed F sono indipendenti

$$P(E|F) = P(E) \iff P(F|E) = P(F)$$

l'indipendenza è simmetrica

Sostituendo nelle due formule sopra:

$P(E \cap F) = P(E)\,P(F)$ nel caso di eventi indipendenti

criterio alternativo per la verifica dell'indipendenza

TEOREMA INDIPENDENZA

Se E ed F sono indipendenti, allora saranno indipendenti anche:

1. E ed $\bar{F}$
2. $\bar{E}$ ed F
3. $\bar{E}$ ed $\bar{F}$

Dimostrazione 1

Se E ed F sono indipendenti $\rightsquigarrow P(E|F) = P(E)$

$$\Downarrow$$

$$1 - P(\bar{E}|F) = 1 - P(\bar{E})$$

$$\Downarrow$$

$$P(\bar{E}|F) = P(\bar{E})$$

$\bar{E}$ ed F sono indipendenti

(2) Si procede come per la prima usando $\bar{E}$ invece di $\bar{F}$

(3)

$$\underbrace{P(\bar{E}\mid F)} = \underbrace{P(\bar{E})}$$ dal teorema (2)

$$P(\bar{E}) = P(\bar{E}\cap F) + P(\bar{E}\cap\bar{F})$$

$$= P(\bar{E}\mid F)P(F) + P(\bar{E}\mid\bar{F})P(\bar{F})$$

$$\frac{P(\bar{E}\mid F)}{P(\bar{E}\mid F)} = \frac{P(\bar{E}\mid F)P(F) + P(\bar{E}\mid\bar{F})P(\bar{F})}{P(\bar{E}\mid F)}$$ dividiamo ambo i membri per $P(\bar{E}\mid F)$

$$1 = P(F) + \frac{P(\bar{E}\mid\bar{F})P(\bar{F})}{P(\bar{E}\mid F)}$$

$$\frac{\cancel{1-P(F)}}{\cancel{P(\bar{F})}} = \frac{P(\bar{E}\mid\bar{F})\cancel{P(\bar{F})}}{P(\bar{E}\mid F)}$$ dividiamo ambo i membri per $P(\bar{F})$

$$\underbrace{P(\bar{E}\mid F)}_{=P(\bar{E})} = P(\bar{E}\mid\bar{F}) \Longrightarrow P(\bar{E}\mid\bar{F}) = P(\bar{E})$$

dal teorema 2

$\Downarrow$

$\bar{E}$ e $\bar{F}$ sono indipendenti

	Task Description	Due Date	Status
	Regola di Bayes		
	Teorema delle probabilità totali		
	Teorema di Bayes: derivazione ed interpretazione; lettura del teorema sfruttando le frequenze naturali		
	Schema di estrazione da un'urna con rimessa e senza rimessa		
	Il fiammifero bruciato		
	Calcolo di probabilità su tabelle doppie. Interpretazione probabilistica dell'indice χ^2 a partire dalla condizione di indipendenza		

REGOLA DI BAYES

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

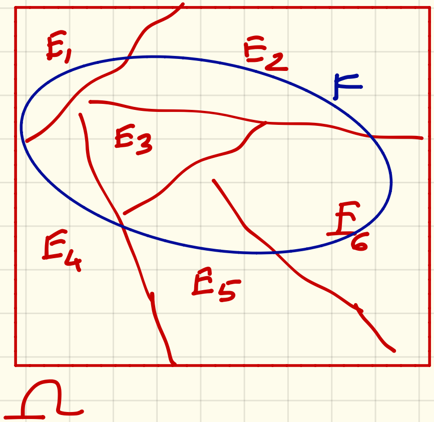
$$P(F|E) = \frac{P(E \cap F)}{P(E)}$$

$$\Downarrow$$

$$P(E \cap F) = P(F|E)\,P(E)$$

$$P(E|F) = \frac{P(F|E)\,P(E)}{P(F)}$$

per calcolare $P(E|F)$ uso le due prob. marginali e la prob. condizionata in cui i due eventi giocano un ruolo diverso



immaginiamo di avere una partizione di Ω in n eventi

- $\bigcup_{i=1}^{n} E_i = \Omega$
- $E_i \cap E_j = \phi$

ad es. le E_i sono tutte le possibili cause che possono aver generato un dato evento F

