

# Come sfruttare le informazioni?

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

La probabilità è uno strumento che usiamo quando abbiamo *incertezza* sull'esito dei nostri esperimenti.

Tuttavia ci sono situazioni in cui abbiamo una *conoscenza parziale* dell'esito.

Questa conoscenza potrebbe alterare la probabilità che assegniamo a certi eventi.

## Esempio: geni e fiori

### Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

### Indipendenza di eventi

Esempio

Caveat

### Teorema delle probabilità totali

### Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

### Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

### Approfondiamo: famiglie indipendenti



Torniamo alle piante in cui gli individui AA, Aa hanno fiori rossi; quelli aa fiori bianchi.

## Esempio: geni e fiori

### Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

### Indipendenza di eventi

Esempio

Caveat

### Teorema delle probabilità totali

### Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

### Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

### Approfondiamo: famiglie indipendenti

Incrociando due piante  $Aa$ : la probabilità che la pianta figlia sia  $AA$  è  $1/4$  (casi possibili  $\{AA, Aa, aA, aa\}$ ).

Supponiamo di sapere che la figlia ha fiori rossi: qual è la probabilità che sia  $AA$ ?

Ora abbiamo un'informazione in più: il colore, quindi i casi possibili sono  $\{AA, Aa, aA\}$  e la probabilità richiesta è  $1/3$ .

Conoscendo il colore, la probabilità di essere  $AA$  è diversa rispetto al caso in cui il colore è sconosciuto ( $\implies$  colore e genoma non sono indipendenti).

# Probabilità condizionata classica

Analizziamo il caso della probabilità classica. Dato  $B \neq \emptyset$ ,  $B \subseteq \Omega$  ed  $A \subseteq \Omega$  la probabilità che avvenga  $A$  sapendo che  $B$  avviene è:

$$\frac{\#(A \cap B)}{\#B} = \frac{\#(A \cap B) / \#\Omega}{\#B / \#\Omega} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Si considera  $B$  come nuovo spazio campionario e si fa riferimento solo agli eventi elementari che appartengono sia ad  $A$  che a  $B$ .

# Probabilità condizionata

## Probabilità condizionata

Esempio: genotipo e fenotipo

### Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

## DEFINIZIONE DI PROBABILITÀ CONDIZIONATA

Dati due eventi  $A$  e  $B$  con  $\mathbb{P}(B) > 0$ , si dice probabilità di  $A$  dato  $B$  (o probabilità di  $A$  condizionata al verificarsi di  $B$ ) il seguente numero:

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

La probabilità  $\mathbb{P}(A|B)$  è la probabilità che attribuiamo ad  $A$  quando sappiamo che **si è verificato  $B$** .

Nell'esempio,  $B$  è "il fiore della pianta figlia è rosso".

# Condizionare = modificare $\Omega$

La conoscenza parziale che abbiamo fa cambiare  $\Omega$ .  
Infatti  $\Omega$  è il “mondo dei casi possibili” e spesso le informazioni ci fanno restringere questo mondo (ad un evento  $B$ ).

Nel caso dei fiori siamo passati da  $\Omega = \{AA, Aa, aA, aa\}$  a  $B = \{AA, Aa, aA\}$ .

La probabilità è quella definita con la formula classica.  
Allora se  $\mathbf{A} = \{AA\}$  è “l’individuo è  $AA$ ”,

$$\begin{aligned}\mathbb{P}(\mathbf{A}|B) &= \frac{\mathbb{P}(\mathbf{A} \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{AA\})}{3/4} \\ &= \frac{1/4}{3/4} = \frac{1}{3}.\end{aligned}$$

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio  
Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo  
Test clinici  
Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg  
Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

# Perché la probabilità condizionata

Definisce la probabilità *utilizzando informazioni aggiuntive venute a disposizione.*

## TEOREMA

Data una probabilità  $\mathbb{P}$  su  $\Omega$  e fissato un evento  $B$  con  $\mathbb{P}(B) > 0$ , allora la funzione  $\mathbb{P}(\cdot|B)$  che all'evento  $A$  associa  $\mathbb{P}(A|B)$  è una probabilità su  $\Omega$ .

Notiamo ancora che la cosa non stupisce, visto che *condizionare a  $B$*  significa sapere che il “mondo dei casi possibili” non è tutto  $\Omega$ , ma solo  $B$  (detto in altro modo, ci si restringe alla *sottopopolazione  $B$* ).

