UNIVERSITÀ DEGLI STUDI DI TRIESTE

Dipartimento di scienze economiche, aziendali, matematiche e statistiche "Bruno de Finetti"

# Statistica

## Probabilità condizionate e teorema di Bayes

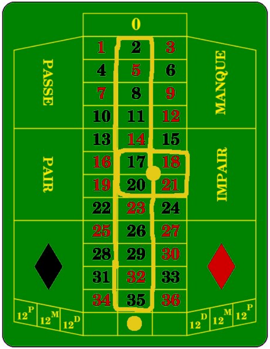
Francesco Pauli

A.A. 2017/2018

# Indice

# Eventi condizionati: esempio

Riprendiamo la roulette




Scommessa multipla:

$C$: seconda colonna

$Q$: quartina 17,18,20,21

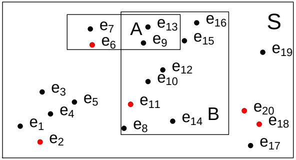
- Abbiamo effettuato due puntate:
  - sulla II colonna: $C$
  - sulla quartina $\{17, 18, 20, 21\}$: $Q$
- sappiamo che $P(C) = 12/37$;
- supponiamo di **sapere che è vincente la scommessa sulla quartina** (ma non che numero esattamente è uscito)
- Cambia la mia valutazione della probabilità che $C$ sia vincente?
  - se è vincente la quartina è uscito uno tra 17, 18, 20, 21
  - se è uscito 17 o 20 è vincente anche la colonna, se è uscito 20 o 21 no
  - Quindi sapendo che è vero $Q$, la probabilità di $C$ è $1/2$.

# Evento condizionato

Dati due eventi $A$ e $B$ si parla di **evento $A$ condizionato all'evento $B$** quando sappiamo che l'evento $B$ si verifica, l'evento condizionato si indica con

$$A|B$$

In pratica, l'evento $B$ diventa il nuovo evento certo.



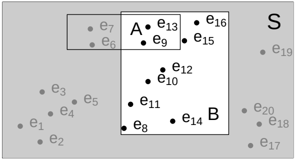
# Evento condizionato

Dati due eventi $A$ e $B$ si parla di **evento $A$ condizionato all'evento $B$** quando sappiamo che l'evento $B$ si verifica, l'evento condizionato si indica con

$$A|B$$

In pratica, l'evento $B$ diventa il nuovo evento certo.

- Ho un nuovo insieme di eventi elementari che comprende solo quelli che costituiscono $B$.



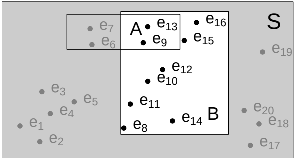
# Evento condizionato

Dati due eventi $A$ e $B$ si parla di **evento $A$ condizionato all'evento $B$** quando sappiamo che l'evento $B$ si verifica, l'evento condizionato si indica con

$$A|B$$

In pratica, l'evento $B$ diventa il nuovo evento certo.

- Ho un nuovo insieme di eventi elementari che comprende solo quelli che costituiscono $B$.




La probabilità di $A$ condizionato a $B$ è

$$P(A|B) = \frac{\sum_{A \cap B} p_i}{\sum_B p_i}$$

# Probabilità condizionate

**Definizione: probabilità condizionata**

Dati due eventi $A$ e $B$, la probabilità di $A$ sapendo che è accaduto $B$ si indica con $P(A|B)$ ed è pari a

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

In pratica, $B$ diventa l'evento certo, quindi

- $A$ può verificarsi solo se si verifica $A \cap B$,
- si divide per $P(B)$ per rinormalizzare.

# Probabilità della congiunzione

Dall'espressione per la probabilità condizionata si ottiene che

**Probabilità di $A \cap B$**

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

La probabilità che due eventi si verifichino congiuntamente è il prodotto della probabilità che se ne verifichi uno e della probabilità che si verifichi l'altro **sapendo che si è verificato il primo**.

# Indipendenza e dipendenza

Peschiamo una carta da un mazzo standard di 52

$$E_1 = \{\text{il seme è } \spadesuit\}, \quad P(E) = 13/52$$

# Indipendenza e dipendenza

Peschiamo una carta da un mazzo standard di 52

$$E_1 = \{\text{il seme è } \spadesuit\}, \quad P(E) = 13/52$$

Consideriamo due esperimenti alternativi

la carta è reinserita nel mazzo, un'altra carta è pescata

la carta NON è reinserita nel mazzo, un'altra carta è pescata

# Indipendenza e dipendenza

Peschiamo una carta da un mazzo standard di 52

$$E_1 = \{\text{il seme è } \spadesuit\}, \quad P(E) = 13/52$$

Consideriamo due esperimenti alternativi

la carta è reinserita nel mazzo, un'altra carta è pescata

la carta NON è reinserita nel mazzo, un'altra carta è pescata

Qual è la probabilità di $E_2 = \{\text{la seconda carta è } \spadesuit\}$?

# Indipendenza e dipendenza

Peschiamo una carta da un mazzo standard di 52

$$E_1 = \{\text{il seme è } \spadesuit\}, \quad P(E) = 13/52$$

Consideriamo due esperimenti alternativi

la carta è reinserita nel mazzo, un'altra carta è pescata

la carta NON è reinserita nel mazzo, un'altra carta è pescata

Qual è la probabilità di $E_2 = \{\text{la seconda carta è } \spadesuit\}$?

la probabilità è $\frac{13}{52}$ qualunque sia la prima

$$P(E_2|E_1) = P(E_2|\bar{E}_1) = \frac{13}{52}$$

la probabilità dipende dal seme della prima

$$P(E_2|E_1) = \frac{12}{51} \neq P(E_2|\bar{E}_1) = \frac{13}{51}$$

# Indipendenza e dipendenza

Peschiamo una carta da un mazzo standard di 52

$$E_1 = \{\text{il seme è } \spadesuit\}, \quad P(E) = 13/52$$

Consideriamo due esperimenti alternativi

la carta è reinserita nel mazzo, un'altra carta è pescata

la carta NON è reinserita nel mazzo, un'altra carta è pescata

Qual è la probabilità di $E_2 = \{\text{la seconda carta è } \spadesuit\}$?