$E_i \rightsquigarrow$ malattie $\qquad F \rightsquigarrow$ sintomo

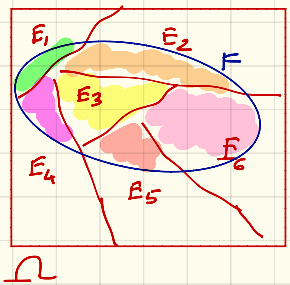
$E_i \rightsquigarrow$ guasti impianti $\qquad F \rightsquigarrow$ difetto

$P(E_i \mid F) \longrightarrow$ dato che si verificato l'evento F qual è la prob. che sia dovuto a E_i

per calcolare questo, si può procedere:

$$P(E_i \mid F) = \frac{P(E_i \cap F)}{P(F)} = \frac{P(F \mid E_i)\, P(E_i)}{P(F)}$$

$F = \bigcup_{i=1}^{m} (F \cap E_i)$ unione di m eventi incompatibili



teorema probab. totali

$$P(F) = \underbrace{P\left(\bigcup_{i=1}^{m} [F \cap E_i]\right)}_{\text{III assioma}} = \sum_{i=1}^{m} P(F \cap E_i) = \boxed{\sum_i P(F \mid E_i)\, P(E_i)}$$

Esempio: un pezzo di un assemblato è sottoposto ad un controllo per verificare se è difettoso

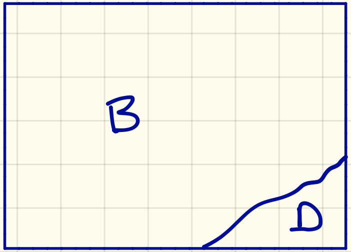
E' noto che la percentuale di pezzi difettosi è dell' 1%

E' noto inoltre che il test dà una risposta esatta nel 99% dei casi se il pezzo è difettoso e nel 98% se il pezzo è buono

Se un pezzo sottoposto al test non lo supera, qual è la probabilità che il pezzo sia effettivamente difettoso ?

$D =$ difettoso

$B =$ buono $= \bar{D}$



$P(D) = 0{,}01$

$P(B) = 0{,}99$

$P(\text{non superi il test} \mid D) = 0{,}99$

$\Downarrow$

$P(\text{superi il test} \mid D) = 0{,}01$

1

$P(\text{superi il test} \mid B) = 0{,}98$

$\Downarrow$

$P(\text{non supera il test} \mid B) = 0{,}02$

1

$$P\left(D \,\middle|\, \begin{matrix}\text{test non}\\ \text{superato}\end{matrix}\right) = \frac{P\left(D \cap \begin{matrix}\text{test non}\\ \text{superato}\end{matrix}\right)}{P\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix}\right)} =$$

$$= \frac{P\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \,\middle|\, D\right) P(D)}{P\left\{\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \cap D\right) \cup \left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \cap B\right)\right\}} =$$

$$= \frac{P\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \,\middle|\, D\right) P(D)}{P\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \,\middle|\, D\right) P(D) + P\left(\begin{matrix}\text{test non}\\ \text{superato}\end{matrix} \,\middle|\, B\right) P(B)} =$$

$$= \frac{0{,}99 \times 0{,}01}{0{,}99 \times 0{,}01 + 0{,}02 \times 0{,}99} = 0{,}333$$

ESEMPIO: interpretazione del teorema di Bayes a partire dalle frequenze naturali

- 1000 pazienti
  - 8 malati
    - 7 positive test
    - 1 test negativo
  - 992 non malattie
    - 70 positive test
    - 922 negative test

$$P(\text{malato} \mid \text{Test positivo}) = \frac{7}{7 + 70} = \frac{7}{77} = 0{,}090$$

- 1000 pazienti
  - 8 malati
    - 7 positive test
    - 1 test negativo
  - 992 non malattie
    - 70 positive test
    - 922 negative test

$P(\text{malati}) = \frac{8}{1000} = 0{,}008$

$P(\overline{\text{malati}}) = 0{,}992$

$P(\text{test posit.} \mid \text{malata}) = \frac{7}{8} = 0{,}9$

$P(\text{test posit.} \mid \overline{\text{malati}}) = 0{,}07$

$$P(\text{malato} \mid TP) = \frac{0{,}008 \times 0{,}9}{0{,}008 \times 0{,}9 + 0{,}992 \times 0{,}07} = 0{,}09$$

Gerd Gigerenzer

Quando i numeri ingannano

Imparare a vivere con l'incertezza

Raffaello Cortina Editore

NOTA: nelle tre pagine seguenti è possibile leggere il dettaglio dell'interpretazione del teorema di Bayes in termini delle frequenze naturali

LA COMPRENSIONE VISTA DALL'ESTERNO

Perché il fatto di presentare le informazioni in termini di frequenze naturali anziché di probabilità favorisce la comprensione? Le ragioni sono due: la semplicità computazionale (è la rappresentazione stessa a fare una parte del calcolo) e il primato sul piano dell'evoluzione e dello sviluppo (la nostra mente è adattata alle frequenze naturali).

