

Probabilità Appunti

Giovanni Palma e Alex Basta

Contents

Chapter 1

Introduzione _____ Page _____

Chapter 2

Spazi di probabilità _____ Page _____

- 2.1 Concetti introduttivi
- 2.2 Regole del calcolo probabilistico
Assiomi della probabilità — • Conseguenze degli assiomi —
- 2.3 Probabilità discreta
Probabilità uniforme —

Chapter 3

Probabilità Condizionata _____ Page _____

- 3.1 Definizione e motivazioni
- 3.2 Regola della catena
- 3.3 Indipendenza di eventi
Generalizzazione su n eventi — • Esercizi —
- 3.4 Bayes
Formula di Boyes —

Chapter 1

Introduzione

Appunti di Probabilità presi in base alle lezioni di Elly Shlein, qui si è piddini

Chapter 2

Spazi di probabilità

2.1 Concetti introduttivi

Innanzitutto andiamo a definire che cosa intendiamo per *esperimento aleatorio*, *esito*, *probabilità*

Con la dicitura *esperimento aleatorio* indicheremo qualunque fenomeno (fisico, economico, sociale, ...) il cui esito non sia determinabile con certezza a priori. Il nostro obiettivo è di fornire una descrizione matematica di un esperimento aleatorio, definendo un modello probabilistico, un *esito* invece è un ipotetico risultato di un'esperimento aleatorio sulla base di un cosiddetto *spazio campionario* un insieme che contiene tutti gli esiti possibili dell'esperimento

Example 2.1.1

- **Esperimento aleatorio:** Lancio di un dado.
- **Spazio campionario:** $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- **Esito:** 4.

Note:

In casi più complessi ci saranno vari sotto-esperimenti aleatori, come 10 lanci di un dado.

Adesso forniamo vere e proprie definizioni

Definition 2.1.1: evento

Si definisce **evento** un'affermazione riguardante l'ipotetico esito univoco dell'esperimento, di cui si può affermare con certezza se è vero o falso una volta noto l'esito

Example 2.1.2

Esper. aleatorio: Lancio del dado
 $A = \text{"esce un numero pari"}$

Definition 2.1.2: Spazio campionario

Lo **spazio campionario** è l'insieme di tutti i possibili esiti di un esperimento casuale e viene denotato con Ω

Notare che non si afferma "tutti e solo tutti", quindi **qualsiasi** insieme che contiene gli esiti possibili può essere considerato uno spazio campionario

Example 2.1.3 (Lancio dado)

Possiamo porre come spazio campionario:

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

ma anche

$$\Omega = \mathbb{R}$$

Definition 2.1.3: Esiti favorevoli

Esiti per cui un evento è vero sono detti esiti favorevoli.

Definition 2.1.4: Evento in termine di insiemi

Un evento si può definire anche come il sottoinsieme dello spazio campionario Ω formato da tutti gli esiti favorevoli dell'evento.

Example 2.1.4

$\Omega = \{1, 2, 3, 4, 5, 6\} \implies A = \text{"esce un numero pari"} \implies \{2, 4, 6\}$ sono gli esiti favorevoli dell'evento A .

Note:

La definizione insiemistica di un evento dipende dallo spazio campionario Ω definito, poiché l'evento è un sottoinsieme di Ω . Tuttavia, l'insieme degli esiti favorevoli di un evento è fisso, e rappresenta l'insieme evento di cardinalità massima possibile, ovvero l'insieme degli esiti favorevoli $A \subseteq \Omega$.

Definition 2.1.5

- Ω è l'evento *certo*
- \emptyset è l'evento *impossibile*
- $\omega \in \Omega$ è un evento *elementare* ($A = \{\omega\}$)

Example 2.1.5

Lancio un dado.

$A = \text{"esce un numero tra 1 e 6"}$

$B = \text{"esce un numero maggiore di 6"}$

$C = \text{"esce il numero 3"}$

- Se $\Omega = \{1, 2, 3, 4, 5, 6\}$, allora:
 - $A = \Omega$ (evento certo),
 - $B = \emptyset$ (evento impossibile),
 - $C = \{3\}$ (evento con un solo esito favorevole).
- Se $\Omega = \mathbb{R}$, allora:
 - $A = \{1, 2, 3, 4, 5, 6\} \subset \Omega$ (evento quasi certo),
 - $B = (6, +\infty)$ (evento quasi impossibile),
 - $C = \{3\}$ (evento con un solo esito favorevole).

2.2 Regole del calcolo probabilistico

Ad ogni relazione logica possiamo associare un'operazione insiemistica:

Connettivi Logici	Connettivi Insiemistici
$A \vee B$	$A \cup B$
$A \wedge B$	$A \cap B$
$\neg A$	A^c
$A \implies B$	$A \subseteq B$
$A \iff B$	$A = B$

Note:

Nella prima colonna, A e B sono eventi come affermazioni, mentre nella colonna di destra sono degli insiemi.