la probabilità è $\frac{13}{52}$ qualunque sia la prima

$$P(E_2|E_1) = P(E_2|\bar{E}_1) = \frac{13}{52}$$

Gli eventi $E_1$ e $E_2$ sono **indipendenti**

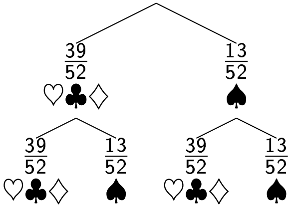
la probabilità dipende dal seme della prima

$$P(E_2|E_1) = \frac{12}{51} \neq P(E_2|\bar{E}_1) = \frac{13}{51}$$

Gli eventi $E_1$ e $E_2$ sono **dipendenti**

# Indipendenza e dipendenza

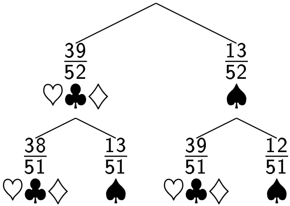
Con reinserimento




quindi

$$P(\spadesuit\spadesuit) = \frac{13}{52}\frac{13}{52} = \frac{1}{16} = 0.0625$$

Senza reinserimento




quindi

$$P(\spadesuit\spadesuit) = \frac{13}{52}\frac{12}{51} = 0.0588$$

# Definizione di indipendenza stocastica

**Definizione: indipendenza stocastica**

Dati due eventi $A$ e $B$ essi sono indipendenti se e solo se

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

L'esito dell'uno non influenza la probabilità attribuita all'esito del secondo

---

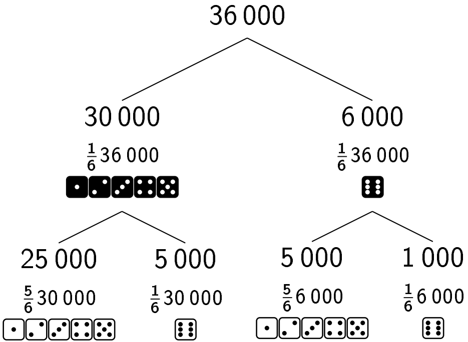
Si noti che potremmo definire l'indipendenza dicendo che

**Definizione: indipendenza stocastica /2**

Dati due eventi $A$ e $B$ essi sono indipendenti se e solo se

$$P(A \cap B) = P(A)P(B)$$

# Perché il prodotto?




| | | |
|---|---|---|
| 25 000 | 5 000 | ⚀⚁⚂⚃⚄ |
| 5 000 | 1 000 | ⚅ |
| ⚀⚁⚂⚃⚄ | ⚅ | |

Lanciamo 36 000 volte una coppia di dadi distinti ⚅⚅

in media esce ⚅ in $\frac{1}{6}$ dei lanci: 6 000,

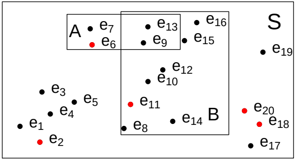
tra questi 6 000 esce ⚅ in $\frac{1}{6}$ (indipendenza): 1 000.

Esce ⚅ ∩ ⚅ in 1 000 occasioni su 36 000:

$$P(⚅ \cap ⚅) = \frac{1\,000}{36\,000} = \frac{1}{36} = P(⚅)P(⚅)$$

# Esempio

- $A = \{e_6, e_7, e_9, e_{13}\}$
- $B = \{e_8, e_9, \ldots, e_{16}\}$
- $C = \{e_2, e_6, e_{11}, e_{18}, e_{20}\}$




Assumendo
$P(e_1) = \ldots = P(e_{10}) = 0.03$
$P(e_{11}) = \ldots = P(e_{20}) = 0.07$
si ottenga

- $P(A|B) = \ldots$
- $P(A \cap C|B) = \ldots$
- $P(B|C) = \ldots$

Assumendo

$$P(e_i) = \begin{cases} 0.04 & i \text{ pari} \\ 0.06 & i \text{ dispari} \end{cases}$$

si ottenga

- $P(A|B) = \ldots$
- $P(A \cap C|B) = \ldots$
- $P(B|C) = \ldots$

# Esempio: vaiolo a Boston, 1721

- I dati (presi da OpenIntroStat) riguardano 6224 persone esposte al virus del vaiolo a Boston nel 1721.
- Di questi, 244 vennero esposti – su base volontaria – alla malattia in modo controllato dai medici (una specie di vaccinazione ante litteram).
- È noto poi chi è sopravvissuto all'epidemia e chi no.

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di sopravvivere all'epidemia?

# Esempio: vaiolo a Boston, 1721

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di sopravvivere all'epidemia?

Risp: La probabilità si ottiene come

$$P(\text{Sopr}) = \frac{\#\text{Sopr}}{\#\text{Totale}} = \frac{5374}{6224} = 0.863$$

# Esempio: vaiolo a Boston, 1721

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di essere inoculato?

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di essere inoculato?

Risp: La probabilità si ottiene come

$$P(\text{Inoc}) = \frac{\#\text{Inoc}}{\#\text{Totale}} = \frac{244}{6224} = 0.039$$

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di essere inoculato **e** di sopravvivere?

# Esempio: vaiolo a Boston, 1721

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo preso a caso, qual è la probabilità di essere inoculato **e** di sopravvivere?

Risp: La probabilità si ottiene come

$$P(\text{Inoc} \cap \text{Sopr}) = \frac{\#\text{Inoc e Sopr}}{\#\text{Totale}} = \frac{238}{6224} = 0.038$$

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo inoculato, qual è la probabilità di sopravvivere all'epidemia?

# Esempio: vaiolo a Boston, 1721

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo inoculato, qual è la probabilità di sopravvivere all'epidemia?

Risp: La probabilità si ottiene come

$$P(\text{Sopr}|\text{Inoc}) = \frac{P(\text{Inoc} \cap \text{Sopr})}{P(\text{Inoc})} = \frac{0.038}{0.039} = 0.974$$

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Dom: Per un individuo inoculato, qual è la probabilità di sopravvivere all'epidemia?

