http://www.mat.uniroma2.it/~macci/didattica-corsoST.htm

CAPITOLO 2: INTRODUZIONE ALLA PROBABILITÀ

Un fenomeno è detto "fenomeno aleatorio" se il suo esito è incerto. L'insieme dei possibili esiti (in genere "numerico", eventualmente multidimensionale) viene indicato con Ω

FENOMENO ALEATORIO
- $\rightarrow$ DISCRETO se Ω è finito o numerabile
- $\rightarrow$ CONTINUO se Ω è più che numerabile

Esempi discreti: Ω finito (possibili risultati del lancio di un dado) $\boxed{\Omega=\{1,\dots,6\}}$

Ω numerabile (numero di telefonate ricevute da un centralino in un intervallo di tempo) $\boxed{\Omega=\{0,1,2,\dots\}}$

Esempio continuo: tempo di funzionamento di una lampadina rispetto ad una certa unità di misura $\boxed{\Omega=(0,\infty)}$

In genere si introduce una "famiglia di eventi" $\mathcal{A}$ che viene individuata da una famiglia di sottoinsiemi di Ω.

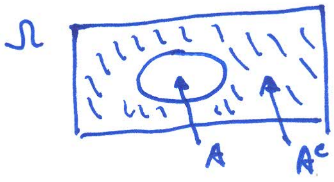
Tornando agli esempi: "esce un numero pari" $\longleftrightarrow \{2,4,6\} \subset \{1,2,3,4,5,6\}$

"il centralino riceve al più 2 telefonate" $\longleftrightarrow \{0,1,2\} \subset \{0,1,2,\dots\}$

"la lampadina si fulmina dopo il tempo $t=5$" $\longleftrightarrow (5,\infty) \subset (0,\infty)$.

Allora viene naturale considerare la corrispondenza tra "operazioni logiche tra eventi" e "operazioni insiemistiche":

SOMMA LOGICA	$A \vee B$	$\longleftrightarrow$	UNIONE	$A \cup B$
PRODOTTO LOGICO	$A \wedge B$	$\longleftrightarrow$	INTERSEZIONE	$A \cap B$
NEGAZIONE	$\overline{A}$	$\longleftrightarrow$	COMPLEMENTAZIONE	$A^c = \Omega \setminus A$.



Si vuole fare riferimento a "famiglie di eventi con buone proprietà".
Si intende che, facendo operazioni insiemistiche su elementi di $\mathcal{A}$, si ottiene ancora un elemento di $\mathcal{A}$.

DEFINIZIONE (σ-algebra)

Sia Ω un insieme non vuoto e sia $\mathcal{A} \subset \mathcal{P}(\Omega)$. insieme delle parti di Ω

Allora $\mathcal{A}$ è una <u>σ-algebra (di eventi)</u> se:

i) $\Omega \in \mathcal{A}$

ii) $\forall A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

iii) $\forall \{A_n\}_{n\geq 1} \subset \mathcal{A} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{A}$

<u>OSS</u>. Si vede facilmente che anche $\emptyset \in \mathcal{A}$ e che $\bigcap_{n\geq 1} A_n \in \mathcal{A}$ nella iii).

<u>OSS</u>. La richiesta della numerabilità viene fatta per semplificare alcune cose successivamente. È una questione che non tratteremo in questo corso.

<u>OSS</u>. Si può prendere $\mathcal{A} = \mathcal{P}(\Omega)$. Questa scelta dà altri problemi se Ω è più che numerabile (è una questione che non trattiamo).
Per questo motivo il caso con Ω discreto viene trattato più diffusamente in questo corso.

DEFINIZIONE (misura di probabilità).

Sia Ω un insieme non vuoto e $\mathcal{A}$ una σ-algebra di eventi.
Allora una funzione $P: \mathcal{A} \to [0,\infty)$ è una <u>misura di probabilità</u> se

1) $P(\Omega) = 1$

2) $\forall \{A_n\}_{n\geq 1} \subset \mathcal{A}$ tale che $\left[\begin{matrix} A_m \cap A_n = \emptyset \\ \text{per } m \neq n \end{matrix}\right]$ si ha

$$P\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} P(A_n)$$

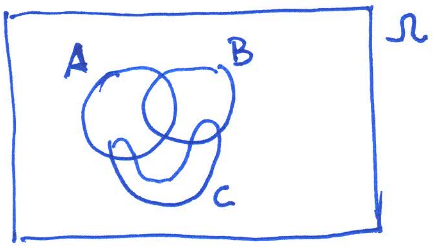
DISGIUNTI A DUE A DUE

<u>TERMINOLOGIA</u>: La terna $(\Omega, \mathcal{A}, P)$ è detta SPAZIO DI PROBABILITÀ.