2.2.1 Assiomi della probabilit 

Poniamo tre assiomi fondamentali da cui possiamo partire per derivare tutte le operazioni e propriet  che ci servono:

Note:

Per noi tutti i sottoinsiemi di Ω sono eventi (anche se non sar  sempre cos )

Definition 2.2.1: Assiomi fondamentali della probabilit 

Assioma 1. A ciascun sottoinsieme (o evento) A di Ω   assegnato un numero $\mathbb{P}(A)$ che verifica:

$$0 \leq \mathbb{P}(A) \leq 1.$$

Tale numero $\mathbb{P}(A)$ si chiama **probabilit ** dell'evento A .

Assioma 2. $\mathbb{P}(\Omega) = 1$.

Assioma 3. Vale la propriet  di **additivit  numerabile**^a: sia $A_1, A_2, \dots, A_n, \dots$ una successione di sottoinsiemi di Ω tra loro disgiunti^b e sia

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Allora

$$\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

^aAnche detta **σ -additivit **.

^bIn formule: $A_i \cap A_j = \emptyset$, per ogni $i \neq j$. In altri termini, non hanno elementi in comune.

Note:

Quindi, per il primo assioma, esiste una funzione probabilit  $\mathbb{P}(A) : \mathcal{P}(\Omega) \rightarrow [0, 1]$.

Definition 2.2.2: Spazio di probabilit 

La coppia (Ω, \mathbb{P}) si dice **spazio di probabilit ** o modello matematico dell'esperimento aleatorio

2.2.2 Conseguenze degli assiomi

Theorem 2.2.1

Sia Ω spazio campionario e \mathbb{P} probabilit  su Ω ((Ω, \mathbb{P})   uno spazio di probabilit  con $\mathbb{P} : \mathcal{P}(A) \rightarrow [0, 1]$). Dagli assiomi 1, 2, 3 deduciamo le cose seguenti:

1. $\mathbb{P}(\emptyset) = 0$

2. **Additività finita:** $(A_i)_{i=1,\dots,n}. \forall i \neq j. A_i \cap A_j = \emptyset \implies \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i)$
3. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
4. **Monotonia:** $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$

Dimostrazione: 1. Devo mostrare che $\mathbb{P}(\emptyset) = 0$. Per semplicità definiamo $p := \mathbb{P}(\emptyset)$. Uso l'assioma 3 con la successione $(A_n)_{n \in \mathbb{N}}$ dove $\forall i \in \mathbb{N}. A_i = \emptyset$, che sono tutti eventi disgiunti. Quindi:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} p.$$

Inoltre:

$$\bigcup_{i=1}^{\infty} A_i = \emptyset \implies p = \sum_{i=1}^{\infty} p.$$

L'equazione è soddisfatta solo per $p = 0$.

2. Supponiamo di avere una sequenza finita disgiunta A_1, \dots, A_n . Definisco $(B_i)_{i \in \mathbb{N}}$ tale che $B_i = A_i$ per $i = 1, \dots, n$ e $B_i = \emptyset$ per $i > n$. Usando l'assioma 3:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) = \sum_{i=1}^n \mathbb{P}(A_i).$$

3. Per definizione di complemento, $A^c \cup A = \Omega$ e $A^c \cap A = \emptyset$. Per additività:

$$\mathbb{P}(A^c) + \mathbb{P}(A) = \mathbb{P}(\Omega) = 1 \quad (\text{per l'assioma 2}).$$

4. Se $A \subseteq B$, allora $B = A \cup (B \setminus A)$, con A e $B \setminus A$ disgiunti. Per additività:

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A).$$

Q

Theorem 2.2.2 Probabilità unione non disgiunta

Siano A e B eventi:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \quad (2.1)$$

Dimostrazione: $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$. Per additività:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B).$$

Osservando che:

$$\mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B) = \mathbb{P}(A), \quad \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) = \mathbb{P}(B),$$

si ottiene la formula.

Q

Note:

La formula si complica con un numero di eventi maggiore di 2. Per $n = 3$:

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(A \cap C) + \mathbb{P}(A \cap B \cap C).$$

2.3 Probabilità discreta

Finora sappiamo solo le "regole" che deve seguire una funzione per essere una probabilità. Passiamo ora a vedere come calcolare il valore di un certo tipo di probabilità, la *probabilità discreta*:

Definition 2.3.1: Probabilità discreta

Chiamo probabilità discreta una funzione probabilità \mathbb{P} su Ω , tale che:

$$\exists \bar{\Omega} \subseteq \Omega, \bar{\Omega} \text{ e' finito o numerabile. } \mathbb{P}(\bar{\Omega}) = 1$$

Ovvero, una probabilità è discreta se il suo spazio campionario minimo è finito o numerabile. Questa condizione è necessaria per poter poi definire un modo per effettivamente calcolare il valore della probabilità (discreta) di un qualunque evento.