Risp: La probabilità si ottiene come

$$P(\text{Sopr}|\text{Inoc}) = \frac{P(\text{Inoc} \cap \text{Sopr})}{P(\text{Inoc})} = \frac{0.038}{0.039} = 0.974$$

Alt: Si noti che la stessa risposta si ottiene come

$$P(\text{Sopr}|\text{Inoc}) = \frac{\#\text{Inoc} \cap \text{Sopr}}{\#\text{Inoc}} = \frac{238}{244}$$

# Esempio: vaiolo a Boston, 1721

| | | inoculato | | |
|---|---|---|---|---|
| | | sì | no | Totale |
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Eserc: Si calcoli $P(\text{Sopr}|\text{non Inoc})$

# Esempio: vaiolo a Boston, 1721

| | | inoculato sì | inoculato no | Totale |
|---|---|---|---|---|
| Sopravvissuto | sì | 238 | 5136 | 5374 |
| | no | 6 | 844 | 850 |
| | Totale | 244 | 5980 | 6224 |

Eserc: Si calcoli $P(\text{Sopr}|\text{non Inoc})$

Risp: (0.859)

# Esempio: elezioni, elettori del M5S

- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo abbia votato M5S al proporzionale?

# Esempio: elezioni, elettori del M5S

- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo abbia votato M5S al proporzionale? Gli elettori sono 46 505 499, i votanti 32 755 044 (proporzionale)

| Partito | voti |
|---|---:|
| Partito Democratico | 6 134 727 |
| Liberi e Uguali | 1 109 198 |
| Lega | 5 691 921 |
| Forza Italia | 4 590 774 |
| UDC | 428 298 |
| Fratelli d'Italia | 1 426 564 |
| M5S | 10 697 994 |

- Gli elettori sono 46 505 499.
- Tra questi, hanno votato M5S in 10 697 994

# Esempio: elezioni, elettori del M5S

- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo abbia votato M5S al proporzionale? Gli elettori sono 46 505 499, i votanti 32 755 044 (proporzionale)

| Partito | voti |
|---|---:|
| Partito Democratico | 6 134 727 |
| Liberi e Uguali | 1 109 198 |
| Lega | 5 691 921 |
| Forza Italia | 4 590 774 |
| UDC | 428 298 |
| Fratelli d'Italia | 1 426 564 |
| M5S | 10 697 994 |

- Gli elettori sono 46 505 499.
- Tra questi, hanno votato M5S in 10 697 994
- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo sia andato a votare?

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo sia andato a votare?
- $P(\text{Ha votato}) = 0.704$

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- $P(\text{Ha votato}) = 0.704$
- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo sia andato a votare **e** abbia votato il M5S?

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- $P(\text{Ha votato}) = 0.704$
- Qual è la probabilità che, preso un individuo a caso tra gli aventi diritto al voto, questo sia andato a votare e abbia votato il M5S?
- $P(\text{M5S} \cap \text{Ha votato}) = P(\text{M5S})$

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- $P(\text{Ha votato}) = 0.704$
- $P(\text{M5S} \cap \text{Ha votato}) = P(\text{M5S})$ Supponiamo ora di sapere che l'individuo preso a caso è andato a votare, qual è la probabilità che abbia votato M5S?

# Esempio: elezioni, elettori del M5S

- $P(\text{M5S}) = \frac{10\,697\,994}{46\,505\,499} = 0.23.$
- $P(\text{Ha votato}) = 0.704$
- $P(\text{M5S} \cap \text{Ha votato}) = P(\text{M5S})$ Supponiamo ora di sapere che l'individuo preso a caso è andato a votare, qual è la probabilità che abbia votato M5S?
- Si ha

$$\begin{aligned} P(\text{voto M5S}|\text{Ha votato}) &= \frac{P(\text{voto M5S} \cap \text{Ha votato})}{P(\text{Ha votato})} \\ &= \frac{P(\text{voto M5S})}{P(\text{Ha votato})} = \frac{0.23}{0.704} = 0.327 \end{aligned}$$

# Compleanni

In una stanza ci sono 30 persone, qual è la probabilità che due di queste abbiano il compleanno nello stesso giorno?

## Compleanni

In una stanza ci sono 30 persone, qual è la probabilità che due di queste abbiano il compleanno nello stesso giorno?

---

Vogliamo calcolare la probabilità dell'evento

$$E = \{\text{due persone condividono il compleanno}\}$$

osserviamo che è

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

dove $\bar{E} = \{\text{tutti hanno compleanni diversi}\}$, per il quale si ha

$$P(\bar{E}) = \frac{364}{365}\frac{363}{365} \times \ldots \times \frac{365-28}{365}\frac{365-29}{365} \approx 0.32$$

e quindi $P(E) \approx 0.68$.

# Compleanni

In una stanza ci sono 30 persone, qual è la probabilità che due di queste abbiano il compleanno nello stesso giorno?

---

In generale, se le persone sono $N$, la probabilità che almeno due abbiano il compleanno nello stesso giorno è

$$P(E) = 1 - \frac{364}{365}\frac{363}{365} \times \ldots \times \frac{365-N+2}{365}\frac{365-N+1}{365}$$

A titolo di esempio, per alcuni valori di $N$ si ottiene quanto segue

| $N$ | 10 | 20 | 30 | 40 | 50 | 60 |
|---|---|---|---|---|---|---|
| $P(E)$ | 0.09 | 0.38 | 0.68 | 0.88 | 0.96 | 0.99 |

# Compleanno

In una stanza ci sono 30 persone, qual è la probabilità che una di queste abbia il compleanno nel mio stesso giorno?

# Compleanno

In una stanza ci sono 30 persone, qual è la probabilità che una di queste abbia il compleanno nel mio stesso giorno?

