Calcolo delle Probabilità e Statistica 2021/2022

Spazi di probabilità

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

1 *L'incertezza e la probabilità*

Che cos'è la probabilità? Il concetto di *probabilità* è strettamente collegato al concetto di *incertezza*. A tal proposito, consideriamo la seguente proposizione:

$$A \;=\; \text{“domani a Bologna piove”}.$$

Essa riguarda il verificarsi di un evento futuro, per tale ragione non è possibile dire con certezza assoluta se è *vera* oppure *falsa*, è possibile però dire che è *probabile*. Mentre in matematica (in particolare, in logica matematica, anche detta *logica del certo*) si studiano solo proposizioni vere oppure false, nella realtà si ha a che fare con proposizioni di cui in genere è possibile solamente dire che sono probabili. Tali proposizioni riguardano infatti eventi (non solo futuri come nel caso di A, ma anche presenti o passati) di cui non si hanno tutte le informazioni a disposizione per dire con certezza assoluta se sono veri oppure falsi. A questo proposito, riportiamo la seguente citazione:

“... il caso della certezza, intesa come certezza assoluta, è, se non un'astrazione illusoria, per lo meno un caso limite, mentre sarebbe da considerarsi normale il caso dell'incertezza.”[1]

La probabilità è una quantificazione o misura dell'incertezza: in matematica si attribuisce il valore numerico 1 ad una proposizione vera e 0 ad una proposizione falsa; la probabilità permette di andare oltre questi due casi limite, assegnando ad una proposizione un qualunque valore numerico nell'intervallo $[0, 1]$. Gli estremi dell'intervallo, 0 e 1, corrispondono ai valori di verità *“falso”* e *“vero”* utilizzati normalmente in matematica.

Se consideriamo ad esempio A = “domani a Bologna piove”, la probabilità di tale affermazione, indicata con $\mathbb{P}(A)$, è un numero appartenente all'intervallo $[0, 1]$ tale che:

- $\mathbb{P}(A) = 0$ significa che con certezza assoluta la proposizione è *falsa*;
- $\mathbb{P}(A) = 1$ significa che con certezza assoluta la proposizione è *vera*;
- $\mathbb{P}(A) \in (0, 1)$ significa che la proposizione è probabile, in particolare più $\mathbb{P}(A)$ è vicino ad 1 più è probabile che l'evento “domani a Bologna piove” si verifichi.

In sintesi, possiamo dunque descrivere la probabilità come una

“misura dell'avverabilità di un evento”[2].

Ma come si assegna la probabilità? Perché dovrei ad esempio assegnare una probabilità del 60% alla proposizione A? L'assegnazione della probabilità si compone in generale di due fasi (anche se spesso si salta la prima fase).

1) In primo luogo, si mettono insieme tutte le informazioni che si hanno già a disposizione riguardo al fenomeno/esperimento che si sta studiando per giungere ad un'assegnazione *iniziale*, anche detta *a priori*. Si pensi ad esempio ai giochi d'azzardo e, in particolare, all'esperimento che consiste nel lanciare una moneta oppure un dado: in tal caso, grazie alla simmetria, è ragionevole assegnare la stessa probabilità ai vari esiti.

[1] Bruno de Finetti, *“Vero, falso, oppure probabile?”*, 1982.
[2] Italo Scardovi, *“Il tempo e il caso”*, 1999.

2) La seconda fase consiste nel ripetere l'esperimento in modo tale da ottenere un *campione* di dati che andranno analizzati tramite gli strumenti della *Statistica descrittiva.*

L'approccio che si basa solo sulla seconda fase è detto *frequentista*, perché di fatto sfrutta solamente le *frequenza* per assegnare la probabilità. Infatti, in tal caso la probabilità di una proposizione A si assegna, o meglio *si stima*, come segue:

$$\mathbb{P}(A) \approx \frac{\text{n}^\text{o} \text{ di volte che si verifica } A \text{ nel campione}}{n},$$

dove n è la numerosità del campione. Se al contrario si combinano le informazioni provenienti dal campione con l'assegnazione fatta *a priori*, si giunge all'approccio cosiddetto *bayesiano*. Chiaramente quest'ultimo approccio è più articolato, quindi spesso si preferisce usare l'approccio *frequentista*. I casi in cui ha senso utilizzare l'approccio *bayesiano* sono essenzialmente i seguenti:

- i *giochi d'azzardo*, che sono tra i pochissimi esempi in cui è molto facile l'assegnazione a priori (in cui si sfrutta la simmetria del problema);
- gli esperimenti che non si possono ripetere, come ad esempio le partite di calcio o quelli riguardanti il tempo metereologico (anche se oggi l'informatica permette di eseguire delle simulazioni al computer, dando così la possibilità di costruire un campione anche in questi casi).

Moltissimi esperimenti della Fisica o della Medicina sono invece facilmente ripetibili quindi ad essi si applica in generale l'approccio *frequentista*. In ogni caso, le conclusioni o *inferenze* così ottenute si definiscono *ingenue* e andrebbero *validate* con gli strumenti della *Statistica inferenziale* (*frequentista* o *bayesiana*).

La matematica della probabilità. Qualunque *interpretazione* si favorisca, qualsiasi *assegnazione* venga effettuata, vi è comunque un consenso generale sulla matematica della probabilità, ovvero sulle *regole* che la probabilità deve verificare. La matematica della probabilità si chiama *Calcolo delle probabilità* o, più semplicemente, *Probabilità*.
Ad esempio, oltre ad A consideriamo le seguenti proposizioni:

$$\begin{aligned} B &= \text{“domani a Bologna la temperatura è superiore a 25 °C”}, \\ C &= A \text{ e } B. \end{aligned}$$

Grazie al *Calcolo delle probabilità* sappiamo che, una volta che sono note le probabilità di A e di B, è nota anche la probabilità di C, infatti come vedremo $\mathbb{P}(C) = \mathbb{P}(A) + \mathbb{P}(B)$.
Nel 1854, George Boole, nel suo trattato *The Laws of Thought*, afferma che il problema generale del Calcolo delle probabilità è il seguente:

“date le probabilità di eventi qualsiasi, semplici o composti, condizionati o incondizionati, calcolare la probabilità di un altro evento qualsiasi, ugualmente arbitrario.”