# Proprietà della probabilità condizionata

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

Si può calcolare la probabilità condizionata a un evento fissato  $B$  utilizzando tutte le proprietà viste per le probabilità.

## PROPOSIZIONE

Data una probabilità  $\mathbb{P}$  su  $\Omega$  e fissato un evento  $B$  con  $\mathbb{P}(B) > 0$ , allora

- 1  $\mathbb{P}(\emptyset|B) = 0$ .
- 2 Se  $A_1, \dots, A_n$  sono a due a due disgiunti, allora

$$\mathbb{P}(A_1 \cup \dots \cup A_n|B) = \mathbb{P}(A_1|B) + \dots + \mathbb{P}(A_n|B).$$

- 3  $\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B)$ .
- 4  $\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) - \mathbb{P}(A_1 \cap A_2|B)$ .



# Indipendenza di eventi

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

La gente comune fa spesso confusione sulle probabilità condizionate, e fra l'evento di cui si valuta la probabilità e l'evento noto.

La cosa diventa ancora più evidente quando si trattano argomenti sensibili come il legame fra comportamenti a rischio e certe malattie.

Analizziamo ad esempio una frase che troviamo sui pacchetti di sigarette: “Il fumo causa cancro ai polmoni”.

# Significato della dipendenza stocastica

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

Cosa si vuole dire?

Che tutti i fumatori contraggono il tumore al polmone?

Questo vorrebbe dire che l'insieme dei fumatori **è contenuto** in quello dei malati.

Ma questa è quella che abbiamo chiamato **implicazione logica**.

Con le formule della probabilità condizionata, sarebbe

$$\mathbb{P}(\text{contrarre la malattia} | \text{essere fumatore}) = 1.$$

I medici vogliono forse dire questo?

Ovviamente no! (chi non ha un conoscente morto nel sonno in tardissima età, che ha sempre fumato in barba ai medici e senza mai ammalarsi?).

# Significato della dipendenza stocastica

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

Allora forse si intende che tutti i malati di tumore al polmone sono fumatori? Questo vorrebbe dire che l'insieme dei malati **è contenuto** in quello dei fumatori.

Ma questa è un'altra **implicazione logica**!

Con le formule della probabilità condizionata, sarebbe

$$\mathbb{P}(\text{essere fumatore} | \text{aver contratto la malattia}) = 1.$$

Anche se questa seconda affermazione si avvicina alla verità (vedremo fra poco i dati numerici) anche questa è falsa (ci sono malati che non hanno mai fumato).

# Probabilità condizionata in medicina

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

Allora cosa intendono i medici per “comportamento a rischio” per una malattia? Di solito intendono che

$$\mathbb{P}(\text{contrarre la malattia} | \text{comportamento a rischio}) > \mathbb{P}(\text{contrarre la malattia}),$$

cioè che l'incidenza della malattia in chi ha un determinato comportamento a rischio è maggiore (spesso di molte volte) dell'incidenza nella popolazione generale.

Dal punto di vista matematico

questo è un esempio di eventi che **non sono indipendenti**.

# Indipendenza stocastica di due eventi

## DEFINIZIONE DI COPPIA DI EVENTI INDIPENDENTI

Due eventi  $A$  e  $B$  si dicono indipendenti se vale l'uguaglianza  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ .

Notate che *nel caso che*  $\mathbb{P}(B) > 0$  la definizione equivale a

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

Si preferisce la definizione sopra perché per verificarla non c'è bisogno che  $\mathbb{P}(B) > 0$  (anzi se  $\mathbb{P}(B) = 0$  allora  $B$  è indipendente da qualsiasi evento: esercizio...).

Notate che  $\mathbb{P}(A|B) = \mathbb{P}(A)$  significa che **sapere che si è verificato  $B$  non cambia la valutazione della probabilità di  $A$ .**

## Esempio di eventi indipendenti

Peschiamo a caso una carta da un mazzo di 52 carte (le francesi). Siano  $A$  l'evento "la carta è un re" e  $B$  "la carta è cuori". Sono indipendenti?

$$\mathbb{P}(A) = 4/52; \mathbb{P}(B) = 13/52$$

$A \cap B$  = "la carta è il re di cuori", quindi

$$\mathbb{P}(A \cap B) = 1/52.$$

Andiamo a vedere se  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$  (se sì risponderemo che  $A$  e  $B$  sono indipendenti).