LA RAPPRESENTAZIONE FA UNA PARTE DEL CALCOLO

La figura 4.2 illustra la differenza fra frequenze naturali e probabilità. A sinistra abbiamo delle frequenze naturali disposte ad albero; l'albero rappresenta il modo in cui una persona incontrerebbe delle informazioni statistiche attraverso la propria esperienza diretta. A destra le stesse informazioni vengono fornite come probabilità (ed è così che sono presentate agli studenti nella maggior parte dei libri di testo di medicina). I numeri sono ancora quelli incontrati nel problema del cancro al seno che il dottor Standing aveva trovato così difficile; i fumetti coi "pensierini" mostrano i calcoli necessari nei due casi per rispondere alla nostra domanda.

Entrambe le equazioni sono varianti della regola di Bayes,[7] che ha preso il nome dal pastore protestante (dissenziente) britannico al quale è attribuita, il reverendo Thomas Bayes (1702 (?)-1761).[8] Si vede subito che calcolare la probabilità di una malattia dato un esame positivo è più facile quando l'informazione è fornita in termini di frequenze naturali:

$$p\,(\text{malattia}|\text{positivo}) = \frac{a}{a+b}$$

La regola di Bayes con le frequenze naturali

Nella figura 4.2 *a* è il numero dei soggetti con un esame positivo *e* la malattia (7) e *b* quello dei soggetti con un esame positivo ma *senza* la malattia (70).[9] Con le probabilità, invece, il calcolo è più impegnativo:

$$p\,(\text{malattia}|\text{positivo}) = \frac{p(\text{malattia})p(\text{positivo}|\text{malattia})}{p(\text{malattia})p(\text{positivo}|\text{malattia}) + p(\text{non malattia})p(\text{positivo}|\text{non malattia})}$$

La regola di Bayes con le probabilità condizionali

Questa regola equivale a quella, più semplice, riportata sopra: entrambe ci dicono qual è la proporzione degli esiti positivi corretti (al numeratore) sulla totalità degli esiti positivi (al denominatore). La sola differenza è che nella seconda versione ogni frequenza naturale è stata sostituita dal prodotto di due probabilità. Quali? Ce lo spiega la tabella 4.1.

Un esame ha in generale quattro esiti possibili: quando una persona ha una malattia, il test può essere positivo (*positività*

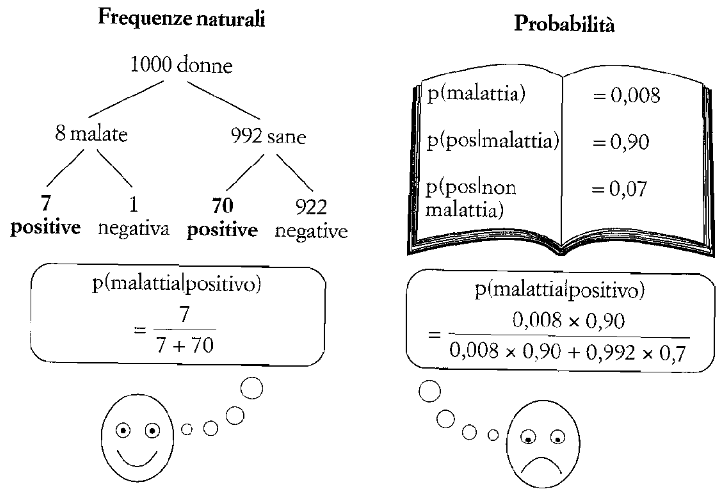


Figura 4.2 *Le frequenze naturali facilitano i computi bayesiani.* La persona dal viso felice ha avuto le informazioni pertinenti in frequenze naturali e può stimare la probabilità della malattia dato un test positivo (o un sintomo) con facilità; infatti, deve considerare due soli numeri, quello delle pazienti con un test positivo e la malattia ($a = 7$) e quello delle pazienti con un test positivo, ma senza la malattia ($b = 70$). La persona dal viso infelice ha avuto le stesse informazioni in linguaggio probabilistico, e per lei fare questa stima è difficile. La struttura della sua equazione è identica a quella usata dalla persona felice, $a/(a + b)$, ma le frequenze naturali a e b sono state trasformate in probabilità condizionali rendendo la formula molto più complessa.