In realtà oggi il *Calcolo delle probabilità* è molto più esteso di allora, per questo si preferisce a volte utilizzare il termine più generico *Probabilità* (anziché *Calcolo delle probabilità*).

Le tre questioni fondamentali della probabilità. Concludiamo questa sezione riportando le tre questioni fondamentali che ruotano attorno al concetto di probabilità.

- Che cos'è la probabilità?
- Come si assegna/stima la probabilità?
- Quali regole/assiomi verifica la probabilità?

La prima questione (a cui abbiamo già fornito una possibile risposta affermando che la probabilità è una "misura dell'avverabilità di un evento") è di pertinenza della *Filosofia*. La seconda è invece di competenza della *Statistica* e può essere affrontata con approcci differenti e su due livelli: *ingenuo* o *formale*. Infine, la terza questione può essere risolta servendosi unicamente di argomentazioni assiomatico-deduttive ed è dunque di pertinenza della *Matematica:* la disciplina che se ne occupa è il *Calcolo delle probabilità* o, semplicemente, la *Probabilità*.

2 Richiami di teoria degli insiemi

Come vedremo in seguito, un "evento" sarà descritto matematicamente da un insieme, oltre che da un'affermazione. Per questo motivo, è utile fare qualche richiamo di teoria degli insiemi.

Indichiamo con la lettera greca maiuscola *omega*, simbolo Ω, un *insieme* qualunque. Un generico *elemento* di Ω verrà indicato con la lettera greca minuscola *omega*: ω. Scriveremo $\omega \in \Omega$ per dire che ω *appartiene ad* Ω o, equivalentemente, che ω *è un elemento di* Ω.

SOTTOINSIEMI. Per indicare che A è un *sottoinsieme* di Ω, scriveremo

$$A \subset \Omega.$$

Si noti che tra tutti i sottoinsiemi di Ω ci sono anche[3] l'insieme vuoto (indicato con il simbolo $\emptyset$) e l'insieme Ω stesso, ovvero $\emptyset \subset \Omega$ e $\Omega \subset \Omega$. Indicheremo con $\mathscr{P}(\Omega)$ l'***insieme delle parti*** di Ω, ovvero l'insieme i cui elementi sono tutti i sottoinsiemi di Ω, compresi l'insieme vuoto $\emptyset$ e Ω stesso.

CARDINALITÀ. Indicheremo con $|\Omega|$ oppure $\#\Omega$ la *cardinalità* di un qualunque insieme Ω, ovvero il numero dei suoi elementi.

[3] L'insieme vuoto e Ω sono anche detti *sottoinsiemi impropri* di Ω, mentre tutti gli altri si chiamano *sottoinsiemi propri*.

Esercizio 2.1. *Si determini $\mathscr{P}(\Omega)$ e se ne calcoli la cardinalità nei seguenti casi:*

1) $\Omega = \{a\}$. Risposta: $\mathscr{P}(\Omega) = \{\emptyset, \{a\}\}$ e $|\mathscr{P}(\Omega)| = 2$.

2) $\Omega = \{a, b\}$. Risposta: $\mathscr{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$ e $|\mathscr{P}(\Omega)| = 4$.

3) $\Omega = \{a, b, c\}$. Risposta: $\mathscr{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$ e $|\mathscr{P}(\Omega)| = 8$.

Sulla base di quanto ottenuto nei tre casi appena studiati, qual è la cardinalità di $\mathscr{P}(\Omega)$ quando Ω ha n elementi? Risposta: $|\mathscr{P}(\Omega)| = 2^n$.

Operazioni insiemistiche. Ricordiamo la definizione di *unione*, *intersezione* e *complementazione*.

$$
\begin{aligned}
A \cup B &= \{\omega \in \Omega \colon \omega \text{ appartiene ad almeno uno tra } A \text{ e } B\}, \\
A_1 \cup A_2 \cup \cdots \cup A_n &= \{\omega \in \Omega \colon \omega \text{ appartiene ad almeno un insieme tra } A_1, A_2, \ldots, A_n\}, \\
A \cap B &= \{\omega \in \Omega \colon \omega \text{ appartiene sia ad } A \text{ che a } B\}, \\
A_1 \cap A_2 \cap \cdots \cap A_n &= \{\omega \in \Omega \colon \omega \text{ appartiene a tutti gli insiemi } A_1, A_2, \ldots, A_n\}, \\
A^c &= \{\omega \in \Omega \colon \omega \text{ non appartiene ad } A\}.
\end{aligned}
$$

Ricordiamo inoltre che il simbolo $B \backslash A$ ha il seguente significato:

$$B \backslash A = \{\omega \in B \colon \omega \text{ non appartiene ad } A\}.$$

Quindi, in particolare, $A^c = \Omega \backslash A$. Ricordiamo anche le ***leggi di De Morgan***:

- per due insiemi

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c,$$

- per n insiemi

$$(A_1 \cup A_2 \cup \cdots \cup A_n)^c = A_1^c \cap A_2^c \cap \cdots \cap A_n^c, \qquad (A_1 \cap A_2 \cap \cdots \cap A_n)^c = A_1^c \cup A_2^c \cup \cdots \cup A_n^c.$$

Ricordiamo infine le ***proprietà distributive*** dell'unione rispetto all'intersezione e dell'intersezione rispetto all'unione:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

e, più in generale,

$$A \cup \left(\bigcap_{i=1}^n B_i \right) = \bigcap_{i=1}^n (A \cup B_i), \qquad A \cap \left(\bigcup_{i=1}^n B_i \right) = \bigcup_{i=1}^n (A \cap B_i).$$

Unioni e intersezioni infinite. Nel seguito considereremo anche unioni e intersezioni di una famiglia numerabile di sottoinsiemi di Ω, che ora definiamo. Siano $A_1, A_2, \ldots, A_n, \ldots$

sottoinsiemi di Ω. Si dice che $A_1, A_2, \dots, A_n, \dots$ costituiscono una *famiglia numerabile* o *successione* di sottoinsiemi di Ω. Definiamo

$$\begin{aligned} \bigcup_{n=1}^{\infty} A_n &= \{\omega \in \Omega \colon \omega \in A_n \text{ per almeno un } n\}, \\ \bigcap_{n=1}^{\infty} A_n &= \{\omega \in \Omega \colon \omega \in A_n \text{ per ogni } n\}. \end{aligned}$$

Le leggi di De Morgan valgono anche per unioni e intersezioni infinite:

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c, \qquad \left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c.$$

Esercizio 2.2. *Per le seguenti successioni di insiemi, determinare* $\cup_{n=1}^{\infty} A_n$ *e* $\cap_{n=1}^{\infty} A_n$.