$\mathbb{P}(A)\mathbb{P}(B) = \frac{4}{52} \frac{13}{52} = \frac{1}{52}$  che è uguale a  $\mathbb{P}(A \cap B) = 1/52$ , quindi la risposta è che sono indipendenti.

Sapere che la carta estratta è cuori non cambia la valutazione della probabilità che sia un re (la frequenza dei re fra i cuori è la stessa che fra tutte le carte).

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

# Indipendenti $\neq$ incompatibili

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

I due concetti sono diversi!

Se due eventi  $A$  e  $B$  sono incompatibili, sapere che si è verificato  $A$  **cambia** la probabilità di  $B$  (a meno che non sia  $\mathbb{P}(B) = 0$ ).

Infatti se si è verificato  $A$  a quel punto so per certo che  $B$  non può essersi verificato (perché?)

# Teorema delle probabilità totali

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

A volte non conosciamo la probabilità di un evento, ma la conosceremmo *se ci fossero altre informazioni*.

Facciamo un esempio: le nostre piante di piselli, che hanno fiori bianchi solo se sono *aa*.

Se incrociamo una pianta con fiori rossi con una con fiori bianchi, qual è la probabilità che la pianta figlia abbia fiori rossi?



## Esempio dei fiori

Non sappiamo rispondere alla domanda ma...

se conoscessimo il genoma della pianta con fiori rossi la risposta sarebbe ovvia.

Infatti se la pianta ha genoma  $Aa$  allora la probabilità che la figlia abbia fiori bianchi (cioè sia  $aa$ ) è  $1/2$ , quindi anche la probabilità che li abbia rossi è  $1/2$ .

Se invece ha genoma  $AA$  allora tutti i figli avranno fiori rossi e la probabilità richiesta è 1.

Notate che la pianta a fiori rossi può avere solo genoma  $Aa$  o genoma  $AA$ , quindi ho esaurito le possibilità.

## Esempio dei fiori

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

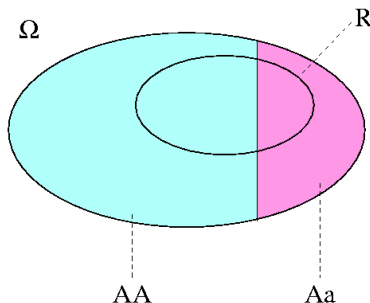
Scriviamo in formule quanto appena visto: siano  $R$  l'evento "la pianta figlia ha fiori rossi";  $AA$  l'evento "la pianta madre è  $AA$ ";  $Aa$  l'evento "la pianta madre è  $Aa$ ".

Allora

$$\mathbb{P}(R|AA) = 1; \mathbb{P}(R|Aa) = 1/2.$$

Inoltre  $AA \cup Aa = \Omega$  cioè questi due genomi esauriscono tutti i casi possibili per la pianta madre (quella con i fiori rossi).

Se dividiamo  $\Omega$  in  $AA$  (azzurro) e  $Aa$  (rosa) possiamo rappresentare



Allora possiamo scrivere

$$\mathbb{P}(R) = \mathbb{P}(R \cap AA) + \mathbb{P}(R \cap Aa).$$

La formula vale perché

$AA \cap Aa = \emptyset$  (= non ci sono parti che conto più volte) e  
 $AA \cup Aa = \Omega$  (=non ci sono parti che dimentico).

Ci siamo quasi: abbiamo ottenuto la formula

$$\mathbb{P}(R) = \mathbb{P}(R \cap AA) + \mathbb{P}(R \cap Aa).$$

e sappiamo che  $\mathbb{P}(R|AA) = 1$  e  $\mathbb{P}(R|Aa) = 1/2$ .

Possiamo ricavare  $\mathbb{P}(R \cap AA)$  da  $\mathbb{P}(R|AA)$ ? Certamente: ricordiamo la formula che definisce la probabilità condizionata:

$$\mathbb{P}(R|AA) = \frac{\mathbb{P}(R \cap AA)}{\mathbb{P}(AA)},$$

da cui  $\mathbb{P}(R \cap AA) = \mathbb{P}(R|AA)\mathbb{P}(AA)$  e idem per l'altra probabilità  $\mathbb{P}(R \cap Aa) = \mathbb{P}(R|Aa)\mathbb{P}(Aa)$ .