Diamo prima una definizione di una tale probabilità:

Definition 2.3.2: Delta di Dirac

Sia $\Omega = \mathbb{R}$, $x_0 \in \mathbb{R}$, allora si chiama delta di Dirac centrato in x_0 la funzione:

$$\delta_{x_0} : \mathcal{P}(\mathbb{R}) \rightarrow [0, 1]$$
$$A \mapsto \delta_{x_0}(A) = \begin{cases} 1 & x_0 \in A \\ 0 & x_0 \notin A \end{cases}$$

Notare che per definizione, la funzione di Dirac è una probabilità discreta, dato che soddisfa tutti gli assiomi per essere una probabilità e il suo spazio campionario minimo è formato da un solo elemento di Ω , quindi è discreta (ma non molto utile dato che può assumere solo due valori). Però, tramite le delta di Dirac, siamo in grado di costruire qualunque altra probabilità discreta:

Sia $\Omega = \mathbb{R}$. Prendiamo un numero contabile n di eventi singoletto $x_1, x_2, \dots, x_n \in \mathbb{R}$ a cui corrispondono $p_1, p_2, \dots, p_n \in \mathbb{R}$ tale che:

$$\forall i = 1, \dots, n. p_i \in [0, 1], \quad \sum_{i=1}^n p_i = 1$$

Definiamo la funzione:

$$\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$$
$$A \mapsto \sum_{i=1}^n p_i \delta_{x_i}(A)$$

\mathbb{P} è una combinazione lineare di delta di Dirac. Essendo una combinazione convessa, $\mathbb{P} \in [0, 1]$ e si può dimostrare che soddisfa gli altri due assiomi (2 e 3), quindi è una probabilità discreta! Variando le x e le p è possibile generare qualsiasi funzione \mathbb{P} discreta.

Example 2.3.1

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $\forall i = 1, \dots, 6. x_i = i, p_i = \frac{1}{6}$, la funzione \mathbb{P} associata è:

$$P(A) = \sum_{i=1}^6 \frac{1}{6} \delta_{x_i}(A)$$

$$A = \{1, 2, 3, 4, 5, 6\} \implies P(A) = 1$$

$$B = (6, +\infty) \implies P(B) = 0$$

$$C = \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \implies P(C) = 1$$

Definition 2.3.3

Si chiama evento quasi certo un evento A tale che $\mathbb{P}(A) = 1$.

Definition 2.3.4

Si chiama evento quasi impossibile un evento A tale che $\mathbb{P}(A) = 0$.

Posso allargare Ω quanto voglio perché tanto fuori dall'insieme minimo che comprende tutti gli eventi possibili le probabilità che aggiungo sono quasi impossibili e quindi hanno probabilità 0 e non cambiano il valore totale della somma.

2.3.1 Probabilità uniforme

TODO: dare un esempio di prob. uniforme e definirla per benino

Chapter 3

Probabilità Condizionata

3.1 Definizione e motivazioni

Supponiamo di sapere che un evento di un'esperimento aleatorio si è avverato. Finora abbiamo visto solo casi in cui gli eventi non si influenzavano (*indipendenti*), ma succede spesso nella realtà che se si sa che un certo evento è avvenuto, allora questo ci dà informazioni aggiuntive che possono cambiare la probabilità di altri eventi di cui ancora non sappiamo gli esiti.

Chiamiamo B l'evento che è avvenuto e A un altro evento di cui vogliamo sapere la probabilità. Prima di avere informazioni su B , la probabilità di A era semplicemente $\mathbb{P}(A)$, ma ora ci poniamo la domanda: "se so che si è verificato B , come cambia $\mathbb{P}(A)$?". Denotiamo questa nuova probabilità con:

$$P(A|B)$$

chiamata la *probabilità condizionata di A dato B* .

Definition 3.1.1: Probabilità Condizionata

Prendo due eventi A, B su uno spazio di probabilità (Ω, \mathbb{P}) . Definisco *probabilità condizionata a B di A* la funzione:

$$\begin{aligned}\mathbb{P}(A|B) : \mathcal{P}(\Omega) &\rightarrow [0, 1] \\ A &\mapsto \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}\end{aligned}$$

È possibile dimostrare che una certa funzione è anch'essa una probabilità (sempre discreta), verifichiamo gli assiomi (fissiamo $B \subseteq \Omega$ con $\mathbb{P}(B) > 0$):

1. $\mathbb{P}(A|B) \in [0, 1], \forall A \subseteq \Omega$
2. $\mathbb{P}(\Omega|B) = 1$
3. σ - addittività: $(A_i)_{i \in \mathbb{N}}$ disgiunti:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i | B\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i | B)$$

Lasciate al lettore in quanto davvero molto facili, quasi banali. Se non riesci a farle fai schifo. Vediamo ora, con un esempio, come mai è proprio questa la definizione utilizzata:

Example 3.1.1

- Lancio del dado: $\Omega = \{1, 2, 3, 4, 5, 6\}$, \mathbb{P} probabilità uniforme:

$\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}, \forall \omega \in \Omega$, ovvero:

$$\mathbb{P}(A) = \frac{\text{casi favorevoli in } A}{\text{casi possibili}}$$

$A = \text{"esce un numero maggiore di 3"} = \{3, 4, 5, 6\}$ e $B = \text{"esce un numero pari"} = \{2, 4, 6\}$, domanda: quanto vale $\mathbb{P}(A|B)$?