1) $A_n = \{n\}$. Risposta: $\cup_{n=1}^{\infty} A_n = \mathbb{N}$ e $\cap_{n=1}^{\infty} A_n = \emptyset$.

2) $A_n = [0, 1/n]$. Risposta: $\cup_{n=1}^{\infty} A_n = [0, 1]$ e $\cap_{n=1}^{\infty} A_n = \{0\}$.

Esercizio 2.3. *Siano* $A_1, A_2, \dots, A_n, \dots$ *tutti uguali all'insieme vuoto. Mostrare che tali insiemi sono tra loro* ***disgiunti***[a] *e la loro unione è uguale all'insieme vuoto.*

[a] $A_i \cap A_j = \emptyset$, per ogni $i \neq j$.

3 Modello matematico di un esperimento aleatorio

3.1 Esperimento aleatorio

Iniziamo con l'introdurre alcuni *concetti fondamentali*[4].

Un ***esperimento aleatorio*** *(detto anche* ***fenomeno aleatorio*** *o* ***prova*** *o* ***situazione d'incertezza****) è un esperimento di cui non conosciamo con certezza il risultato.*

Un ***esito*** *è un ipotetico risultato dell'esperimento aleatorio.*

Esempi classici di esperimenti aleatori sono i seguenti:

- lancio di una *moneta*, i cui esiti sono generalmente indicati con "testa" e "croce";
- lancio di un *dado*, i cui esiti sono generalmente indicati con i numeri naturali da 1 a 6.

[4] Le nozioni di *esperimento aleatorio* ed *esito* qui riportate, insieme alla nozione di *probabilità*, non sono vere e proprie definizioni matematiche. Tali concetti sono, più precisamente, *enti primitivi* (come *punto*, *retta*, *piano* in geometria). Possono quindi essere definiti solo intuitivamente e a partire da essi vengono formulate tutte le altre (vere) definizioni.

Osservazione. *L'esperimento aleatorio può essere suddiviso in diverse fasi che si chiamano* ***sotto-esperimenti aleatori****. Ad esempio, il lancio di due dadi è composto da due sotto-esperimenti aleatori.*

Grazie ai concetti di *esperimento aleatorio* ed *esito*, possiamo ora dare una prima definizione matematica, ovvero quella di *evento*.

Definizione 3.1. *Un* ***evento*** *è un'*affermazione *riguardante l'ipotetico risultato dell'esperimento aleatorio, di cui è possibile dire con certezza se è* vera *oppure* falsa *una volta noto l'esito dell'esperimento aleatorio.*

Gli esiti per cui un evento è vero si chiamano casi favorevoli *(per l'evento in questione), gli altri si chiamano* casi contrari*.*

Un evento viene solitamente indicato con una lettera maiuscola dell'alfabeto. Spesso si utilizzano le prime lettere dell'alfabeto: A, B, C, ...

Esempio 3.1. *Si lancia un dado. L'affermazione*

$$A \;=\; \text{“esce un numero pari”}$$

è un evento.

3.2 Modello matematico di un esperimento aleatorio

Il modello matematico di un esperimento aleatorio è una descrizione sintetica dell'esperimento stesso che fa uso della teoria degli insiemi.

Spazio campionario ed eventi. Diamo le seguenti definizioni (che sono vere e proprie definizioni matematiche).

Definizione 3.2. *Si chiama* ***spazio campionario*** *un qualunque insieme che contiene tutti gli ipotetici risultati dell'esperimento aleatorio, rappresentati secondo un opportuno codice.*

Lo spazio campionario si indica generalmente con la lettera greca maiuscola omega: Ω.

Un generico elemento di Ω verrà chiamato ***esito*** *e sarà indicato con la lettera greca minuscola omega: ω.*

Definizione 3.3. *Ogni* ***evento****, inteso come* affermazione*, è rappresentato dal sottoinsieme di Ω costituito dai casi favorevoli, ovvero dagli esiti per cui l'evento è vero (è rappresentato dall'insieme vuoto $\emptyset$ se è sempre falso).*

Un qualunque sottoinsieme di Ω lo chiameremo ancora ***evento****.*

Osservazione. *Il termine "evento" lo useremo dunque indistintamente per indicare sia l'*affermazione *che il* sottoinsieme. *Per tale ragione, nel seguito indicheremo entrambi con lo stesso simbolo (tipicamente una lettera maiuscola dell'alfabeto), come accade nell'esempio seguente in cui la lettera maiuscola A indica sia l'affermazione che il sottoinsieme.*

Esempio 3.2. *Si lancia un dado. Consideriamo l'evento*

$$A \;=\; \text{“}\textit{esce un numero pari}\text{”}.$$

Uno spazio campionario naturale per questo esperimento aleatorio è l'insieme

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

Dunque l'evento A è rappresentato dal sottoinsieme

$$A \;=\; \{2, 4, 6\}.$$

Alcuni eventi hanno nomi specifici.

- L'***evento certo*** è l'insieme Ω, che corrisponde ad un'affermazione sempre vera, qualunque sia l'esito dell'esperimento aleatorio.
- L'***evento impossibile*** è l'insieme $\emptyset$, che corrisponde ad un'affermazione sempre falsa, qualunque sia l'esito dell'esperimento aleatorio.
- Un ***evento elementare***[5] è un sottoinsieme di Ω che contiene un solo elemento, quindi corrisponde ad un'affermazione che è vera per un solo esito.