Eccoci finalmente alla formula:

$$\mathbb{P}(R) = \mathbb{P}(R|AA)\mathbb{P}(AA) + \mathbb{P}(R|Aa)\mathbb{P}(Aa).$$

Le sole cose che ancora mancano sono  $\mathbb{P}(AA)$  e  $\mathbb{P}(Aa)$ , queste si spera di averle dal problema.

Ad esempio supponiamo che il 25% delle nostre piante a fiori rossi sia  $AA$  e il 75% sia  $Aa$ , allora la risposta è

$$\mathbb{P}(R) = 1 \cdot 0.25 + \frac{1}{2} \cdot 0.75 = 0.625.$$

Le partizioni di  $\Omega$ 

## Nota bene

Abbiamo visto nell'esempio che per ottenere la formula era **fondamentale** suddividere  $\Omega$  in "pezzi" (= eventi) che lo "coprissero tutto" e non si sovrapponevano.

Infatti nell'esempio i "pezzi" sono gli eventi  $AA$  e  $Aa$ , e  
 $AA \cup Aa = \Omega$ ,  $AA \cap Aa = \emptyset$ .

Diamo la definizione generale.

## DEFINIZIONE DI PARTIZIONE

Dati uno spazio campionario  $\Omega$ , una sua partizione è una famiglia di eventi  $B_1, B_2, \dots, B_n$  tali che  $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$  e  $B_i \cap B_j = \emptyset$  per ogni  $i \neq j$ .

In probabilità hanno particolare rilevanza anche le partizioni numerabili, i.e.  $\{B_0, B_1, \dots, B_n, \dots\}$  tali che  $\bigcup_{i \in \mathbb{N}} B_i = \Omega$  e  $B_i \cap B_j = \emptyset$  per ogni  $i \neq j$ .

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio  
Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo  
Test clinici  
Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg  
Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

# Il teorema delle probabilità totali

## TEOREMA DELLE PROBABILITÀ TOTALI

Dato uno spazio campionario  $\Omega$ , una sua partizione  $B_1, \dots, B_n$  tale che  $\mathbb{P}(B_i) > 0$  per ogni  $i$  e un evento  $A$ , vale la seguente formula

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A \cap B_1) + \dots + \mathbb{P}(A \cap B_n) \\ &= \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \dots + \mathbb{P}(A|B_n)\mathbb{P}(B_n) \\ &= \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i).\end{aligned}$$

La formula precedente vale anche per partizioni numerabili (sostituendo una serie alla sommatoria) i.e.

$$\mathbb{P}(A) = \sum_{i \in \mathbb{N}} \mathbb{P}(A \cap B_i) = \sum_{i \in \mathbb{N}} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

Altri esempi

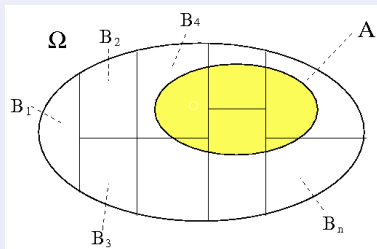
Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

D. Bertacchi  
F. Zucca

Disegniamo gli eventi:



Poiché la famiglia  $B_1, \dots, B_n$  è una partizione, vale

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A \cap B_i).$$

Dalla formula per la probabilità condizionata

$$\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Sostituire nella formula precedente e avete finito!

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti



## Formula di Bayes

Nel teorema delle probabilità totali si ha una partizione  $\{B_j\}_j$  di  $\Omega$  e si conoscono tutte le probabilità condizionate  $\mathbb{P}(A|B_j)$ . A volte ci interessa invece un particolare  $\mathbb{P}(B_k|A)$ .

Ad esempio se  $A$  indica una categoria di individui e i  $B_j$  sono una suddivisione della popolazione in sottogruppi (per esempio malattie più il sottogruppo “sani”),

$\mathbb{P}(A|B_1)$  indica la probabilità di appartenere a quella categoria se si è contratta la malattia  $B_1$  (dunque la frequenza, fra i malati  $B_1$  degli individui di categoria  $A$ ), mentre  $\mathbb{P}(B_1|A)$  indica la probabilità che un individuo della categoria  $A$  contragga la malattia  $B_1$  (dunque l'incidenza della malattia nella categoria  $A$ ).

La formula di Bayes ci fornisce un modo per calcolare  $\mathbb{P}(B_k|A)$ .