$P(A) = \frac{4}{6}$ come abbiamo già visto.

Ora abbiamo un'informazione in più: sappiamo che B si è avverato. Questo significa che si restringe l'insieme di valori che possono essere usciti al lancio del dado. ATTENZIONE! ciò non vuol dire che cambia lo spazio campionario perché l'esperimento è lo stesso, ma cambiano i *veri* casi favorevoli e i *veri* casi possibili:

$$P(A|B) = \frac{\text{"veri casi favorevoli di A"}}{\text{veri casi possibili}} = \frac{|A \cap B|}{|B|} = \frac{2}{3}$$

- Vediamo anche cosa accade quando la probabilità non è uniforme, come con un dado a 4 facce truccato:

$$\Omega = \{1, 2, 3, 4\}, \mathbb{P}(4) = \frac{1}{15}, \mathbb{P}(3) = \frac{2}{15}, \mathbb{P}(2) = \frac{4}{15}, \mathbb{P}(1) = \frac{8}{15}$$

$$A = \{3, 4\}, B = \{2, 4\}$$

$$\mathbb{P}(A|B) = \frac{\text{"probabilità dei veri casi favorevoli di A"}}{\text{probabilità dei veri casi possibili}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Note:

Se \mathbb{P} è la probabilità uniforme allora:

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

Note:

B è fissato nella definizione di probabilità condizionata, ovvero:

$$\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$$

Quindi il ruolo di A e B è completamente diverso

Note:

Se $B = \Omega$, allora $\mathbb{P}(A|B) = \mathbb{P}(A)$ dato che la conoscenza del fatto che si è avverato Ω è ovvio e non ci cambia.
Se $A = \Omega$, allora $\mathbb{P}(\Omega|B) = 1$ (per proprietà, dato che è sempre una probabilità)

3.2 Regola della catena

La probabilità condizionata in genere è nota e si usa per calcolare la probabilità dell'intersezione:

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

Questa formula è detta *regola della catena* e vale in generale con n eventi:

Proposition 3.2.1 Regola della catena (generalizzata)

$(A_i)_{i=1, \dots, n}$, $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$, allora:

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \dots \mathbb{P}(A_n|A_1 \cap \dots \cap A_{n-1})$$

Note:

La condizione funziona grazie alla monotonia, dato che $0 < \mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \leq \mathbb{P}(A_1 \cap \dots \cap A_j)$, $1 \leq j \leq n-1$ quindi siamo certi che l'intersezione degli eventi che sono avvenuti è maggiore di 0.

$$: P(A_1) \frac{P(A_1 \cap A_2)}{P(A_1)} \dots \frac{P(A_1 \cap \dots \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})} = P(A_1 \cap \dots \cap A_n)$$



TODO: migliora un po

Example 3.2.1

Un'urna contiene tre palline bianche, due palline nere e una pallina rossa. Si eseguono tre estrazioni senza reimmissione. Qual è la probabilità di estrarre nell'ordine una bianca, una rossa e una nera? Sono interessato solo ad alcuni eventi, quindi non c'è bisogno di descrivere l'intero esperimento aleatorio. Per prima cosa definisco l'evento:

A = "estrarre in ordine una bianca, una rossa e una nera"

Voglio trovare $P(A)$. Notiamo che dobbiamo determinare tre sottoesperimenti in relazione (dato che non c'è reimmissione). Quindi dopo ogni sottoesperimento cambia la composizione dell'urna, e sappiamo come calcolare la probabilità condizionata:

B_i = "estraggo una pallina bianca all' i -esimo turno"

R_i = "estraggo una pallina rossa all' i -esimo turno"

N_i = "estraggo una pallina nera all' i -esimo turno"

Esistono tre famiglie di eventi: $(B_i)_{i=1,\dots,k}$, $(R_i)_{i=1,\dots,k}$, $(N_i)_{i=1,\dots,k}$ dove i indica il turno al quale viene estratta la pallina. Quindi possiamo scrivere A come relazione fra sottoeventi:

$$A = B_1 \cap R_2 \cap N_3$$

Quindi:

$$P(A) = P(B_1 \cap R_2 \cap N_3) = P(B_1)P(R_2|B_1)P(N_3|B_1 \cap R_2)$$

Solo ora possiamo passare ai valori numerici. Dato che gli esiti sono equiprobabili e lo spazio campionario è finito, la probabilità è uniforme:

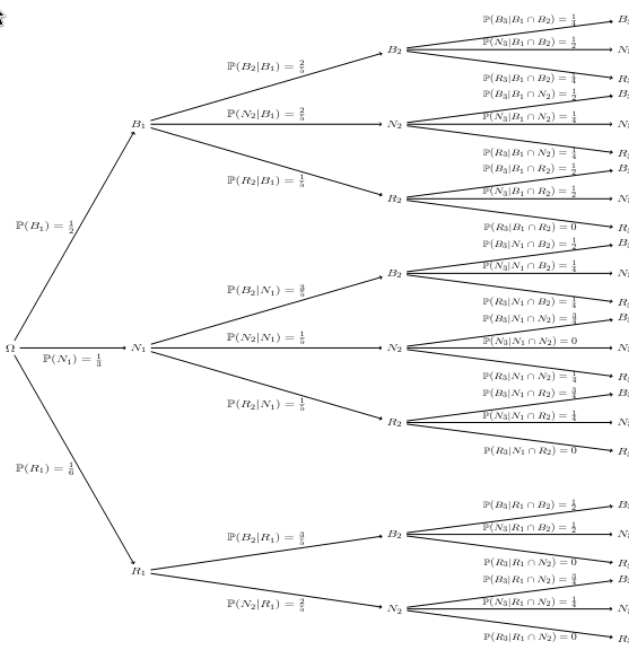
$$P(B_1) = \frac{1}{2}, P(R_2) = \frac{1}{5}, P(N_3) = \frac{1}{2}$$

$$P(A) = \frac{1}{20}$$

Per esperimenti formati da sottoesperimenti di cui conosco le probabilità condizionate, è possibile rappresentare ogni evento come un nodo:

Ω = primo nodo

e ogni probabilità come un ramo che partiziona il nodo (tanti rami quanti gli insiemi della partizione) che rappresenta poi un altro evento (condizionato dalla seconda in poi).



La regola della catena la leggo sul diagramma ad albero:

percorso: $\Omega \rightarrow B_1 \rightarrow R_2 \rightarrow N_3$ ha probabilita' $P(B_1 \cap R_2 \cap N_3)$ che si calcola facendo il prodotto delle probabilita' dei relativi rami che si usano nel percorso.

E' uno strumento utile per convincerci che stiamo usando le formule giuste, ma non le sostituisce e puo' diventare laborioso per problemi complessi.

Example 3.2.2

Ci sono due urne: la prima contiene due palline rosse e una bianca; la seconda contiene tre palline rosse e due bianche. Si lancia una moneta: se esce testa si estrae una pallina dalla prima urna, se esce croce si estrae una pallina dalla seconda urna. Qual e' la probabilita' che l'esito del lancio della moneta sia testa e la pallina estratta sia bianca?

2 sottoesperimenti:

- lancio della moneta
- estrazione da un'urna

Nota che i sottoesperimenti sono indipendenti dall'esito di altri esperimenti. Sono gli esiti, ovvero i risultati, che possono dipendere dagli esiti di altri esperimenti.

A = "esce testa ed estraggo una pallina bianca"

Devo esprimere A con eventi che

T = "esce testa"

U = estraggo una pallina bianca

Disegna lo zio pera di diagramma che non mi metto a fare, se @GiovanniPalma vuole puo' farlo

$$A = T \cap U, \quad P(A) = P(T)P(U|T)$$

3.3 Indipendenza di eventi

E' possibile che sapere che un evento B e' avvenuto non altera la probabilita' di un altro evento A . Possiamo esprimere questa relazione in modo matematico cosi':

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

Utilizzando la definizione di probabilit  condizionale, possiamo usare un'identit  equivalente che useremo come definizione:

Definition 3.3.1: Eventi indipendenti

Due eventi A, B si dicono indipendenti se:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \quad (3.1)$$

E viene denotato $A \perp B$

Usiamo questa definizione dato che   esplicitamente simmetrica, ovvero se A   indipendente a B allora vale anche il contrario:

$$A \perp B \iff B \perp A$$

ed   definita (e banalmente vera) anche quando $\mathbb{P}(A) = 0$ o $\mathbb{P}(B) = 0$. In particolare si noti il seguente teorema:

Theorem 3.3.1 Teorema della simmetria tra eventi indipendenti

Sia $\mathbb{P}(B) > 0$ allora:

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

Dall'altro lato, sia $\mathbb{P}(A) > 0$ allora:

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

Dimostrazione: Verr  fornita solo la dimostrazione del primo punto, la seconda parte   analoga. Assumo $\mathbb{P}(B) > 0$, si ha:

- $A \perp B \implies \mathbb{P}(A|B) = \mathbb{P}(A)$:

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

- $\mathbb{P}(A|B) = \mathbb{P}(A) \implies A \perp B$:

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



Note:

Si noti che se $\mathbb{P}(A) > 0$ e $\mathbb{P}(B) > 0$ allora, le tre uguaglianze seguenti sono equivalenti:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \iff \mathbb{P}(A|B) = \mathbb{P}(A) \iff \mathbb{P}(B|A) = \mathbb{P}(B)$$

Note:

- L'indipendenza   diversa dalla disgiunzione:

$$A \perp B \neq A \cap B = \emptyset$$

infatti sono relazioni ortogonali:

$$A \perp B \wedge A \cap B = \emptyset \iff \mathbb{P}(A)\mathbb{P}(B) = 0 \iff \mathbb{P}(A) = 0 \vee \mathbb{P}(B) = 0$$

- L'indipendenza   diverso dall'essere sottoinsieme non-vuoto:

$$A \perp B \neq A \subseteq B \vee B \subseteq A$$

infatti:

$$A \perp B \wedge A \subseteq B \iff \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A) \iff \mathbb{P}(B) = 1$$

Quindi in generale due eventi sono indipendenti quando la loro intersezione ha "le giuste proporzioni".