COMMENTI

→ La misura di probabilità $P : \mathcal{A} \to [0, \infty)$ in realtà assume valori in $[0,1]$; questo sarà chiaro successivamente.

→ La richiesta $A_m \cap A_n = \emptyset$ per $m \neq n$, cioè "insiemi disgiunti a due a due", è più forte della condizione $\bigcap_{n \geq 1} A_n = \emptyset$. Per fissare le idee consideriamo una famiglia di 3 insiemi A, B, C; allora in questo caso



$$A \cap B \neq \emptyset \qquad A \cap B \cap C = \emptyset.$$
$$A \cap C \neq \emptyset$$
$$B \cap C \neq \emptyset$$

CONSEGUENZE DELLA DEFINIZIONE DI MISURA DI PROBABILITÀ

1) $P(\emptyset) = 0$

Infatti, se consideriamo $A_n = \emptyset$ $\forall n \geq 1$, si ha $A_m \cap A_n = \emptyset$ per $m \neq n$ (anche per $m = n$) da cui segue

$$\begin{cases} P\left(\bigcup_{n\geq1} A_n\right) = P\left(\bigcup_{n\geq1} \emptyset\right) = P(\emptyset) \\ \sum_{n\geq1} P(A_n) = \sum_{n\geq1} P(\emptyset) \end{cases} \Rightarrow P(\emptyset) = \sum_{n\geq1} P(\emptyset)$$

Questa condizione non può essere vera se $P(\emptyset) > 0$; infatti il 2° membro sarebbe infinito (e il 1° membro è finito). Quindi si deve avere $P(\emptyset) = 0$ e l'uguaglianza vale come $0 = 0$.

~~2) Sia $m \geq 1$ (finito) e $B_1, \ldots, B_m \in \mathcal{A}$ con $B_m \cap B_n = \emptyset$ per $m \neq n$. Allora $P\left(\bigcup_{n=1}^{m} B_n\right) = \sum_{n=1}^{m} P(B_n)$. Infatti basta~~

2) Sia $h \geq 1$ intero e $B_1, \dots, B_h \in \mathcal{A}$ con $B_m \cap B_n = \emptyset$ per $m \neq n$ $(m, n \in \{1, \dots, h\})$.

Allora $P\left(\bigcup_{n=1}^{h} B_n\right) = \sum_{n=1}^{h} P(B_n)$.

Infatti basta fare riferimento alla condizione ii) nella definizione con

$$A_1 = B_1,\ A_2 = B_2, \dots, A_h = B_h,\ A_{h+1} = A_{h+2} = A_{h+3} = \dots = \emptyset.$$

Infatti, con queste scelte (e per le ipotesi), si ha $A_m \cap A_n = \emptyset$ per $m \neq n$ e si ha

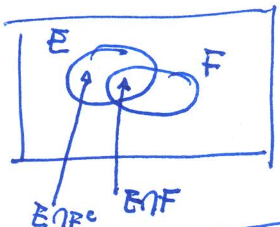
$$P\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} P(A_n) \quad \text{da cui segue:}$$

$$\begin{cases} \bigcup_{n\geq 1} A_n = B_1 \cup \dots \cup B_h \cup \emptyset \cup \emptyset \cup \emptyset \dots = \bigcup_{n=1}^{h} B_n \Rightarrow P\left(\bigcup_{n\geq 1} A_n\right) = P\left(\bigcup_{n=1}^{h} B_n\right) \\ \sum_{n \geq 1} P(A_n) = P(B_1) + \dots + P(B_h) + \underbrace{P(\emptyset)}_{=0} + \dots + \underbrace{P(\emptyset)}_{=0} + \dots = \sum_{n=1}^{h} P(B_n) \end{cases}$$

e quindi si ottiene $P\left(\bigcup_{n=1}^{h} B_n\right) = \sum_{n=1}^{h} P(B_n)$.