# Formula di Bayes (caso finito)

## TEOREMA

Sia  $\{B_j\}_{j=1}^n$  una partizione di  $\Omega$  con  $\mathbb{P}(B_j) > 0$  per ogni  $j$  e sia  $A$  un evento con  $\mathbb{P}(A) > 0$ . Allora per ogni  $k \in \{1, 2, \dots, n\}$  vale

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{j=1}^n \mathbb{P}(A|B_j)\mathbb{P}(B_j)}.$$

## Dimostrazione.

Basta scrivere la definizione di  $\mathbb{P}(B_k|A)$  e poi ricavare i termini in funzione delle  $\mathbb{P}(A|B_j)$ .

$$\begin{aligned} \mathbb{P}(B_k|A) &= \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{j=1}^n \mathbb{P}(A|B_j)\mathbb{P}(B_j)}, \end{aligned}$$

dove nell'ultima uguaglianza abbiamo applicato il teorema delle probabilità totali per ricavare  $\mathbb{P}(A)$  al denominatore.

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

# Formula di Bayes (caso infinito)

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

Naturalmente la formula di Bayes vale anche nel caso di una partizione numerabile (la dimostrazione è identica).

### TEOREMA

Sia  $\{B_j\}_{j \in \mathbb{N}}$  una partizione di  $\Omega$  con  $\mathbb{P}(B_j) > 0$  per ogni  $j$  e sia  $A$  un evento con  $\mathbb{P}(A) > 0$ . Allora per ogni  $k \in \mathbb{N}$  vale

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{j \in \mathbb{N}} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}.$$

# Un esempio

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

In Italia una persona presa a caso ha una probabilità pari a 0.06 di morire di tumore al polmone. Delle morti per tumore al polmone, il 90% si riscontra in fumatori, inoltre i fumatori sono il 25% della popolazione. (i dati sono presi da indagini OMS, registro dei tumori in Italia e ISTAT)

Qual è la probabilità di morire di tumore al polmone per un fumatore? E per un non fumatore?

Scriviamo i dati come probabilità degli eventi  $T$ ="la persona morirà di tumore al polmone" e  $F$ ="la persona è un fumatore/trice".

$$\mathbb{P}(T) = 0.06; \quad \mathbb{P}(F|T) = 0.90; \quad \mathbb{P}(F) = 0.25.$$

La probabilità di morire di tumore al polmone per un fumatore è

$$\mathbb{P}(T|F) = \frac{\mathbb{P}(F \cap T)}{\mathbb{P}(F)} = \frac{\mathbb{P}(F|T)\mathbb{P}(T)}{\mathbb{P}(F)} = \frac{0.90 \cdot 0.06}{0.25} = 0.216,$$

mentre per un non fumatore è

$$\begin{aligned} \mathbb{P}(T|F^c) &= \frac{\mathbb{P}(F^c \cap T)}{\mathbb{P}(F^c)} = \frac{\mathbb{P}(F^c|T)\mathbb{P}(T)}{\mathbb{P}(F^c)} = \frac{(1 - 0.90) \cdot 0.06}{(1 - 0.25)} \\ &= 0.008. \end{aligned}$$

# I test clinici

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

In un test clinico un individuo ottiene un responso che può essere *positivo* o *negativo*.

L'idea è che la positività sia associata alla presenza della malattia cercata mentre la negatività sia associata con la sua assenza.

Purtroppo possono esserci errori, ovvero dei *falsi positivi* (sani che il test dichiara malati) o *falsi negativi* (malati che il test dichiara sani).

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

## Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Fissiamo i nomi di eventi di interesse:

$M$  = l'individuo è malato;

$S$  = l'individuo è sano;

$Pos$  = il test è positivo;

$Neg$  = il test è negativo.

Sono interessanti alcune probabilità:

$\mathbb{P}(Pos|M)$  = sensibilità del test;

$\mathbb{P}(Neg|S)$  = specificità del test;

$\mathbb{P}(M|Pos)$  = predittività del test.

Queste quantità sono legate fra loro dalla formula di Bayes.