Adesso fornirò un altro teorema piuttosto importante:

Proposition 3.3.1 Sull'indipendenza di eventi complementari

Siano A, B due eventi indipendenti, allora:

$$A \perp B \iff A^c \perp B, A \perp B^c, A^c \perp B^c$$

Dimostrazione: Dimostro solo la prima parte, le altre sono analoghe.
Assumo A, B due eventi indipendenti, debbo dimostrare la seguente uguaglianza:

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(A^c) \cdot \mathbb{P}(B)$$

Dato che

$$B = \Omega \cap B = (A \cup A^c) \cap B = (A \cap B) \cup (A^c \cap B)$$

E dato che $(A \cap B)$ e $(A^c \cap B)$ sono disgiunti, per 2 (*additività finita*) si ha:

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)$$

E dato che $A \perp B$ si ha:

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) = \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) + \mathbb{P}(A^c \cap B)$$

Quindi:

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(B) \cdot (1 - \mathbb{P}(A)) = \mathbb{P}(A^c) \cdot \mathbb{P}(B)$$



3.3.1 Generalizzazione su n eventi

Come nel caso con solo due eventi, $n > 2$ eventi si dicono indipendenti quando, sapendo che qualsiasi numero degli altri eventi si è avverato, la probabilità dell'evento non cambia. Questo deve valere per tutti gli n eventi, ovvero:

$$\mathbb{P}\left(A_i \left| \bigcup_{j=1, j \neq i}^n A_j\right.\right) = \mathbb{P}(A_i) \quad \forall i = 1, \dots, n$$

Si può dimostrare, usando la definizione di probabilità condizionata, che questa identità equivale a dire:

Definition 3.3.2: Eventi indipendenti (per n eventi)

Sia $(A_i)_{i \in I}$ una famiglia di eventi in uno spazio di probabilità. Si dice che questi eventi sono indipendenti quando **per ogni sottoinsieme** finito $J \subseteq I$, $|J| > 2$:

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$$

3.3.2 Esercizi

Example 3.3.1 (Calcolo di eventi indipendenti con probabilità condizionata)

TESTO:

Si lancia un dado a 6 facce

A = "Esce un numero > 4 "

B = "Esce un numero pari"

Determinare $P(A)$ e $P(A|B)$

DETTAGLIO SVOLGIMENTO:

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

$$A = \{5, 6\}$$

$$B = \{2, 4, 6\}$$

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

Dato che $P(A) = P(A|B)$, per il teorema 3.3 si ha che $A \perp B$

Ecco un altro esercizio:

Example 3.3.2

TESTO

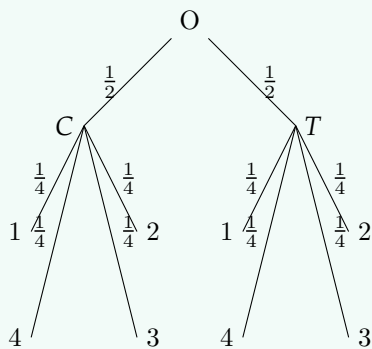
Lanciamo una moneta e un dado a 4 facce.

Determinare uno spazio di prob. che descriva l'esperimento aleatorio

SOLUZIONE

$$\Omega = \{(T, 1), (T, 2), (T, 3), (T, 4), (C, 1), (C, 2), (C, 3), (C, 4)\}$$

$$P(T, 1) = \frac{1}{8} = P(C, 4) = \frac{1}{8}$$



T = "esito del lancio moneta e testa"

C = "esito del lancio moneta è croce"

D_i = "è uscito il numero i "

$$P(C) = \frac{1}{2}, \quad P(T) = \frac{1}{2}$$

$$P(D_i|C) = \frac{1}{4}$$

$$P(D_i|T) = \frac{1}{4}$$

A = "è uscito testa e il numero i "

$$P(A) = P(T) \cdot P(D_i|T) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

Analogamente per $C \cap D_i$

3.4 Bayes

Ci sono delle situazioni nelle quali può esserci richiesto di calcolare la probabilità condizionata "inversa" conoscendo quella diretta. Possiamo fare ciò usando la definizione di probabilità condizionate

Theorem 3.4.1

Siano A, B due eventi t.c. $P(A), P(B) > 0$:

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

Example 3.4.1 (4.1)

Ci sono due urne: la prima urna contiene una pallina bianca e due palline rosse, mentre la seconda contiene due palline bianche e cinque palline rosse. Si lancia una moneta: se esce testa si estrae una pallina dalla prima urna, se esce croce si estrae una pallina dalla seconda. Sapendo che 'e stata estratta una pallina bianca, calcolare la probabilit'a che l'esito del lancio della moneta sia stato testa.