Esempio 3.3. *Si lancia un dado. Consideriamo gli eventi:*

$$\begin{aligned} A &= \text{“}\textit{esce un numero naturale compreso tra } 1 \textit{ e } 6\text{”}, \\ B &= \text{“}\textit{esce un numero maggiore o uguale a } 7\text{”}, \\ C &= \text{“}\textit{esce il numero } 2\text{”}. \end{aligned}$$

Si consideri lo spazio campionario

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

Allora gli eventi A, B e C sono rappresentati dai seguenti sottoinsiemi:

$$A \;=\; \Omega, \qquad B \;=\; \emptyset, \qquad C \;=\; \{2\}.$$

Quindi A è l'evento certo, B è l'evento impossibile, C è un evento elementare.

[5] In alcuni libri *esito* ed *evento elementare* sono considerati sinonimi.

Operazioni tra eventi. Come abbiamo visto gli eventi sono descritti sia da affermazioni che da insiemi. Sulle affermazioni possiamo eseguire le operazioni date dai *connettivi logici* (ovvero *disgiunzione*, *congiunzione*, *negazione*), le quali corrispondono alle relative *operazioni insiemistiche* (ovvero *unione*, *intersezione*, *complementazione*). Si tratta in entrambi i casi di una struttura algebrica nota come *algebra booleana*. Abbiamo inoltre delle corrispondenze tra opportune *relazioni*, come riportato nella tabella seguente:

Connettivi logici	Operazioni/Relazioni insiemistiche
Disgiunzione A o B	**Unione** $A \cup B$
Congiunzione A e B	**Intersezione** $A \cap B$
Negazione non A	**Complementazione** A^c
Implicazione $A \Longrightarrow B$	**Inclusione** $A \subset B$
Doppia implicazione $A \Longleftrightarrow B$	**Uguaglianza** $A = B$

Esercizio 3.1. *A cosa corrisponde "A* ***ma non*** *B"?*

Osservazione*. ***Quali sottoinsiemi di Ω sono eventi?*** *Distinguiamo due casi.*

- *Quando Ω ha cardinalità finita o, al più, numerabile è naturale supporre che tutti i sottoinsiemi di Ω siano eventi. Quindi la famiglia degli eventi corrisponde a $\mathscr{P}(\Omega)$, l'insieme delle parti di Ω. Lo stesso vale nel caso in cui Ω è non numerabile, ma la probabilità è discreta (si veda l'*Osservazione 2 *nella sezione successiva).*

- *Se invece Ω è un insieme non numerabile, ad esempio $\Omega = \mathbb{R}$, e la probabilità non è discreta, risulta necessario considerare una famiglia di eventi più piccola rispetto a $\mathscr{P}(\Omega)$. Tale famiglia, generalmente indicata con $\mathcal{F}$, non contiene certi sottoinsiemi "anomali" di Ω (ad esempio, nel caso $\Omega = \mathbb{R}$, $\mathcal{F}$ non conterrà l'insieme di Vitali). Affinché $\mathcal{F}$ sia una famiglia di eventi si chiede che sia una σ-**algebra** di sottoinsiemi di Ω, ovvero:*

  - *Ω è un evento, quindi $\Omega \in \mathcal{F}$;*
  - *se A è un evento, quindi $A \in \mathcal{F}$, allora anche $A^c \in \mathcal{F}$;*
  - *se $A_1, A_2, \dots, A_n, \dots$ è una successione di eventi, allora anche $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.*

Per semplicità, nel seguito supporremo sempre che tutti i sottoinsiemi di Ω siano eventi.

3.3 Assiomi della probabilità

L'ultimo elemento del modello matematico di un esperimento aleatorio è la *probabilità*, che, come già detto, è un *concetto primitivo* e sulla cui definizione (non matematica) non torniamo, ma ci limitiamo a ricordare la seguente: *la probabilità è un numero reale compreso tra zero ed uno che rappresenta la "misura dell'avverabilità di un evento".* Veniamo ora agli ***assiomi***[6] della probabilità.

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

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

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

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

Assioma III. *Vale la proprietà di* ***additività 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

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

[a]Anche detta σ-**additività**.
[b]In formule: $A_i \cap A_j = \emptyset$, per ogni $i \neq j$. In altri termini, non hanno elementi in comune.

Osservazione 1. *La probabilità $\mathbb{P}$ è dunque una funzione che ad ogni sottoinsieme A di Ω fa corrispondere un numero in $[0,1]$:*

$$A \in \mathscr{P}(\Omega) \;\mapsto\; \mathbb{P}(A) \in [0,1],$$

dove $\mathscr{P}(\Omega)$ è l'insieme delle parti di Ω e rappresenta il dominio[7] *della funzione $\mathbb{P}$, mentre il codominio è l'intervallo $[0,1]$. In simboli:*

$$\mathbb{P} \colon \mathscr{P}(\Omega) \;\longrightarrow\; [0,1].$$

Osservazione 2. *Una probabilità $\mathbb{P}$ si dice* ***discreta*** *se assume un numero finito o al più un'infinità numerabile di valori. Ad esempio, la delta di Dirac (introdotta nell'Esempio 3.4 riportato qui di seguito) è una probabilità discreta in quanto assume due soli valori:* 0 *e* 1.

Definizione 3.4. *La coppia*[a] $(\Omega, \mathbb{P})$ *si dice* ***spazio di probabilità*** *o* ***modello matematico dell'esperimento aleatorio****.*

[a]Quando $\mathbb{P}$ è definita su una σ-algebra $\mathcal{F}$, allora si chiama spazio di probabilità la terna $(\Omega, \mathcal{F}, \mathbb{P})$.

[6]Tali assiomi sono stati formulati dal matematico sovietico Kolmogorov nel 1933.
[7]Come abbiamo detto alla fine della sezione precedente, quando Ω ha cardinalità non numerabile e $\mathbb{P}$ non è discreta risulta necessario definire la probabilità $\mathbb{P}$ su un dominio più piccolo, ossia su una σ-algebra $\mathcal{F}$ di sottoinsiemi di Ω.