Vogliamo calcolare la probabilità dell'evento

$$E = \{\text{uno dei 30 è nato il } \ldots\}$$

osserviamo che è

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

dove $\bar{E} = \{\text{nessuno è nato il } \ldots\}$, per il quale si ha

$$P(\bar{E}) = \frac{364}{365}\frac{364}{365} \times \ldots \times \frac{364}{365}\frac{364}{365} = \left(\frac{364}{365}\right)^{30} \approx 0.92$$

e quindi $P(E) \approx 0.08$.

# Compleanno

In una stanza ci sono 30 persone, qual è la probabilità che una di queste abbia il compleanno nel mio stesso giorno?

---

In generale, se le persone sono $N$, la probabilità che una abbia il compleano in una specifica data è

$$P(E) = 1 - \left(\frac{364}{365}\right)^N$$

A titolo di esempio, per alcuni valori di $N$ si ottiene quanto segue

| $N$ | 10 | 20 | 30 | 40 | 50 | 60 | 100 | 250 | 500 |
|---|---|---|---|---|---|---|---|---|---|
| $P(E)$ | 0.03 | 0.05 | 0.08 | 0.10 | 0.13 | 0.15 | 0.24 | 0.50 | 0.75 |

# Esempio

Sono dati due eventi $E_1$ ed $E_2$ tali che $P(E_2) = 0.7$ e $P(E_1 \cup E_2) = 0.8$. Trovare $P(E_1)$ nelle seguenti circostanze:

a. i due eventi $E_1$ ed $E_2$ sono incompatibili;
b. i due eventi $E_1$ ed $E_2$ sono tali che $P(E_1 \mid E_2) = 0.55$;
c. i due eventi $E_1$ ed $E_2$ sono indipendenti.

# Esempio

Siano dati due eventi $A$ e $B$ dello spazio fondamentale $\Omega$, è noto che $P(\overline{A}) = 0,3, p(B) = 0,4$ e $P(A \cap \overline{B}) = 0,5$. Determinare

a. $P(A)$
b. $P(A \cup B)$
c. $P(A \cap B)$.

# Esempio

Uno studente intende presentarsi agli esami di Matematica e Statistica in una sessione, si ritiene che sia pari a 0.6 la probabilità di superare Matematica e 0.32 quella di superare Statistica, la probabilità di superarli entrambi è 0.24.

a. Il fatto di superare il secondo esame è indipendente dal fatto di superare il primo?
b. Sapendo che lo studente ha superato il primo esame, qual è la probabilità che superi il secondo?
c. Qual è la probabilità che nella sessione superi un solo esame?

# Esempio

Due tetraedri regolari con le facce numerate da 1 a 4 vengono lanciati ripetutamente finché non compare un totale pari a 5 sulle facce rivolte verso il basso.

a. Qual è la probabilità che siano necessari più di due lanci?
b. Se si effettuano 10 lanci della coppia di tetraedri, qual è la probabilità che si ottenga almeno 1 volta un totale pari a 8?

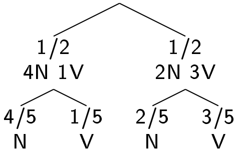
# Esempio

Un campione di prodotti manufatturieri viene sottoposto a controllo. Si scopre che il 5% ha un difetto di tipo A, il 20% un difetto di tipo B e che il 3% manifesta difetti di entrambi i tipi. Si valuti la probabilità che un prodotto dell'azienda

a. sia (in qualche modo) difettoso
b. abbia il difetto A oppure B ma non entrambi
c. abbia il difetto A ma non B
d. non sia difettoso

# Esempio

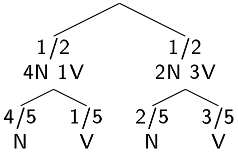
L'urna A contiene 4 palline nere e 1 verde; l'urna B contiene 2 palline nere e 3 verdi. Si sceglie a caso un'urna e si estrae dall'urna prescelta una pallina. Qual è la probabilità che la pallina sia verde?




| | | Estrazione | |
|---|---|---|---|
| | | N | V |
| Urna | A | $A \cap N$ | $A \cap V$ |
| | B | $B \cap N$ | $B \cap V$ |
| | | | |

# Esempio

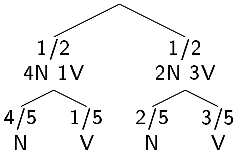
L'urna A contiene 4 palline nere e 1 verde; l'urna B contiene 2 palline nere e 3 verdi. Si sceglie a caso un'urna e si estrae dall'urna prescelta una pallina. Qual è la probabilità che la pallina sia verde?




| | | Estrazione | |
|---|---|---|---|
| | | N | V |
| Urna | A | $P(A \cap N) = 4/10$ | $P(A \cap V) = 1/10$ |
| | B | $P(B \cap N) = 2/10$ | $P(B \cap V) = 3/10$ |

# Esempio

L'urna A contiene 4 palline nere e 1 verde; l'urna B contiene 2 palline nere e 3 verdi. Si sceglie a caso un'urna e si estrae dall'urna prescelta una pallina. Qual è la probabilità che la pallina sia verde?




| | | Estrazione | |
|---|---|---|---|
| | | N | V |
| Urna | A | $P(A \cap N)=4/10$ | $P(A \cap V)=1/10$ |
| | B | $P(B \cap N)=2/10$ | $P(B \cap V)=3/10$ |
| | | | $P(V)=4/10$ |

$$P(V) = P(A \cap V) + P(B \cap V)$$

# Indice

# Formula delle probabilità totali

Nei due esempi sopra si è impiegata, per ricavare la probabilità di un evento $E$ la regola

**Probabilità totali /1**

$$P(E) = P(H)P(E|H) + P(\bar{H})P(E|\bar{H})$$

Per ottenere la probabilità non condizionata dell'evento $E$

- si sono calcolate le probabilità di $E$ condizionate a due eventi incompatibili e esaustivi (la cui unione dà l'evento certo),
- si sono combinate le due condizionate pesandole per le probabilità degli eventi condizionanti.