Tabella 4.1 *Risultati degli esami.* Un esame può dare quattro esiti possibili: (*a*) positivo, essendoci la malattia (o qualche altra condizione ignota), (*b*) positivo senza malattia, (*c*) negativo essendoci la malattia, (*d*) negativo senza malattia. Le percentuali dei casi in cui si verificano questi quattro esiti sono dette, rispettivamente, (*a*) sensibilità (o proporzione delle positività vere), (*b*) proporzione delle positività false, (*c*) proporzione delle negatività false e (*d*) specificità (proporzione delle negatività vere). Le due aree ombreggiate indicano i due tipi di errore possibili. La frequenza delle positività vere e di quelle false sono, rispettivamente, la *a* e la *b* della regola di Bayes.

	Malattia	
Risultato dell'esame	**Sì**	**No**
Positivo	(*a*) sensibilità	(*b*) tasso di false positività
Negativo	(*c*) tasso di false negatività	(*d*) specificità

vera) o negativo (*negatività falsa*). La probabilità *p*(positivo|malattia) è la *sensibilità* dell'esame; la sensibilità della mammografia è la proporzione delle donne con un risultato positivo fra quelle col cancro al seno (di solito varia fra l'80% e il 95%, con valori più bassi fra quelle più giovani). La somma delle proporzioni di questi due esiti (sensibilità e tasso di false negatività) è uguale a 1.

Quando una persona non ha la malattia, l'esito può essere positivo (*falsa positività*) o negativo (*negatività vera*). Anche qui la somma delle rispettive proporzioni – tasso di false positività e specificità – è uguale a 1. La probabilità *p*(positivo|non malattia) è il tasso di false positività di un esame; il tasso di false positività della mammografia è la percentuale delle donne con un esito positivo fra quelle che non hanno il cancro al seno; varia fra il 5 e il 10% e ha valori più elevati fra le più giovani.

Fra i quattro esiti possibili, due (quelli ombreggiati nella tabella 4.1) sono errati. Le proporzioni di questi due errori sono interdipendenti: se il tasso di false positività di un esame diminuisce, quello di false negatività aumenta, e viceversa. Le quattro probabilità della tabella 4.1 sono dette *condizionali*

perché esprimono la probabilità di un certo evento (per esempio, un esito positivo) nell'ipotesi che se ne verifichi un certo altro (per esempio, la malattia) – *dato* cioè questo altro evento. La probabilità non condizionale *p*(malattia) è il *tasso di base* della malattia stessa; in altre parole, il tasso di base è la proporzione delle persone con una determinata malattia su una certa popolazione e in un momento specifico – ed è ben noto che le probabilità condizionali ci fanno confondere, il tasso di base no.

Ma adesso capiamo finalmente la ragione esatta per cui le cose stanno così. Quando trasformiamo le frequenze naturali in probabilità condizionali, togliamo di mezzo il tasso di base (è la cosiddetta *normalizzazione*), cosa che ha il vantaggio di far cadere tutti i valori risultanti in uno stesso intervallo, quello fra 0 e 1, ma anche lo svantaggio che quando facciamo inferenze a partire dalle probabilità anziché dalle frequenze naturali dobbiamo ri-moltiplicare queste probabilità condizionali per i rispettivi tassi di base.[10]

Ricapitolando, le frequenze naturali facilitano le inferenze effettuate sulla base di informazioni numeriche. È la rappresentazione a eseguire una parte del ragionamento fornendoci direttamente il risultato di certe moltiplicazioni che dovremmo fare noi, se ci venissero fornite delle probabilità. È in questo senso che la *comprensione* può venire alla mente dall'*esterno*.*

LA MENTE È ADATTATA ALLE FREQUENZE NATURALI

Le frequenze naturali derivano dal processo che ha messo gli umani e gli animali di fronte all'informazione sul rischio per la maggior parte della loro evoluzione, mentre le probabilità, le percentuali e altri modi formalizzati di rappresentarsi il rischio sono relativamente recenti. Gli animali possono regi-

* Gioco di parole intraducibile. *Insight* significa "comprensione" ma anche, letteralmente, "visione interna" (*in* = "dentro", *sight* = "visione"). [NdT]