Esempio 3.4. *Siano* $\Omega = \mathbb{R}$ *e* x_0 *un numero reale fissato. Si consideri la funzione*

$$\delta_{x_0} \colon \mathscr{P}(\mathbb{R}) \longrightarrow [0,1]$$

data da

$$\delta_{x_0}(A) = \begin{cases} 1, & se\ x_0 \in A, \\ 0, & se\ x_0 \notin A, \end{cases} \qquad \forall\, A \subset \mathbb{R}.$$

δ_{x_0} *si chiama* ***delta di Dirac in*** x_0. *Si verifica che* δ_{x_0} *soddisfa gli* Assiomi I-II-III, *quindi* δ_{x_0} *è una probabilità e* $(\mathbb{R}, \delta_{x_0})$ *è uno spazio di probabilità.*

Esempio 3.5. *Siano* $\Omega = \mathbb{R}$, $x_1, \dots, x_n$ *e* $p_1, \dots, p_n$ *numeri reali, con*

$$0 \le p_i \le 1, \qquad \sum_{i=1}^{n} p_i = 1. \tag{3.1}$$

Si consideri la funzione

$$\mathbb{P} \colon \mathscr{P}(\mathbb{R}) \longrightarrow [0,1]$$

data da

$$\mathbb{P}(A) = \sum_{i=1}^{n} p_i\, \delta_{x_i}(A), \qquad \forall\, A \subset \mathbb{R}.$$

$\mathbb{P}$ *è dunque una combinazione lineare*[a] *di delta di Dirac. Si verifica che* $\mathbb{P}$ *soddisfa gli* Assiomi I-II-III, *quindi* $\mathbb{P}$ *è una probabilità e* $(\mathbb{R}, \mathbb{P})$ *è uno spazio di probabilità.*

[a]Si tratta in particolare di una *combinazione convessa*, ovvero di una combinazione lineare con coefficienti che verificano (3.1).

3.4 Conseguenze degli assiomi

Teorema 3.1. *Sia* $(\Omega, \mathbb{P})$ *uno spazio di probabilità. Le seguenti proprietà della probabilità* $\mathbb{P}$ *discendono dagli* Assiomi I-II-III*:*

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

V) ***Additività finita:*** *se* A *e* B *sono disgiunti allora*

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

Più in generale, se $A_1, \dots, A_n$ *sono tra loro disgiunti allora*

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

VI) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

VII) ***Monotonia:*** *se* $A \subset B$, *allora* $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Dimostrazione.

IV) Sappiamo dall'Assioma I che $\mathbb{P}(\emptyset)$ è un numero che verifica $0 \leq \mathbb{P}(\emptyset) \leq 1$. Per semplificare la notazione, poniamo $p := \mathbb{P}(\emptyset)$. Dobbiamo mostrare che $p = 0$. A tale scopo utilizziamo l'*additività numerabile* (Assioma III) prendendo come successione $A_1, A_2, \ldots, A_n, \ldots$ gli insiemi

$$A_1 = \emptyset, \qquad A_2 = \emptyset, \qquad \cdots \qquad A_n = \emptyset, \qquad \cdots$$

Tali insiemi sono evidentemente disgiunti tra loro (cioè non hanno elementi in comune; questo è evidente dato che sono tutti uguali all'insieme vuoto, quindi ciascun insieme non contiene alcun elemento). Inoltre, posto

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

si ha che $A = \emptyset$. Segue allora dall'additività numerabile (Assioma III)

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

ovvero (ricordando che $p = \mathbb{P}(\emptyset) = \mathbb{P}(A) = \mathbb{P}(A_n)$)

$$p = \sum_{n=1}^{+\infty} p.$$

Questa è un'equazione nell'incognita p, che è verificata solo per $p = 0$. Infatti il primo termine è uguale a p (quindi è uguale a zero se $p = 0$), mentre il secondo è dato da

$$\sum_{n=1}^{+\infty} p = \begin{cases} 0, & \text{se } p = 0, \\ +\infty, & \text{se } 0 < p \leq 1. \end{cases}$$

V) Dimostriamo il caso con due insiemi A e B (il caso con n insiemi si dimostra in modo analogo). Utilizziamo l'*additività numerabile* (Assioma III) prendendo come successione $A_1, A_2, \ldots, A_n, \ldots$ gli insiemi

$$A_1 = A, \qquad A_2 = B, \qquad A_n = \emptyset, \quad \text{per ogni } n \geq 3.$$

Tali insiemi sono tra loro disgiunti, inoltre

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

Segue allora dall'additività numerabile (Assioma III) e dal fatto che $\mathbb{P}(A_n) = 0$ quando $n \geq 3$

$$\mathbb{P}(A \cup B) \;=\; \sum_{n=1}^{\infty} \mathbb{P}(A_n) \underset{\substack{\uparrow \\ \mathbb{P}(A_n)=0,\, n\geq 3}}{=} \mathbb{P}(A) + \mathbb{P}(B).$$

VI) Utilizziamo l'*additività finita* dimostrata al punto precedente prendendo gli insiemi

$$A \quad \text{e} \quad A^c.$$

Si noti che A e A^c sono disgiunti. Inoltre

$$\Omega \;=\; A \cup A^c.$$

Per l'additività finita abbiamo che

$$\mathbb{P}(\Omega) \;=\; \mathbb{P}(A) + \mathbb{P}(A^c).$$

Per l'Assioma II si ha che $\mathbb{P}(\Omega) = 1$. Quindi

$$\mathbb{P}(A^c) \;=\; 1 - \mathbb{P}(A).$$

VII) Se $A \subset B$ allora possiamo scrivere B come unione di A e $B \backslash A$, cioè $B = A \cup (B \backslash A)$. Si noti che A e $B \backslash A$ sono disgiunti. Quindi, per l'*additività finita*,

$$\mathbb{P}(B) \;=\; \mathbb{P}\big(A \cup (B \backslash A)\big) \;=\; \mathbb{P}(A) + \mathbb{P}(B \backslash A) \underset{\substack{\uparrow \\ \mathbb{P}(B \backslash A)\geq 0}}{\geq} \mathbb{P}(A).$$

$\square$

Alcuni eventi, in base alla loro probabilità, hanno nomi specifici.