# Formula delle probabilità totali: dimostrazione

**Probabilità totali /1**

$$P(E) = P(H)P(E|H) + P(\bar{H})P(E|\bar{H})$$

Si noti che

$$E = E \cap (H \cup \bar{H}) = (E \cap H) \cup (E \cap \bar{H})$$

e quindi

$$\begin{aligned} P(E) &= P((E \cap H) \cup (E \cap \bar{H})) \\ &= P(E \cap H) + P(E \cap \bar{H}) \\ &= P(H)P(E|H) + P(\bar{H})P(E|\bar{H}) \end{aligned}$$

# Formula delle probabilità totali, in generale

## Probabilità totali /2

Dati gli eventi $H_1, \dots, H_k$

- a due a due incompatibili ($H_i \cap H_j = \varnothing$ se $i \neq j$)
- esaustivi $\left(\bigcup_{j=1}^{k} H_j = \Omega\right)$

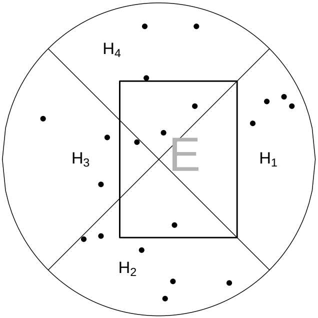
si ha

$$P(E) = \sum_{j=1}^{k} P(H_j)P(E|H_j)$$

# Formula delle probabilità totali, dimostrazione

$$
\begin{aligned}
\sum_{j=1}^{k} P(H_j)P(E|H_j) &= \sum_{j=1}^{k} P(E \cap H_j) \quad \text{[teorema Bayes]} \\
&= P\left(\bigcup_{j=1}^{k}(E \cap H_j)\right) \quad \text{[prob. unione]} \\
&= P\left(E \cap \bigcup_{j=1}^{k} H_j\right) \quad \text{[prop. operaz. insiemi]} \\
&= P\left(E \cap S\right) \quad [H_j \text{ partizione}] \\
&= P(E)
\end{aligned}
$$

# Formula delle probabilità totali, raffigurazione




- si assuma che gli eventi elementari (20) siano equiprobabili
- i quattro “spicchi” rappresentano la partizione
- il rettangolo è l’evento $E$

si calcoli dunque $P(E)$ direttamente o secondo la formula delle probabilità totali

# Esempio

Si ha un'urna con 2 palline nere.
Si lancia un dado a sei facce e si aggiunge all'urna un numero di palline bianche pari al risultato del lancio del dado.
Dall'urna si estrae una pallina, qual è la probabilità sia nera?

# Esempio

Da un'urna con 8 palline bianche e 2 nere si estraggono due palline senza reimbussolamento.
Queste vengono inserite in una seconda urna, che contiene, inizialmente, 5 bianche e tre nere.
Si estrae infine una pallina dalla seconda urna, qual è la probabilità che sia bianca?

# Indice

## Un giallo: il delitto dell'isola

- Causa una tempesta, l'isola di Bayes è isolata dalla terraferma
- Le 100 001 persone che si trovano sull'isola non possono perciò lasciarla, né qualcuno può arrivarvi.

*Alberto Pullicino, XVIII secolo*

# Un giallo: il delitto dell'isola

**Bayes oggi**

Domenica 30 febbraio 2016

## Un delitto casuale?

Si trovava sull'isola di Bayes in gita di piacere l'uomo di mezz'età trovato ucciso questa mattina. Il corpo esanime è stato ritrovato grazie alla segnalazione di un pensionato che portava a passeggio il cane. Secondo i primi rilievi delle forze dell'ordine, l'uomo è stato ucciso con un coltello da cucina nella notte di ieri. La polizia brancola nel buio, le circostanze fanno infatti pensare a un delitto casuale, la vittima non conosceva nessuno sull'isola e quindi la ricerca di un movente pare vana. Chiunque potrebbe essere il colpevole.

# Le indagini

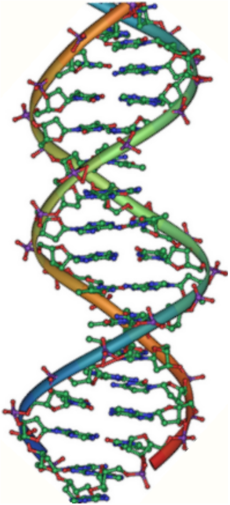
Identità della vittima: sconosciuta
Sospetti: nessuno

<u>Esame del luogo del delitto:</u> a seguito dell'ispezione sul corpo della vittima, si è accertato che
- il fatto è ad opera di un unico individuo,
- sono state rinvenute tracce biologiche che certamente appartengono all'assassino e da cui è stato possibile estrarre il DNA dello stesso

**Si procederà perciò al confronto del DNA trovato con quello delle persone presenti sull'isola.**

(Ipotesi semplificatrice: si assume che solo un DNA sia compatibile (no gemelli))

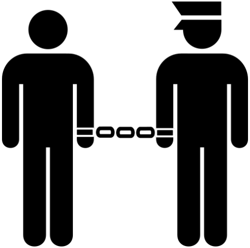
# Il test del DNA



- Il DNA del colpevole è confrontato con quello di ciascuno dei 100 000 abitanti.
- Esito positivo al test significa che il DNA è compatibile secondo il test con quello del colpevole.
- Il test non è infallibile!
  - In un caso su 10000, un DNA non compatibile dà esito positivo.
  - In un caso su 100000, un DNA compatibile dà esito negativo.

(La probabilità che si attribuisce generalmente al falso positivo in un test del DNA è molto più bassa, andrebbe però tenuto conto della probabilità di errore umano in laboratorio, che alcuni valutano all'1%: insomma, 1 su 10000 è una misura non esagerata.)

# Il test del DNA



- Il primo test dà esito "positivo", cioè il DNA del soggetto risulta compatibile.
- Il soggetto che ha dato esito positivo viene arrestato.
- Non vengono effettuati ulteriori test.
- Al processo, la prova del DNA è l'unico elemento portato dall'accusa, che però sostiene
- "Solo in un caso su 10000 la prova del DNA potrebbe essere sbagliata, l'imputato va condannato!"