## Test clinici e Bayes

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Se conosciamo l'incidenza della malattia nella popolazione, cioè  $\mathbb{P}(M)$ , ricaviamo

$$\begin{aligned}\mathbb{P}(M|Pos) &= \frac{\mathbb{P}(Pos|M)\mathbb{P}(M)}{\mathbb{P}(Pos|M)\mathbb{P}(M) + \mathbb{P}(Pos|S)\mathbb{P}(S)} \\ &= \frac{\mathbb{P}(Pos|M)\mathbb{P}(M)}{\mathbb{P}(Pos|M)\mathbb{P}(M) + (1 - \mathbb{P}(Neg|S))(1 - \mathbb{P}(M))}.\end{aligned}$$



## Esempio: il test Elisa

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Per il test Elisa per la ricerca degli anticorpi dell'HIV si hanno i seguenti valori (si veda il libro di Bramanti):

$$\begin{aligned}\mathbb{P}(Pos|M) &= 0.993; & \mathbb{P}(Neg|S) &= 0.9999; \\ \mathbb{P}(M) &= 0.000025.\end{aligned}$$

Da qui si ricava  $\mathbb{P}(M|Pos)$

$$\frac{0.993 \cdot 0.000025}{0.993 \cdot 0.000025 + (1 - 0.9999)(1 - 0.000025)} = 0.19888.$$

Dunque la probabilità che un sieropositivo abbia la malattia è solo del 20% circa!

## Esempio: il test Elisa

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Osserviamo che comunque

$$\frac{\mathbb{P}(M|Pos)}{\mathbb{P}(M)} = 7955.2,$$

quindi la probabilità di essere malati è notevolmente aumentata una volta noto l'esito positivo del test. Il paradosso nasce comunque dal fatto che  $\mathbb{P}(M)$  è molto piccola, se ci si restringe a sottogruppi (es: i tossicodipendenti) con  $\mathbb{P}(M)$  più elevata, il test risulta ancora più indicativo.

## Esempio: il test Elisa

Altro dato interessante è il calcolo di  $\mathbb{P}(S|Neg)$

$$\frac{0.9999 \cdot (1 - 0.000025)}{0.9999 \cdot (1 - 0.000025) + (1 - 0.993) \cdot 0.000025} \\ = 1 - 1.7502 \cdot 10^{-7}.$$

Quindi in caso di risultato negativo del test, la probabilità di non avere la malattia è altissima.

Soluzione?

In caso di doppio test (supponendo i due test indipendenti (condizionati a  $M$  e  $S$ ) si ha

$$\mathbb{P}(\{Pos, Neg\} \cup \{Neg, Pos\}) = 2.0032 \cdot 10^{-4} \\ \mathbb{P}(M|Pos, Pos) = 0.9996.$$

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

## Il modello di Hardy-Weinberg



Hardy(1877-1947)



Weinberg(1862-1937)

Un gene in una popolazione diploide ha due alleli,  $A$  e  $a$ , dunque gli individui possono essere  $AA$ ,  $Aa$  oppure  $aa$ . Le frequenze dei tre tipi sono rispettivamente  $p$ ,  $q$  e  $r$  ( $p + q + r = 1$ ).

Possiamo prevedere le frequenze nella prossima generazione? e in quella successiva?

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

# La prima generazione

Ipotizziamo che gli accoppiamenti siano casuali e avvengano con la stessa probabilità che si avrebbe pescando a caso i due genitori (come palline da un'urna). Diamo un nome ad alcuni eventi:

$G_{AA}^1$  = il primo genitore è  $AA$ ;

$G_{Aa}^1$  = il primo genitore è  $Aa$ ;

$G_{aa}^1$  = il primo genitore è  $aa$ ;

$F_A^1$  = il primo genitore fornisce l'allele  $A$ ;

$F_a^1$  = il primo genitore fornisce l'allele  $a$ ;

mentre con  $G_{AA}^2$ ,  $G_{Aa}^2$ ,  $G_{aa}^2$  e  $F_A^2$ ,  $F_a^2$  indicheremo i rispettivi eventi riguardanti il secondo genitore.

## La prima generazione

Per le ipotesi fatte

$$\mathbb{P}(G_{AA}^1) = p; \quad \mathbb{P}(G_{Aa}^1) = q; \quad \mathbb{P}(G_{aa}^1) = r;$$

$$\mathbb{P}(G_{AA}^2) = p; \quad \mathbb{P}(G_{Aa}^2) = q; \quad \mathbb{P}(G_{aa}^2) = r.$$

Inoltre valgono le seguenti probabilità condizionate:

$$\mathbb{P}(F_A^1 | G_{AA}^1) = 1; \quad \mathbb{P}(F_A^1 | G_{Aa}^1) = \frac{1}{2}; \quad \mathbb{P}(F_A^1 | G_{aa}^1) = 0;$$

$$\mathbb{P}(F_A^2 | G_{AA}^2) = 1; \quad \mathbb{P}(F_A^2 | G_{Aa}^2) = \frac{1}{2}; \quad \mathbb{P}(F_A^2 | G_{aa}^2) = 0;$$