- Un evento A si dice ***quasi certo*** (spesso abbreviato ***q.c.***) se $\mathbb{P}(A) = 1$.
- Un evento A si dice ***quasi impossibile*** se $\mathbb{P}(A) = 0$.

Esempio 3.6. *Si lancia un dado. Consideriamo gli eventi:*

$$\begin{aligned} A &= \text{“esce un numero naturale compreso tra 1 e 6”}, \\ B &= \text{“esce un numero maggiore o uguale a 7”}. \end{aligned}$$

Abbiamo già visto in un esempio precedente che scegliendo come spazio campionario l'insieme

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

si ottiene che A è l'evento certo e B è l'evento impossibile. In formule, $A = \Omega$ e $B = \emptyset$. Consideriamo ora come spazio campionario l'insieme

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

munito della probabilità discreta

$$\mathbb{P}(E) \;=\; \sum_{i=1}^{6} \frac{1}{6}\,\delta_i(E), \qquad \forall\, E \subset \mathbb{R}.$$

In tal caso, otteniamo

$$A \;=\; \{1,2,3,4,5,6\} \;\neq\; \Omega, \qquad\qquad B \;=\; [7,+\infty) \;\neq\; \emptyset.$$

Quindi A non è l'evento certo e B non è l'evento impossibile. Tuttavia, A è un evento quasi certo e B è un evento quasi impossibile, infatti vale che

$$\mathbb{P}(A) \;=\; 1, \qquad\qquad \mathbb{P}(B) \;=\; 0.$$

Esercizio 3.2. *Lanciamo un dado* ***perfettamente bilanciato***[a] *a sei facce.*

Qual è la probabilità che esca un numero maggiore o uguale a 3?

[a] ***equilibrato***, ***regolare***, ***non truccato***, . . ., oppure, equivalentemente, se il testo non dice nulla è sottinteso che sia *perfettamente bilanciato* (ciò è una conseguenza del cosiddetto *"principio di ragione non sufficiente"* di Laplace: se non si ha alcuna informazione a riguardo, si suppone che tutti i risultati dell'esperimento siano tra loro equiprobabibili, dato che, non avendo alcuna informazione più precisa, non c'è alcuna ragione per cui uno debba essere più probabile degli altri).

Soluzione. Nel testo dell'esercizio si fa riferimento all'evento

$$A \;=\; \text{“esce un numero maggiore o uguale a 3”}.$$

Per risolvere l'esercizio dobbiamo innanzitutto trovare uno spazio campionario per l'esperimento aleatorio in questione. In tal caso, è naturale scegliere come spazio campionario l'insieme

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

Dunque l'evento A è rappresentato dal sottoinsieme

$$A \;=\; \{3,4,5,6\}.$$

Resta da determinare $\mathbb{P}(A)$. Dal testo dell'esercizio sappiamo che il dado è *perfettamente bilanciato*, ovvero che gli eventi elementari

$$\{1\}, \qquad \{2\}, \qquad \{3\}, \qquad \{4\}, \qquad \{5\}, \qquad \{6\}$$

sono *equiprobabili*:

$$\mathbb{P}(\{1\}) \;=\; \mathbb{P}(\{2\}) \;=\; \mathbb{P}(\{3\}) \;=\; \mathbb{P}(\{4\}) \;=\; \mathbb{P}(\{5\}) \;=\; \mathbb{P}(\{6\}). \tag{3.2}$$

Dato che gli eventi $\{1\},\ldots,\{6\}$ sono tra loro disgiunti e la loro unione è pari a

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

utilizzando l'*additività finita* e la proprietà $\mathbb{P}(\Omega) = 1$ (Assioma II), otteniamo

$$\mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) \;=\; 1.$$

Quest'ultima equazione, insieme alle cinque equazioni (3.2), fornisce un sistema di sei equazioni in sei incognite:

$$\begin{cases} \mathbb{P}(\{1\}) \;=\; \mathbb{P}(\{2\}), \\ \mathbb{P}(\{2\}) \;=\; \mathbb{P}(\{3\}), \\ \mathbb{P}(\{3\}) \;=\; \mathbb{P}(\{4\}), \\ \mathbb{P}(\{4\}) \;=\; \mathbb{P}(\{5\}), \\ \mathbb{P}(\{5\}) \;=\; \mathbb{P}(\{6\}), \\ \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) \;=\; 1. \end{cases}$$

Tale sistema si risolve facilmente, ponendo $x = \mathbb{P}(\{1\}) = \cdots = \mathbb{P}(\{6\})$ e sostituendo nell'ultima equazione, che diventa $6\,x = 1$. Si conclude che l'unica soluzione del sistema è data da

$$\mathbb{P}(\{1\}) \;=\; \mathbb{P}(\{2\}) \;=\; \mathbb{P}(\{3\}) \;=\; \mathbb{P}(\{4\}) \;=\; \mathbb{P}(\{5\}) \;=\; \mathbb{P}(\{6\}) \;=\; \frac{1}{6}.$$

Ricordando che $A = \{3, 4, 5, 6\}$, dall'*additività finita* si ottiene

$$\mathbb{P}(A) \;=\; \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{5\}) + \mathbb{P}(\{6\}) \;=\; \frac{4}{6}.$$

$\square$

Nell'esercizio appena svolto, dato un qualunque evento A vale la formula

$$\mathbb{P}(A) \;=\; \frac{\text{n}^\text{o} \text{ di eventi elementari che compongono } A}{6} \;=\; \frac{\textbf{\textit{casi favorevoli}}}{\textbf{\textit{casi possibili}}}.$$

Questa proprietà vale *solamente* quando l'esperimento aleatorio può essere descritto da uno spazio campionario Ω ***finito*** con esiti ***equiprobabili***, come affermato nel seguente Teorema.