3) Specializziamo l'uguaglianza appena verificata prendendo $E, F \in \mathcal{A}$, $h = 2$, $B_1 = E \cap F$, $B_2 = E \cap F^c$. Allora

$$\underbrace{P((E \cap F) \cup (E \cap F^c))}_{= P(E)} = P(E \cap F) + P(E \cap F^c)$$



$(E \cap F) \cup (E \cap F^c) = E$

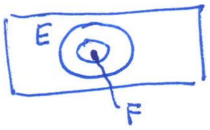
$$\Rightarrow \boxed{P(E) = P(E \cap F) + P(E \cap F^c)}$$

3.1) (con $E = \Omega$)

$$\boxed{\begin{array}{l} 1 = P(F) + P(F^c) \quad \forall F \in \mathcal{A} \\ \text{e quindi } \begin{cases} P(F) = 1 - P(F^c) \\ P(F^c) = 1 - P(F) \end{cases} \end{array}}$$

3.2) (con $F \subset E$) $\to E \cap F = F$

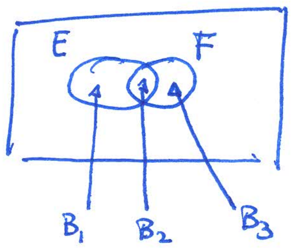
$$\Rightarrow P(E) = P(F) + \underbrace{P(E \cap F^c)}_{\geq 0} \geq P(F)$$



$$\Rightarrow \boxed{P(E) \geq P(F)}$$

Da questo segue che $P(A) \leq P(\Omega) = 1 \quad \forall A \in \mathcal{A}$

3.3) Specializziamo la formula al punto 2) con $E, F \in \mathcal{A}$,
$h=3$, $B_1 = E \cap F^c$, $B_2 = E \cap F$, $B_3 = F \cap E^c$



Ovviamente $B_1 \cup B_2 \cup B_3 = E \cup F$, e quindi

$$\begin{aligned}
P(E\cup F) &= P(B_1)+P(B_2)+P(B_3) \\
&= P(E\cap F^c)+P(E\cap F)+P(F\cap E^c) \\
&= \underbrace{P(E\cap F^c)+P(E\cap F)}_{=P(E)}+\underbrace{P(F\cap E^c)+P(E\cap F)}_{=P(F)}-P(E\cap F)
\end{aligned}$$

si somma e si sottrae la stessa quantità

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

OSSERVAZIONE: L'uguaglianza ottenuta ha una interpretazione naturale. Prendendo $P(E)+P(F)$, si ha che $P(E\cap F)$ viene "presa due volte" e quindi dobbiamo sottrarre questa quantità. Se $E\cap F=\emptyset$ ritroviamo la 2) con $h=2$, $B_1=E$ e $B_2=F$.

La formula ottenuta nella 3.3) si estende al caso di più di due eventi (IDENTITÀ DI BONFERRONI). Qui richiamo quella per $n=3$: $\forall E,F,G\in\mathcal{A}$

$$P(E\cup F\cup G) = P(E)+P(F)+P(G)-P(E\cap F)-P(E\cap G)-P(F\cap G)+P(E\cap F\cap G).$$

COMMENTO GENERALE

Le proprietà della misura di probabilità sono svincolate dalla costruzione del modello, e quindi da come si definisce la misura di probabilità in questione per descrivere il fenomeno aleatorio (tale definizione può dipendere dallo stato di conoscenza dell'osservatore).

SPAZIO DI PROBABILITÀ UNIFORME DISCRETO.

Questa terminologia si usa nel caso in cui si ha la seguente situazione:

- Ω insieme finito (per fissare le idee con n elementi, $n \geq 1$).
- $\mathcal{A} = \mathcal{P}(\Omega)$
- $\forall A \in \mathcal{A}$ $\qquad P(A) = \dfrac{\#A}{\#\Omega} = \dfrac{\#A}{n}$. (dove $\#A$ è la cardinalità di A)

Questa situazione viene fuori imponendo la seguente condizione:

$\forall \omega \in \Omega \qquad P(\{\omega\})$ assume sempre lo stesso valore.

Infatti, se questo valore lo indichiamo con p, allora ~~abbiamo~~ abbiamo le seguenti uguaglianze:

$$1 = P(\Omega) = P\Big(\bigcup_{\omega\in\Omega}\{\omega\}\Big) = \sum_{\omega\in\Omega} \underbrace{P(\{\omega\})}_{=p} = \sum_{\omega\in\Omega} p = np$$

$\boxed{\text{cioè se } \Omega = \{\omega_1, \dots, \omega_n\} \text{ allora } \Omega = \{\omega_1\} \cup \{\omega_2\} \cup \dots \cup \{\omega_n\}}$

$$\Longrightarrow p = \frac{1}{n}.$$

Allora, per ogni $A \in \mathcal{A}$, si ha

$$P(A) = P\Big(\bigcup_{\omega\in A}\{\omega\}\Big) = \sum_{\omega\in A} \underbrace{P(\{\omega\})}_{=p} = p \cdot \#A = \frac{\#A}{n}.$$

COMMENTI.