Schema estrazione palline da un'urna

↓

molti esperimenti possono essere schematizzati usando il parallelo di estrazioni successive da un'urna

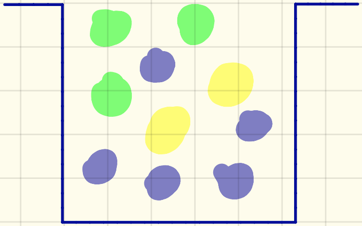
① estrazione con rimessa (con ripetizione)

l'unità estratta viene rimessa nell'urna prima di estrarre le palline successive → Ω rimane invariato sulle varie estrazione → singole estrazioni (sottoprove) sono indipendenti

② estrazione in blocco (senza rimessa o senza ripetizione)

↓

Ω cambia di estrazione in estrazione → sottoprove dipendenti



$$N = 10 \text{ palline} \begin{cases} 3 \; V \\ 2 \; G \\ 5 \; B \end{cases}$$

$$P(V) = \frac{3}{10} \qquad P(G) = \frac{2}{10} \qquad P(B) = \frac{5}{10}$$

Due estrazioni successive $\longrightarrow$ P(estrarre le due gialle)

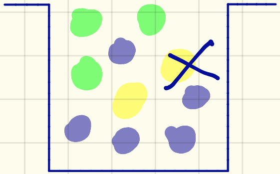
$$P(G_1 \cap G_2) = \underbrace{P(G_1)\, P(G_2 \mid G_1)}$$

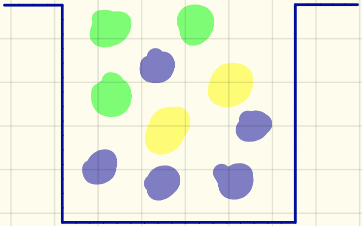
① con rimessa

$$\frac{2}{10} \times \overbrace{\frac{2}{10}}^{P(G_2 \mid G_1) = P(G)}$$

② in blocco

$$\frac{2}{10} \times \frac{1}{9}$$





$$P(\underbrace{\text{estrarre 1 sola pallina gialla}}) =$$

gialla alla prima

$$\begin{bmatrix} G_1 \cap V_2 \\ G_1 \cap B_2 \end{bmatrix} \cup \begin{bmatrix} V_1 \cap G_2 \\ B_1 \cap G_2 \end{bmatrix}$$

gialla alla seconda

$$= P\{(G_1 \cap V_2) \cup (G_1 \cap B_2) \cup (V_1 \cap G_2) \cup (B_1 \cap G_2)\} =$$

eventi incompatibili (III assioma

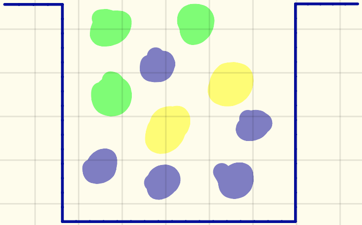
$$= \underbrace{P(G_1 \cap V_2) + P(G_1 \cap B_2) + P(V_1 \cap G_2) + P(B_1 \cap G_2)}$$

1

$$P(G_1)P(V_2) + P(G_1)P(B_2) + \dots$$

2

$$P(G_1)P(V_2 | G_1) + P(G_1)P(B_2 | G_1) + \dots$$



in molti casi conviene dividere gli eventi in eventi di interesse (successi) ed altro (insuccessi)

$$P(\text{estrarre 1 pallina } G) = P\{(G_1 \cap \bar{G}_2) \cup (\bar{G}_1 \cap G_2)\} =$$

$$= P(G_1 \cap \bar{G}_2) + P(\bar{G}_1 \cap G_2)$$

(1)

$$P(G_1)P(\bar{G}_2) + P(\bar{G}_1)P(G_2)$$

$$\frac{2}{10}\,\frac{8}{10} + \frac{8}{10}\,\frac{2}{10}$$

(2)

$$P(G_1)P(\bar{G}_2|G_1) + P(\bar{G}_2)P(G_2|\bar{G}_1)$$

$$\frac{2}{10}\,\frac{8}{9} + \frac{8}{10}\,\frac{2}{9}$$

Utilizzando questo approccio i calcoli diventano più semplici (in particolare all'aumentare del n. di colori differenti e del n. di estrazioni)