Teorema 3.2. *Consideriamo un esperimento aleatorio descritto da uno spazio campionario* ***finito****:*

$$\Omega \;=\; \{\omega_1, \ldots, \omega_N\}$$

con esiti ***equiprobabili****:*

$$\mathbb{P}(\{\omega_1\}) \;=\; \mathbb{P}(\{\omega_2\}) \;=\; \cdots \;=\; \mathbb{P}(\{\omega_N\}).$$

In tal caso diciamo che $\mathbb{P}$ è la ***probabilità uniforme*** *e valgono le seguenti proprietà:*

1) *dato un qualunque evento elementare $\{\omega_i\}$, vale che*

$$\mathbb{P}(\{\omega_i\}) \;=\; \frac{1}{N};$$

2) *dato un qualunque evento A, vale la* ***formula di Laplace***

$$\mathbb{P}(A) \;=\; \frac{n^o \text{ di eventi elementari che compongono } A}{N} \;=\; \frac{\textbf{\textit{casi favorevoli}}}{\textbf{\textit{casi possibili}}}.$$

Dimostrazione. Si ragiona come nell'Esercizio 3.2. Più precisamente, sappiamo che

$$\mathbb{P}(\{\omega_1\}) \;=\; \mathbb{P}(\{\omega_2\}) \;=\; \cdots \;=\; \mathbb{P}(\{\omega_N\}) \tag{3.3}$$

e

$$\mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) + \cdots + \mathbb{P}(\{\omega_N\}) \;=\; 1.$$

Quest'ultima equazione, insieme alle $N-1$ equazioni (3.3), fornisce un sistema di N equazioni in N incognite:

$$\begin{cases} \mathbb{P}(\{\omega_1\}) \;=\; \mathbb{P}(\{\omega_2\}), \\ \mathbb{P}(\{\omega_2\}) \;=\; \mathbb{P}(\{\omega_3\}), \\ \mathbb{P}(\{\omega_3\}) \;=\; \mathbb{P}(\{\omega_4\}), \\ \vdots, \\ \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) + \cdots + \mathbb{P}(\{\omega_N\}) \;=\; 1. \end{cases}$$

Risolvendo il sistema, si ottiene

$$\mathbb{P}(\{\omega_1\}) \;=\; \mathbb{P}(\{\omega_2\}) \;=\; \cdots \;=\; \mathbb{P}(\{\omega_N\}) \;=\; \frac{1}{N},$$

che dimostra la proprietà 1) enunciata nel Teorema. Infine, la proprietà 2) segue direttamente dall'additività finita. □

Nell'ambito degli spazi di probabilità con spazio campionario *finito*, il caso di esiti *equiprobabili* è estremamente particolare. Tuttavia, se guardiamo ai libri di testo di Calcolo delle probabilità, la stragrande maggioranza degli esercizi ha esiti equiprobabili. Questo perché, se gli esiti non sono equiprobabili, il testo dell'esercizio deve in aggiunta specificare esattamente in che modo non lo sono, dato che ci sono infinite[8] possibilità. A tal proposito, si veda l'esercizio seguente.

[8]Infatti, se gli esiti sono equiprobabili, dal Teorema 3.2 sappiamo che necessariamente $\mathbb{P}(\{\omega_1\}) = \cdots = \mathbb{P}(\{\omega_N\}) = \frac{1}{N}$. Se invece non sono equiprobabili, sappiamo solo (dagli Assiomi I e II) che

$$0 \;\le\; \mathbb{P}(\{\omega_i\}) \;\le\; 1, \qquad\qquad \mathbb{P}(\{\omega_1\}) + \cdots + \mathbb{P}(\{\omega_N\}) \;=\; 1. \tag{3.4}$$

Possiamo dunque scegliere come vogliamo $\mathbb{P}(\{\omega_1\}), \ldots, \mathbb{P}(\{\omega_N\})$, purché i vincoli (3.4) siano verificati.

Esercizio 3.3. *Si dispone di un dado* ***non bilanciato*** *a forma di tetraedro regolare con le facce numerate da* 1 *a* 4*. Lanciando il dado, la probabilità che esca* 1 *è il doppio della probabilità che esca* 2*, che a sua volta è il doppio della probabilità che esca* 3*, che a sua volta è il doppio della probabilità che esca* 4*.*

Se si lancia il dado, qual è la probabilità che esca un numero pari?

Soluzione. Nel testo dell'esercizio si fa riferimento all'evento

$$A \;=\; \text{“esce un numero pari”}.$$

In tal caso, è naturale scegliere come spazio campionario l'insieme

$$\Omega \;=\; \{1,2,3,4\},$$

dunque

$$A \;=\; \{2,4\}.$$

Resta da determinare $\mathbb{P}(A)$. Dal testo dell'esercizio sappiamo che il dado *non è bilanciato*, è più precisamente che

> *“... la probabilità che esca* 1 *è il doppio della probabilità che esca* 2*,*
> *che a sua volta è il doppio della probabilità che esca* 3*,*
> *che a sua volta è il doppio della probabilità che esca* 4*.”*

Tradotto in formule, diventa

$$\mathbb{P}(\{1\}) \;=\; 2\,\mathbb{P}(\{2\}), \qquad \mathbb{P}(\{2\}) \;=\; 2\,\mathbb{P}(\{3\}), \qquad \mathbb{P}(\{3\}) \;=\; 2\,\mathbb{P}(\{4\}). \tag{3.5}$$

Inoltre, deve valere che

$$\mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) \;=\; 1.$$

Quest'ultima equazione, insieme alle tre equazioni (3.5), fornisce un sistema di quattro equazioni in quattro incognite:

$$\begin{cases} \mathbb{P}(\{1\}) \;=\; 2\,\mathbb{P}(\{2\}), \\ \mathbb{P}(\{2\}) \;=\; 2\,\mathbb{P}(\{3\}), \\ \mathbb{P}(\{3\}) \;=\; 2\,\mathbb{P}(\{4\}), \\ \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) \;=\; 1. \end{cases}$$

Tale sistema si risolve facilmente, ponendo $x = \mathbb{P}(\{4\})$, da cui si ottiene $\mathbb{P}(\{3\}) = 2\,x$, $\mathbb{P}(\{2\}) = 4\,x$ e $\mathbb{P}(\{1\}) = 8\,x$. Sostituendo nell'ultima equazione, si ottiene $15\,x = 1$. Si conclude che l'unica soluzione del sistema è data da

$$\begin{cases} \mathbb{P}(\{1\}) \;=\; \frac{8}{15}, \\ \mathbb{P}(\{2\}) \;=\; \frac{4}{15}, \\ \mathbb{P}(\{3\}) \;=\; \frac{2}{15}, \\ \mathbb{P}(\{4\}) \;=\; \frac{1}{15}. \end{cases}$$

Ricordando che $A = \{2, 4\}$, dall'*additività finita* si ottiene

$$\mathbb{P}(A) \;=\; \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) \;=\; \frac{1}{3}.$$

$\square$

Concludiamo questa sezione enunciando un'ultima proprietà della probabilità. Abbiamo visto che se A e B sono eventi *disgiunti*, allora

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

Cosa possiamo dire se A e B *non sono disgiunti*?