EVENTI

T = esce testa

B = estrazione pallina bianca

$\Omega = \{t, c\} \times \{b, r\}$

$$P(T|B) = \frac{P(B|T)P(T)}{P(B)} = \frac{P(B|T)P(T)}{P(B|T)P(T) + P(B|T^c)P(T^c)} = \frac{1/3 \cdot 1/2}{1/3 \cdot 1/2 + 2/7 \cdot 1/2} \text{ probabilita' totali e bayes}$$

Example 3.4.2

Esempio 5.1. In un'urna ci sono due palline che possono essere rosse (R) o bianche (B). La composizione esatta non 'e nota, quindi le composizioni possibili sono: RR, RB, BB. (H_0, H_1, H_2) Supponiamo che, in base alle informazioni a disposizione, sia ragionevole assegnare uguale probabilit'a pari a $1/3$ alle tre composizioni (ipotesi) possibili, che denotiamo H_0, H_1 e H_2 . 1) Se si estrae una pallina dall'urna, qual 'e la probabilit'a che sia bianca? 2) Si effettuano tre estrazioni con reimmissione: sapendo che le prime due palline estratte sono bianche, qual 'e la probabilit'a che anche la terza pallina estratta sia bianca?

Il paradosso viene dal fatto che dobbiamo tenere traccia dell'asimmetria iniziale e denotare le diverse composizioni come ipotesi possibili.

EVENTI

B_i = estraggo una pallina bianca all'i-esima estrazione

1) $P(B_1) = P(B_1|H_0)P(H_0) + P(B_1|H_1)P(H_1) + P(B_1|H_2)P(H_2) = 1/2$ probabilita' totali e prob. uniforme delle H

Ce l'aspettiamo dato che c'e' completa simmetria fra i rami delle H con la distribuzione di palline

2) $P(B_3|B_1 \cap B_2) = P(B_3 \cap B_2 \cap B_1)/P(B_1 \cap B_2) =$ usare ancora il partizionamento per H per applicare probabilita' totali

3.4.1 Formula di Bayes

Quando si deve calcolare una probabilita' condizionata di eventi "nell'ordine temporale sbagliato", ad esempio se si calcolare $P(A|B)$ ma si conosce solo $P(B|A)$, si puo' utilizzare uno strumento utilissimo facilmente derivabile dalla matematica della probabilita' condizionata, ovvero la formula di Bayes:

Theorem 3.4.2 Formula di Bayes

Siano A e B due eventi tali che $P(A) > 0$ e $P(B) > 0$, allora:

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

Dimostrazione: Per definizione di $P(A|B)$ si ha:

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

Utilizzando la regola della catena possiamo rimuovere il numeratore come segue:

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

Quindi:

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



Example 3.4.3

Ci sono due urne: la prima urna contiene una pallina bianca e due palline rosse, mentre la seconda contiene due palline bianche e cinque palline rosse. Si lancia una moneta: se esce testa si estrae una pallina dalla prima urna, se esce croce si estrae una pallina dalla seconda. Sapendo che è stata estratta una pallina bianca, calcolare la probabilità che l'esito del lancio della moneta sia stato testa.

Svolgimento:

T = "esce testa"

Chapter 4

Calcolo Combinatorio

Quando abbiamo numero finito di elementi, e' possibile contare gli elementi dell'evento per ridurre il problema di probabilita' a un problema di conteggio -> dobbiamo imparare a contare
Dato un certo insieme, dobbiamo calcolarne la cardinalita' utilizzando i giusti strumenti matematici

Theorem 4.0.1 Calcolo probabilita' uniforme (Laplace)

Se Ω e' finito, $\Omega = \{w_1, \dots, w_N\}$ e $\forall i = 1, \dots, N. \mathbb{P}(\{w_i\}) = \frac{1}{N}$ (probabilita' uniforme), allora $\forall A \subseteq \Omega$ (evento):

$$\mathbb{P}(A) = \frac{|A|}{N}$$

Dobbiamo introdurre:

- **Fattoriali:**

- $0! := 1$
- $(n+1)! := (n+1) \cdot n!$

- **Coefficienti Binomiali:**

$$(n, k) = \frac{n!}{k!(n-k)!} \quad \forall n, k \in \mathbb{N}. n \geq k$$

In generale:

$$(n, k) = (n, n-k)$$

Dal triangolo di Tartaglia, possiamo visualizzare altre proprieta' (oltre alla simmetria)

Vedremo che (n, k) sono il numero di modi diversi che abbiamo per selezionare k sottoinsiemi diversi da un insieme di cardinalita' n .