# Colpevole o innocente?

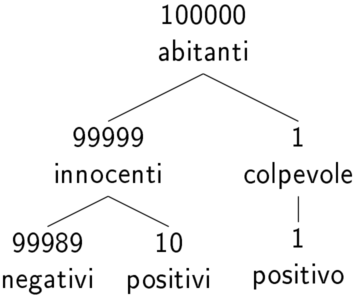
Qual è (circa) la probabilità che sia innocente?

A. Una su 100000
B. Una su 10000
C. Una su 10
D. Dieci su undici

# Esiti attesi del test su tutti gli abitanti

Supponiamo di fare il test a tutti, cosa succederebbe?

*I 99999 innocenti hanno tutti DNA incompatibile, ma per effetto dei falsi positivi, mediamente lo 0.1% è positivo, cioè 10.*




*Il colpevole risulterà positvo con elevata probabilità.*

# Esiti attesi del test su tutti gli abitanti

Insomma, ci ritroveremmo con questa situazione

| | colpevole | innocente |
|---|---|---|
| positivo | 1 | 10 |
| negativo | 0 | 99989 |

Cioè con 11 positivi, di cui uno solo colpevole, alla luce di questo, qual è la probabilità che il nostro positivo sia colpevole?

## In termini di probabilità

Definiamo gli eventi (riferiti a un individuo)

- C : è colpevole (equivale a “il DNA è compatibile’);
- T : risulta positivo al test

quello che ci interessa è: “Sapendo che il test è positivo, qual è la probabilità che l’individuo sia colpevole?”, cioè vogliamo

$$P(C|T)$$

Ricordando che

$$P(T \cap C) = P(T)P(C|T) = P(C)P(T|C)$$

possiamo scrivere che

$$P(C|T) = \frac{P(C)P(T|C)}{P(T)}$$

e di questa espressione abbiamo tutti gli elementi.

# In termini di probabilità

$$P(C|T) = \frac{P(C)P(T|C)}{P(T)}$$

- essendo tutti potenzialmente colpevoli è

$$P(C) = 1/100000 = 0.0001$$

- per quanto riguarda $T$ sappiamo che
  - $P(T|C) = 1 - 1/100000$
  - $P(T|\bar{C}) = 1/10000$

  usando la formula per le probabilità totali quindi

$$\begin{aligned} P(T) &= P(T|C)P(C) + P(T|\bar{C})P(\bar{C}) = \\ &= \frac{99999}{100000}\frac{1}{100000} + \frac{10}{100000}\frac{99999}{100000} \approx \frac{11}{100000} \end{aligned}$$

# In termini di probabilità

$$P(C|T) = \frac{P(C)P(T|C)}{P(T)}$$

- essendo tutti potenzialmente colpevoli è

$$P(C) = 1/100000 = 0.0001$$

- per quanto riguarda $T$ : $P(T) \approx \frac{11}{100000}$

# In termini di probabilità

$$P(C|T) = \frac{P(C)P(T|C)}{P(T)}$$

- essendo tutti potenzialmente colpevoli è

$$P(C) = 1/100000 = 0.0001$$

- per quanto riguarda $T$ : $P(T) \approx \frac{11}{100000}$

Sostituendo si ottiene

$$P(C|T) = \frac{\frac{1}{100000}\frac{99999}{100000}}{\frac{11}{100000}} \approx \frac{1}{11}$$

# Teorema di Bayes

**Teorema di Bayes**

Dati due eventi $C$ e $T$ si ha

$$P(C|T) = \frac{P(C)P(T|C)}{P(T)}$$

# Teorema di Bayes 2

Il teorema di Bayes si può scrivere con riferimento a un evento $E$ e a $k$ eventi $H_1, \ldots .H_k$ che costituiscano una partizione, cioè tali che

$$H_i \cap H_j = \varnothing \text{ per ogni } i \neq j$$

$$\bigcup_{i=1}^{k} H_i = S \text{ (evento certo)}$$

si ha allora

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{P(E)} = \frac{P(H_i)P(E|H_i)}{\sum_{j=1}^{k} P(H_j)P(E|H_j)}$$

# Esempio, test diagnostici

L'incidenza di tumore alla mammella nella popolazione è l'1%. La mammografia è un comune test diagnostico per la patologia, come tutti i test non è infallibile: nel 10% dei pazienti affetti da tumore al seno il test non lo rileva (falso negativo); d'altra parte, il test rileva la malattia nel 10% dei pazienti che non hanno il cancro (falso positivo).

—

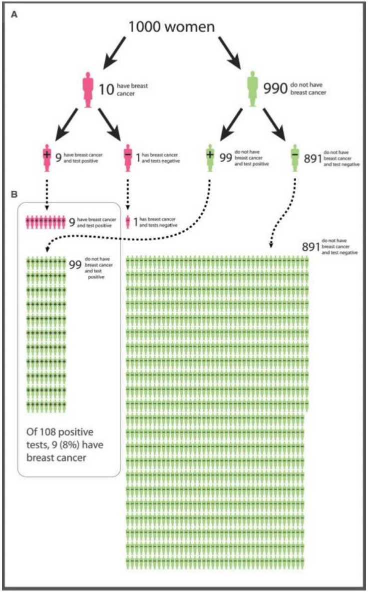
Se una donna, presa a caso, viene sottoposta a mamografia e questa fornisce un risultato positivo, qual è la probabilità che sia effettivamente malata?

# Esempio, test diagnostici

L'incidenza di tumore alla mammella nella popolazione è l'1%. La mammografia è un comune test diagnostico per la patologia, come tutti i test non è infallibile: nel 10% dei pazienti affetti da tumore al seno il test non lo rileva (falso negativo); d'altra parte, il test rileva la malattia nel 10% dei pazienti che non hanno il cancro (falso positivo).

—

Se una donna, presa a caso, viene sottoposta a mamografia e questa fornisce un risultato positivo, qual è la probabilità che sia effettivamente malata?