3 estrazioni

$P(3 \text{ palline blu}) = P(B_1 \cap B_2 \cap B_3) =$

1

$P(B_1)P(B_2)P(B_3)$

2

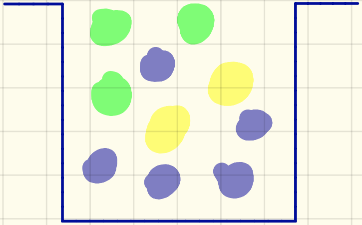
$P(B_1)P(B_2|B_1)P(B_3|B_1 \cap B_2)$

$\frac{5}{10} \quad \frac{4}{9} \quad \frac{3}{8}$

NOTA

Se il n. di palline nell'urna è molto grande $(\to \infty)$ i due schemi sono interscambiabili

conviene in questo usare il più semplice



NOTA: la dipendenza tra sottoprove nello schema di campionamento senza rimessa "si perde" nel caso in cui non si hanno informazioni sull'esito delle prove precedenti

↓

la dipendenza è legata al fatto che lo spazio campione varia da prova a prova e che si ha conoscenza della composizione dello spazio campione in ciascuna sottoprova

↓

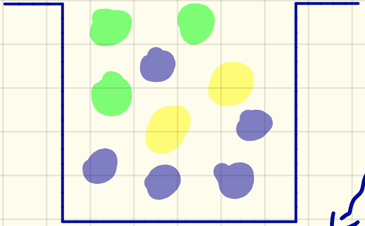
se non si hanno informazioni sugli esiti delle prove precedenti, la probabilità di osservare un evento ad ogni sottoprova è costante (ed è uguale alla probabilità di osservare l'evento alla prima prova)

↓

proviamo a calcolare ad esempio la probabilità di estrarre una pallina gialla dall'urna alla seconda estrazione senza avere informazione sull'esito della prima estrazione

↓

$$P\left(\begin{matrix}\text{gialla alla} \\ \text{II estrazione}\end{matrix}\right) = P(G_2)$$



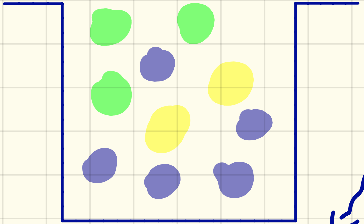
$$P(G_2) \longrightarrow$$ non avendo informazioni sull'esito della prima prova è necessario considerare le varie combinazioni possibili di uscita (alla prima prova)

$$P(G_2) = P\left\{ (G_1 \cap G_2) \cup (\bar{G}_1 \cap G_2) \right\} =$$

$$= P(G_1 \cap G_2) + P(\bar{G}_1 \cap G_2) =$$

$$= P(G_1)\,P(G_2 | G_1) + P(\bar{G}_1)\,P(G_2 | \bar{G}_1) =$$

$$= \frac{2}{10} \times \frac{2}{9} + \frac{7}{10} \times \frac{2}{9} = \frac{2}{10}\left(\frac{2}{9} + \frac{7}{9}\right) = \frac{2}{10} = P(G_1)$$



$P(G_3) \longrightarrow$ non avendo informazioni sull'esito della prima prova è necessario considerare le varie combinazioni possibili di uscita (alla prima prova)

$$P(G_3) = P\{(\bar{G}_1 \cap \bar{G}_2 \cap G_3) \cup (G_1 \cap \bar{G}_2 \cap G_3) \cup (\bar{G}_1 \cap G_2 \cap G_3)\} =$$

$$= P(\bar{G}_1 \cap \bar{G}_2 \cap G_3) + P(G_1 \cap \bar{G}_2 \cap G_3) + P(\bar{G}_1 \cap G_2 \cap G_3) =$$

$$= P(\bar{G}_1)P(\bar{G}_2|\bar{G}_1)P(G_3|\bar{G}_1 \cap \bar{G}_2) + \quad \frac{8}{10} \times \frac{7}{9} \times \frac{2}{8} +$$

$$+ P(G_1)P(\bar{G}_2|G_1)P(G_3|G_1 \cap \bar{G}_2) + \quad \frac{2}{10} \times \frac{8}{9} \times \frac{1}{8} +$$

$$+ P(\bar{G}_1)P(G_2|\bar{G}_1)P(G_3|\bar{G}_1 \cap \bar{G}_2) = \quad \frac{8}{10} \times \frac{2}{9} \times \frac{1}{8} =$$

$$= P(G_1)$$

$$= \frac{8}{10} \times \frac{2}{8}\left[\frac{7}{9} + \frac{1}{9} + \frac{1}{9}\right] =$$

$$= \frac{2}{10}$$