Per il teorema delle probabilità totali

$$\mathbb{P}(F_A^1) = p + \frac{1}{2}q = \mathbb{P}(F_a^2).$$

Allo stesso modo si calcola  $\mathbb{P}(F_a^1) = \mathbb{P}(F_A^2) = r + \frac{1}{2}q$ .Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventiEsempio  
CaveatTeorema delle  
probabilità  
totaliFormula di  
BayesEsempio del fumo  
Test clinici  
Test Elisa

Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

## La prima generazione

Sempre per le nostre ipotesi, gli alleli forniti dai due genitori sono indipendenti, quindi la probabilità  $p_1$  che un individuo della prima generazione sia  $AA$  è il prodotto delle probabilità degli eventi  $F_A^1$  e  $F_A^2$ :

$$p_1 = \left(p + \frac{1}{2}q\right)^2.$$

Analogamente la probabilità  $r_1$  che un individuo della prima generazione sia  $aa$  è

$$r_1 = \left(r + \frac{1}{2}q\right)^2.$$

Infine la probabilità  $q_1$  che un individuo della prima generazione sia  $Aa$  è

$$q_1 = 1 - p_1 - r_1 = 2 \left(p + \frac{1}{2}q\right) \left(r + \frac{1}{2}q\right).$$

## La seconda generazione

Per le frequenze nella seconda generazione basta usare le formule appena ottenute per  $p_1, q_1, r_1$  in funzione di  $p, q, r$  e inserirvi  $p_1, q_1, r_1$  al posto di  $p, q, r$ :

$$p_2 = (p_1 + \frac{1}{2}q_1)^2; \quad r_2 = (r_1 + \frac{1}{2}q_1)^2;$$

$$q_2 = 2(p_1 + \frac{1}{2}q_1)(r_1 + \frac{1}{2}q_1).$$

Bastano pochi calcoli per scoprire che

$$p_2 = p_1, \quad q_2 = q_1, \quad r_2 = r_1.$$

Questo significa che dopo la prima generazione le frequenze non cambiano di generazione in generazione e quindi la popolazione raggiunge un *equilibrio genetico*.

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio  
Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo  
Test clinici  
Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg  
Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti



# Commenti sul risultato ottenuto

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

In realtà le cose sono più complesse, infatti esiste la possibilità, ad esempio, che a riprodursi siano solo gli individui  $AA$  e allora dalla generazione successiva non avremmo alcun individuo con allele  $a$ .

Per uno studio maggiormente realistico occorrerebbe la teoria delle catene di Markov, ma per popolazioni molto numerose anche l'approccio di Hardy-Weinberg può andare bene.

# Il problema delle 3 carte

## Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

## Approfondiamo: famiglie indipendenti

Vediamo un esempio in cui occorre fare un po' di attenzione. Abbiamo 3 carte: una ha le due facce di colore rosso, una ha le due facce di colore nero, la terza ha una faccia rossa, l'altra nera.

Scegliamo a caso una carta e la mettiamo sul tavolo: se la faccia che vediamo è rossa, qual è la probabilità che anche la faccia che non vediamo sia rossa?

Chiaramente se non sapessimo nulla la probabilità di scegliere la carta tutta rossa sarebbe  $1/3$ .

Invece sappiamo che una faccia è rossa.

# Il problema delle 3 carte

Fissiamo gli eventi di interesse:

$R_1 = \{\text{la faccia visibile è rossa}\}$

$R_2 = \{\text{la faccia coperta è rossa}\}$

quello che ci chiedono è  $\mathbb{P}(R_2|R_1)$ .

Usiamo la definizione

$$\mathbb{P}(R_2|R_1) = \frac{\mathbb{P}(R_1 \cap R_2)}{\mathbb{P}(R_1)},$$

dove l'evento  $R_1 \cap R_2$  è  $\{\text{la carta è tutta rossa}\}$ . Allora

$\mathbb{P}(R_1 \cap R_2) = 1/3$  perché la carta è scelta con uguale probabilità fra le 3 e  $\mathbb{P}(R_1) = 3/6 = 1/2$  perché la faccia che vediamo è scelta con uguale probabilità fra le 6 possibili, e 3 facce sono rosse.