# Esercizio per spettatori di Dr. House

*"It's never lupus"*
Il Lupus è una patologia che porta gli anticorpi ad attaccare le proteine del plasma come se fossero cellule estranee, il che comporta un elevato rischio di trombi.
Si ritiene che il 2% della popolazione soffra di Lupus. Il test per la malattia è accurato al 98% in presenza della malattia, al 74% in assenza della stessa. Nella serie televisiva "Dr. House" viene spesso ripetuta la battuta "Non è mai Lupus" quando qualcuno ha un risultato positivo al test. Ala luce delle informazioni sopra, ritenete che la battuta sia appropriata?

# Gemelli

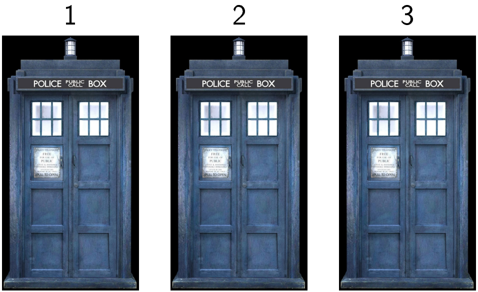
Circa il 30% dei gemelli è omozigote (gemelli identici), il restosono eterozigoti (gemelli fraterni).
I gemelli identici sono necessariamente dello stesso sesso, nel 50% dei casi maschi, nel 50% femmine. Le coppie di gemelli fraterni sono invece per il 25% entrambi maschi, per il 25% entrambi femmine, per il restante 50% misti.
Sapendo che una coppia di gemelli è costituita da due ragazze, qual è la probabilità che queste siano gemelle identiche?

# Indice

# Monty hall

- Il problema di Monty Hall fu proposto nel 1975 su *American Statistician* da *Steve Selvin*;
- ispirato al gioco a premi americano *Let's Make a Deal* (Monty Hall era lo pseudonimo del presentatore).

# Monty hall




- Dietro una porta c'è un automobile, dietro le altre due capre.
- Il concorrente apre una porta e vince quello che c'è dietro.

# Monty hall

1 2 3


scelta

- È stata scelta la porta 1.
- Conduttore (che sa dove sono i premi): "Per aiutarla, ora le aprirò, delle altre due porte, una che cela una capra"

# Monty hall

1 2 3


scelta

- È stata scelta la porta 2.
- Conduttore (che sa dove sono i premi): "Per aiutarla, ora le aprirò, delle altre due porte, una che cela una capra"

# Monty hall

1


2


3


scelta

- È stata scelta la porta 3.
- Conduttore (che sa dove sono i premi): “Per aiutarla, ora le aprirò, delle altre due porte, una che cela una capra”

# Monty hall

1 2 3


scelta

- Conduttore: “Ora, se vuole, può cambiare la sua scelta”
- Conviene cambiare? ▸ Soluzione

# Monty hall

1 2 3


scelta

- Conduttore: "Ora, se vuole, può cambiare la sua scelta"
- Conviene cambiare? Soluzione

# Monty hall

1


2


scelta

3


- Conduttore: “Ora, se vuole, può cambiare la sua scelta”
- Conviene cambiare? ▸ Soluzione

# Monty hall

1


2


3

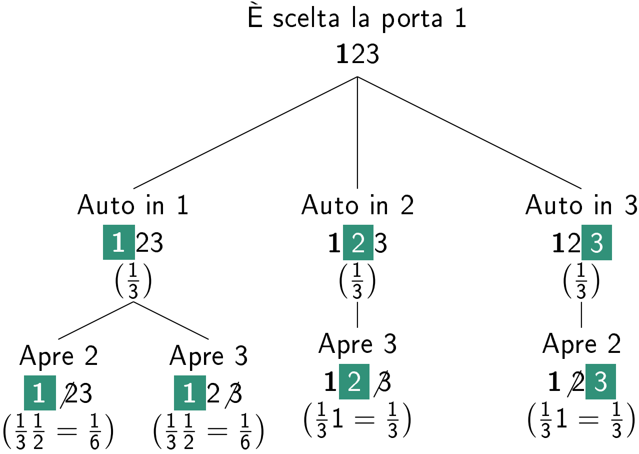

scelta

- Conduttore: “Ora, se vuole, può cambiare la sua scelta”
- Conviene cambiare? Soluzione

# Monty hall: la risposta

- $E_i$ ="il premio è dietro la porta $i$";
- $P(E_1) = P(E_2) = P(E_3) = \frac{1}{3}$;
- Il concorrente sceglie la porta 1, quindi vince se e solo se è vero $E_1$, con probabilità 1/3.
- Consideriamo le due porte rimanenti
  - (a) se è vero $E_1$ il conduttore apre indifferentemente la porta 2 o 3;
  - (b) se è vero $E_2$ il conduttore apre la porta 3 e la porta non scelta che rimane chiusa è quella vincente;
  - (c) se è vero $E_3$ il conduttore apre la porta 2 e la porta non scelta che rimane chiusa è quella vincente;
- i tre casi sono equiprobabili, in due su tre conviene cambiare perché la porta non scelta e rimasta chiusa è quella vincente.
- Cambiando, si vince con probabilità 2/3.

# Monty hall: albero



# Monty hall: porte aperte

1


2


3

# Monty Hall con 100 porte

Ci sono 100 porte, di cui una sola contiene il premio.

Il concorrente sceglie la porta 1.

Il conduttore apre 98 delel restanti mostrando che hanno una capra.

# Monty Hall con 100 porte

Ci sono 100 porte, di cui una sola contiene il premio.

Il concorrente sceglie la porta 1.

Il conduttore apre 98 delel restanti mostrando che hanno una capra.

# Indice

# La storia

Una storia in cui il calcolo delle probabilità entra in un'aula giudiziaria in Gran Bretagna: il caso di SC.

In breve, la storia è questa

- 26/9/1996 nasce il primo figlio di SC
- 13/12/1996 il bambino viene trovato morto, il decesso viene considerato morte naturale.
- 29/11/1997 SC ha un secondo figlio.
- 26/1/1998 anche il secondo figlio viene trovato morto in circostanze analoghe al primo.
- 23/2/1998 Nonostante non ci siano prove materiali di maltrattamento o simili, SC viene accusata di duplice infanticidio.