Theorem 4.0.2 Formula di Newton

$$(a+b)^n = \sum_{k=0}^n (n, k) a^k b^{n-k}$$

Quindi anche il coefficiente deriva da un problema di conteggio.

4.1 Metodo (Principio) delle Scelte Successive

E' un algoritmo per determinare la cardinalita' di un insieme. Vediamo un esempio:

Example 4.1.1

Alfabeto di 36 caratteri dove ognuno dei numeri corrisponde a un carattere alfanumerico.

Domanda: "Quante possibili password distinte di 8 caratteri esistono in questo alfabeto?"

Ω = Alfanumerici $\times \dots \times$ Alfanumerici 8 volte.

• Scelte:

1. un carattere dei 36 totali

⋮

2. (e' cosi 8 volte)

Quindi $|\Omega| = 36 \times \dots \times 36 = 36^8$

E se vogliamo evitare di ripetere i caratteri? Vediamo le scelte:

1. un carattere dei 36 totali

2. un carattere dei **35 possibili**

⋮

3.

4. un carattere dei 29 possibili

Quindi $|\Omega| = 36 \times 35 \times \dots \times 29 = \frac{36!}{28!}$

Definiamo il principio generale:

Theorem 4.1.1 Non proprio un teorema

Ciascun elemetno di un insieme A puo' essere determinato tramite sola sequenza di k scelte, dove per ogni scelta ci sono n_1, \dots, n_k possibilita', allora:

$$|A| = n_1 \times \dots \times n_k$$

Note:

Puo' essere riscritto come teorema formale ma **E' TROPPO DIFFICILE PER NOI INFORMATICI** quindi non lo facciamo!!!!

Example 4.1.2

52 carte (13 tipi 4 semi)

1. A = iniseme dei full (un tris e una coppia), $|A| = ?$

4 scelte:

- tipo di carta nel tris (13)
- tipo di carta nella coppia (12)
- semi nel tris (4)
- semi nella coppia (6)

$|A| = 131246 =$ casi favorevoli

2. A = doppie coppie (due coppie di tipo diverso e una carta libera), $|A| = ?$

Scelte:

- tipo nella prima coppia (13)

- tipo nella seconda coppia (12)
- semi nella prima coppia (6)
- semi seconda (6)
- tipo singolo (11)
- seme singolo (4)

$|A| = 131266114$ **SBAGLIATO!!**

Perche' non ci interessa dell'ordine delle due coppie (bisogna dedurlo dalla definizione di A), quindi NON c'e' una prima e seconda coppia (anche sopra era sbagliato vederlo cosi') dato che non c'e' l'ordine.

Combinazioni dei tipi che compongono 2 coppie ($13 \times 12/2$)

4.2 Disposizioni e Combinazioni

Dato un insieme E con n elementi, indichiamo con $DR_{n,k}$ le sequenze ordinate di k elementi di E . Sostanzialmente, $DR_{n,k} = E \times E \dots \times E$ k volte, ovvero:

$$DR_{n,k} = E^k$$

Quindi usando il principio delle scelte successive:

$$|DR_{n,k}| = n^k$$

dato che per ogni E abbiamo una scelta fra n elementi.

Note:

Indicando tale insieme con $DR_{n,k}$, l'insieme E sparisce, dato che ci interessa solo la **cardinalita'** di tale insieme e non ce ne frega un cavolo dei suoi elementi.

Example 4.2.1 (Iniziamo a calcolare le probabilita'!)

Presa un'urna con n palline numerate ($E = \{1, \dots, n\}$) e si estraggono k palline con *reimbussolamento*.

$$\Omega = E \times \dots \times E = DR_{n,k}$$

Quindi $|\Omega| = n^k$, e la probabilita' uniforme e' data da:

$$\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|} = n^{-k}$$

Example 4.2.2

Determinare spazi campionari per i seguenti esperimenti:

- Scelta casuale di una parola di 8 lettere
- Scelta di colonne del totocalcio (risultato di 13 partite)

Quindi lo usiamo nei casi di estrazione con *reimbussolamento* quando ci interessa l'ordine.

Dato un insieme E di n elementi, l'insieme delle disposizioni (senza ripetizione) di k elementi dell'insieme E e' l'insieme delle sequenze ordinate di k elementi *distinti*, ovvero:

$$D_{n,k} := \{(x_1, \dots, x_k) | x_1, \dots, x_k \in E \wedge x_i \neq x_j \text{ se } i \neq j\}$$

Note:

$D_{n,k}$ e' un sottoinsieme **stretto** di $DR_{n,k}$.

Usando le scelte successive:

$$|D_{n,k}| = n \times (n-1) \times \dots \times (n-k+1) = \frac{n!}{(n-k)!}$$

L'esperimento aleatorio di riferimento e' l'estrazione **senza** reimmisione.

Definition 4.2.1: Permutazioni

$P_n = D_{n,n}$, quindi: $|P_n| = n!$