In conclusione,

$$\mathbb{P}(R_2|R_1) = \frac{1/3}{1/2} = \frac{2}{3}.$$

Probabilità  
condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio  
Caveat

Teorema delle  
probabilità  
totali

Formula di  
Bayes

Esempio del fumo  
Test clinici  
Test Elisa

Altri esempi

Il modello di  
Hardy-Weinberg  
Il problema delle 3  
carte

Approfondiamo:  
famiglie  
indipendenti

## Il problema delle 3 carte

### Probabilità condizionata

Esempio: genotipo e  
fenotipo

Definizione

Proprietà

### Indipendenza di eventi

Esempio

Caveat

### Teorema delle probabilità totali

### Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

### Altri esempi

Il modello di  
Hardy-Weinberg

Il problema delle 3  
carte

### Approfondiamo: famiglie indipendenti

Allora abbiamo concluso che, *sapendo che una delle due facce è rossa*, la probabilità che anche l'altra sia rossa è  $2/3$ .

Esempio di ragionamento errato: siccome una faccia è rossa, i casi sono due. O si tratta della carta tutta rossa, oppure della carta bicolore. Quindi la probabilità che la carta scelta sia quella tutta rossa è  $1/2$ .

Dove sta l'errore?

Le due carte possibili non sono equiprobabili! La carta tutta rossa, *sapendo che una delle due facce è rossa*, ha il doppio di probabilità di quella bicolore.

# Famiglie indipendenti di eventi

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

### DEFINIZIONE DI FAMIGLIA INDIPENDENTE

La famiglia di eventi  $A_1, A_2, \dots, A_n$  si dice indipendente se valgono le uguaglianze

- $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$  per ogni  $i, j$  (tutte le coppie).
- $\mathbb{P}(A_i \cap A_j \cap A_k) = \mathbb{P}(A_i)\mathbb{P}(A_j)\mathbb{P}(A_k)$  per ogni  $i, j, k$  (tutte le terne).
- ...
- $\mathbb{P}(A_1 \cap \dots \cap A_n) = \prod_{i=1}^n \mathbb{P}(A_i)$  (tutti e  $n$ ).

Nota: le uguaglianze di cui sopra vanno verificate tutte, non basta verificare per le coppie.

# Coppie indipendenti, terna non indipendente

## Probabilità condizionata

Esempio: genotipo e fenotipo

Definizione

Proprietà

## Indipendenza di eventi

Esempio

Caveat

## Teorema delle probabilità totali

## Formula di Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di Hardy-Weinberg

Il problema delle 3 carte

## Approfondiamo: famiglie indipendenti

Lancio due dadi.  $A$  = “il primo dado esce pari”,  $B$  = “il secondo dado esce pari”,  $C$  = “la somma è dispari”.

Allora  $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$ ;

$\mathbb{P}(A \cap B) = \mathbb{P}(B \cap C) = \mathbb{P}(A \cap C) = 1/4$  ma

$\mathbb{P}(A \cap B \cap C) = 0$ .

Qui ad esempio sapere che il primo dado esce pari **non cambia** la valutazione della probabilità che la somma sia dispari ( $\Rightarrow$  i due eventi sono indipendenti), ma sapere che il primo dado esce pari **e** il secondo pure **cambia** la valutazione della probabilità che la somma sia dispari ( $\Rightarrow$  i tre eventi non sono indipendenti).

## Altro controesempio

Probabilità  
condizionataEsempio: genotipo e  
fenotipo

Definizione

Proprietà

Indipendenza  
di eventi

Esempio

Caveat

Teorema delle  
probabilità  
totaliFormula di  
Bayes

Esempio del fumo

Test clinici

Test Elisa

## Altri esempi

Il modello di  
Hardy-WeinbergIl problema delle 3  
carteApprofondiamo:  
famiglie  
indipendenti

Estraggo un numero intero tra 1 e 8 con uguale probabilità. Si considerino i tre eventi seguenti:  $A$  = “esce un numero pari”,  $B$  = “esce 1, 3, 7 oppure 8”,  $C$  = “esce 4,5,6 oppure 8”.

Allora  $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$ ;  
 $\mathbb{P}(A \cap B) = 1/8$ ,  $\mathbb{P}(B \cap C) = 1/8$ ,  $\mathbb{P}(A \cap C) = 3/8$   
e  $\mathbb{P}(A \cap B \cap C) = 1/8 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ .

Non basta quindi verificare solo l'ultima uguaglianza!