Per maggiori dettagli si veda, ad esempio, wikipedia.

## Tra le prove ...

Un esperto medico testimonia che

- l'incidenza delle morti in culla nelle famiglie benestanti e di non fumatori (come quella di SC) è circa 8 500 : 1 (cioè un caso ogni 8 500 nati);
- di conseguenza la probabilità che a S.C. sia successo per entrambi i figli è circa

$$\frac{1}{8\,500} \times \frac{1}{8\,500} = \frac{1}{8\,500 \times 8\,500} \approx \frac{1}{73\,000\,000}$$

- Questo significa, dice l'esperto, che una circostanza del genere dovrebbe accadere una volta ogni cento anni in GB.

**Siamo (siete) d'accordo?**

# Errore: indipendenza

- Diamo per buono che la probabilità 1/8 500 sia applicabile al caso,
- il dire che la probabilità che accada due volte è il prodotto delle probabilità significa che riteniamo le due cose indipendenti.
- Questo è poco ragionevole se non altro perché non si possono escludere concause
  - ambientali: i due bambini vivono nelle stesse condizioni;
  - genetiche: condividono parte del patrimonio genetico.
- (Incidentalmente, vi sono teorie secondo cui esiste una predisposizione genetica alla morte in culla, anche se non vi fossero, comunque, la loro possibilità metterebbe in dubbio il calcolo.)

# Errore: probabilità dell'alternativa

La probabilità calcolata (erroneamente) è quello di una *duplice morte in culla (mic)*

$$P(duplice\ mic)$$

In realtà noi sappiamo che c'è stato un*duplice decesso*, per cui quello che interessa è la condizionata

$$P(duplice\ mic|duplice\ decesso)$$

e in particolare interessa quanto è probabile che ci siano state due morti in culla rispetto a quanto è probabile che ci sia stato un *duplice infanticidio*

$$\begin{aligned}\frac{P(dup\ mic|dup\ decesso)}{P(dup\ inf|dup\ decesso)} &= \frac{P(dup\ decesso|dup\ mic)}{P(dup\ decesso|dup\ inf)}\frac{P(dup\ mic)}{P(dup\ inf)}\\ &= \frac{P(dup\ mic)}{P(dup\ inf)}\end{aligned}$$

## La storia giudiziaria

9/11/1999 SC viene ritenuta colpevole a maggioranza (10-2) e condannata al carcere a vita.

10/2000 La sentenza viene confermata in appello.

Successivamente, la 'Criminal Cases Review Commission' reinvia il caso alla corte d'appello, per via di una prova medica non resa nota alla difesa.

1/2003 La (seconda) corte d'appello scagiona SC, dichiarando che anche la sola fallacia della testimoniaza dell'esperto medico sarebbe stata ragione sufficiente per invalidare il verdetto di primo grado.

# Calcolo corretto: un unico decesso

Qual è la probabilità di una morte in culla (che indichiamo con $C$)

- In uno studio affidabile si è osservata un incidenza generale di $P(C) = 1/1303$;
- Nello stesso studio si osservano delle caratteristiche familiari che mitigano il rischio (fam. benestante, non fumatori, madre di più di 27 anni e senza fratelli/sorelle), in presenza di queste $P(C|F) = 1/8543$.

Qual è la probabilità che un decesso ($D$) sia un infanticidio ($I$)?

$$P(I|D) = \frac{P(D|I)P(I)}{P(D)} = \frac{P(D|I)P(I)}{P(D|I)P(I) + P(D|A)P(C)} = \frac{P(I)}{P(I) + P(C)}$$

(Si suppone per semplicità che non vi siano altre cause possibili.)

- Supponendo $P(I) = 30/650000$ (sulla base delle statistiche sulla criminalità) si ha $P(I|D) = 0.28$
- Se supponessimo $P(C) = 1/1303$ otterremmo $P(I|D) = 0.06$.

## Calcolo corretto: due decessi

Qual è la probabilità di due morti in culla ($C_1 C_2$)?

- nello studio citato su 423326 casi vi sono 323 morti in culla e 5 doppie morti in culla da cui si avrebbe $P(C_1 C_2) = 5/423326 = 1/84000$.
- (Il che suggerisce non vi sia indipendenza.)

La probabilità del duplice infanticidio ($I_1 I_2$) date due morti ($D_1 D_2$)

$$\begin{aligned} P(I_1 I_2 | D_1 D_2) &= \frac{P(D_1 D_2 | I_1 I_2) P(I_1 I_2)}{P(D_1 D_2)} = \\ &= \frac{P(D_1 D_2 | I_1 I_2) P(I_1 I_2)}{P(D_1 D_2 | I_1 I_2) P(I_1 I_2) + P(D_1 D_2 | C_1 C_2) P(C_1 C_2)} \\ &= \frac{P(I_1 I_2)}{P(I_1 I_2) + P(C_1 C_2)} \end{aligned}$$

Il problema è valutare $P(I_1 I_2)$, cioè valutare $P(I_2 | I_1)$

# Calcolo corretto: due decessi

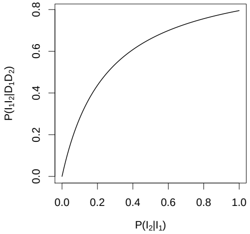
Qual è la probabilità di due morti in culla ($C_1 C_2$)?

- nello studio citato su 423326 casi vi sono 323 morti in culla e 5 doppie morti in culla da cui si avrebbe $P(C_1 C_2) = 5/423326 = 1/84000$.
- (Il che suggerisce non vi sia indipendenza.)

La probabilità del duplice infanticidio ($I_1 I_2$) date due morti ($D_1 D_2$)

$$P(I_1 I_2 | D_1 D_2) = \frac{P(I_1 I_2)}{P(I_1 I_2) + P(C_1 C_2)}$$

Il problema è valutare $P(I_1 I_2)$, cioè valutare $P(I_2 | I_1)$




Nel grafico a sinistra riportiamo la probabilità di duplice infanticidio condizionato al verificarsi delle due morti in funzione della probabilità di un secondo infanticidio condizionato al primo infanticidio.