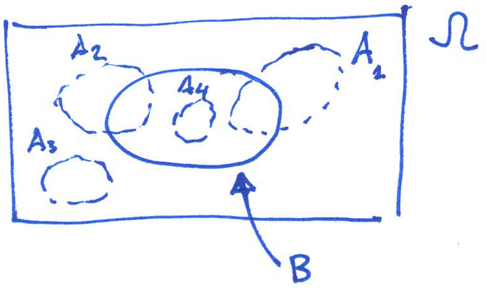
1) In questo caso $P(A) = 0$ se e solo se $A = \emptyset$.
2) Questa situazione esce fuori quando si compiono "estrazioni a caso da un insieme di n oggetti"
3) Questa costruzione non può essere fatta ~~...~~ nel caso in cui Ω è infinito numerabile perché si avrebbe infinito a denominatore e quindi si avrebbe sempre $P(A) = 0$. In altri termini non si riesce a modellizzare il caso di estrazioni a caso da un insieme infinito numerabile di oggetti.
4) Questo modello si può usare nel caso del lancio di un dado equo con $n = 6$ e $\Omega = \{1, 2, 3, 4, 5, 6\}$.

DEFINIZIONE

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Siano $A, B \in \mathcal{A}$ con $P(B) \neq 0$. Allora si definisce "probabilità condizionata di A dato B" (oppure sapendo che si è verificato l'evento B) la seguente quantità:

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

MOTIVAZIONE



Nel voler definire $P(A|B)$ è naturale considerare una quantità che dipende da $P(A \cap B)$ proporzionalmente, con una costante di proporzionalità che non dipende da A (e che dipenderà da B):

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

Inoltre si vuole fare in modo che $(\Omega, \mathcal{A}, P(\cdot|B))$ sia uno spazio di probabilità. Quindi per $A = \Omega$ si ha

$$\underbrace{P(\Omega|B)}_{=1} = c_B \underbrace{P(\Omega \cap B)}_{=P(B)} \implies c_B = \frac{1}{P(B)} \implies P(A|B) = \frac{P(A \cap B)}{P(B)}$$

COMMENTO

La costruzione fatta nella "motivazione" consente di trovare un valore per c_B solo se $P(B) \neq 0$; infatti, se fosse $P(B) = 0$, si avrebbe $1 = c_B \cdot 0$.

Quindi si riesce a dare significato alla probabilità condizionata per $P(B) \neq 0$ ma questa restrizione non è grave.

~~COMMENTO Se P(B)=1, si ha~~

COMMENTO

Supponiamo che $P(B)=1$. Allora $P(A|B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{P(A\cap B)}{1}=P(A\cap B)$.

Inoltre $P(A)=P(A\cap B)+P(A\cap B^c)$ dove $0\overset{\text{ovvio}}{\leq} P(A\cap B^c)\overset{\boxed{A\cap B^c\subset B^c}}{\leq} P(B^c)=0$;

quindi $P(A)=P(A\cap B)$, e sostituendo nell'uguaglianza precedente si ha

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

Questa conclusione si può spiegare come segue. Il verificarsi di un evento di probabilità 1 è una informazione "banale" e quindi la prob. condizionata coincide con quella che si ha senza il condizionamento.

COMMENTO (PROB. CONDIZIONATA PER SPAZIO DI PROBABILITÀ UNIFORME DISCRETO).

Sia $B\in\mathcal{A}$ t.c. $P(B)\neq 0$; quindi $B\neq\emptyset$. Allora

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{\#(A\cap B)/\cancel{n}}{\#B/\cancel{n}}=\frac{\#(A\cap B)}{\#B}\quad\forall A\in\mathcal{A}.$$

ESEMPIO

Un'urna ha 10 palline numerate da 1 a 10.
Si estrae una pallina a caso.

1) Calcolare la probabilità di estrarre un numero maggiore di 5 sapendo che è stato estratto un numero pari
2) Calcolare la stessa probabilità nel caso in cui vengano aggiunte due palline, una con il numero 1 e l'altra con il numero 10.

RISPOSTA

Abbiamo $A=\{6,7,8,9,10\}$ e $B=\{2,4,6,8,10\}$; quindi $A\cap B=\{6,8,10\}$.

Nel 1° caso abbiamo uno spazio di prob. uniforme discreto e quindi

$$P(A|B)=\frac{\#(A\cap B)}{\#B}=\frac{3}{5}.$$

Nel ~~secondo~~ 2° caso abbiamo

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(\{6,8,10\})}{P(\{2,4,6,8,10\})}=\frac{\frac{1+1+2}{\cancel{12}}}{\frac{1+1+1+1+2}{\cancel{12}}}=\frac{4}{6}=\frac{2}{3}.$$