Teorema 3.3 (***Formula dell'unione di due eventi***)**.** *Siano A e B due eventi qualunque (non necessariamente disgiunti), allora*

$$\mathbb{P}(A \cup B) \;=\; \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \tag{3.6}$$

OSSERVAZIONE 1. *Si noti che nel caso in cui A e B sono* disgiunti*, si ha che $A \cap B = \emptyset$, quindi $\mathbb{P}(A \cap B) = 0$. Applicando la formula* (3.6) *del Teorema 3.3 ritroviamo dunque la formula dell'*additività finita

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

OSSERVAZIONE 2. *Per convincersi della validità della formula* (3.6)*, basta osservare che la somma $\mathbb{P}(A) + \mathbb{P}(B)$ conta due volte l'intersezione, per tale ragione dobbiamo sottrarre $\mathbb{P}(A \cap B)$.*

Dimostrazione. Consideriamo gli insiemi

$$C_1 \;=\; A \backslash B, \qquad\qquad C_2 \;=\; B \backslash A.$$

È facile convincersi (utilizzando ad esempio un diagramma di Eulero-Venn) che gli insiemi $C_1, C_2, A \cap B$ sono disgiunti e la loro unione è $A \cup B$. Quindi, dall'additività finita si ha che

$$\mathbb{P}(A \cup B) \;=\; \mathbb{P}(C_1) + \mathbb{P}(C_2) + \mathbb{P}(A \cap B).$$

A secondo membro, aggiungiamo e sottraiamo la quantità $\mathbb{P}(A \cap B)$, trovando

$$\mathbb{P}(A \cup B) \;=\; \mathbb{P}(C_1) + \mathbb{P}(A \cap B) + \mathbb{P}(C_2) + \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B). \tag{3.7}$$

Ora, notiamo che

$$\begin{aligned} \mathbb{P}(A) \;&=\; \mathbb{P}(C_1) + \mathbb{P}(A \cap B), \\ \mathbb{P}(B) \;&=\; \mathbb{P}(C_2) + \mathbb{P}(A \cap B). \end{aligned}$$

Utilizzando queste ultime due uguaglianze in (3.7), otteniamo la formula (3.6). $\square$

Infine, è interessante osservare come la formula dell'unione divenga "ingombrante" quando si passa alla probabilità dell'unione di tre o più eventi (non necessariamente disgiunti). Ad esempio, nel caso di tre eventi, vale che

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) \; &= \; \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ &\quad - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\ &\quad + \mathbb{P}(A \cap B \cap C).\end{aligned}$$

4* Una breve storia della Probabilità

Ripercorriamo qui le tappe principali della storia della Probabilità.

Gli albori. La formalizzazione dell'incertezza è un'acquisizione recente, iniziata pochi secoli fa, anzi pochissimi se confrontati con quelli impiegati per indagare sul concetto di numero intero o di retta.

Gli albori della Probabilità risalgono al Rinascimento, dove la volontà di un'espansione del dominio delle conoscenze razionali spingono più o meno consciamente verso la razionalizzazione anche dell'incerto. I primi studi riguardano problemi tratti dai giochi d'azzardo, che tuttavia esistevano già da qualche[9] millennio.

Dadi
Iran
$\sim$ 3000 a.C.

Achille gioca a dadi con Aiace
Anfora in tomba etrusca
$\sim$ 540 a.C.

I bari
Caravaggio
1594

Nel Rinascimento, infatti, i nobili hanno tra i loro passatempi preferiti proprio i giochi d'azzardo e affidano a studiosi del tempo il compito di risolvere i loro quesiti. Questo è il motivo che spinge Galileo Galilei a scrivere il libro *Sopra le scoperte dei dadi* del 1596 nel quale, su richiesta del Granduca di Toscana, tratta del lancio di tre dadi. Altri matematici dell'epoca che affrontano questo genere di problemi sono Luca Pacioli, Girolamo Cardano, Niccolò Tartaglia e Pietro Cataneo. Al tema del gioco dei dadi, Cardano dedica un intero trattato, *"De ludo aleae"* (1523, ma pubblicato postumo nel 1663). Nonostante questi lavori, non si può ancora dire che sia presente una concezione probabilistica. Questo si capisce da come vengono affrontati alcuni problemi, come ad esempio il cosiddetto *problema di ripartizione della posta in gioco.*

Problema della posta. *In un gioco a due giocatori, ogni partita vinta frutta* 1 *punto e vince chi per primo raggiunge* 7 *punti. Due giocatori A e B che in ciascuna partita hanno la stessa*

[9]Ciò è confermato anche dal fatto che il termine "aleatorio", sinonimo di "incerto", viene dal latino "alea", che significa "gioco dei dadi". Così anche "azzardo" viene dall'arabo "az-zahr", che significa "dado".

probabilità di vincere si sfidano.
Sapendo che il premio è 22 ducati, se per una qualche ragione il gioco viene interrotto quando A e B hanno rispettivamente 5 e 3 punti, ***come va suddivisa la posta in gioco****?*

Soluzione. *Poniamo*

$$P_{5,3} \;=\; \text{“}\textit{probabilità che il giocatore } A \textit{ vinca quando il risultato parziale è } 5 \textit{ a } 3\text{”}.$$

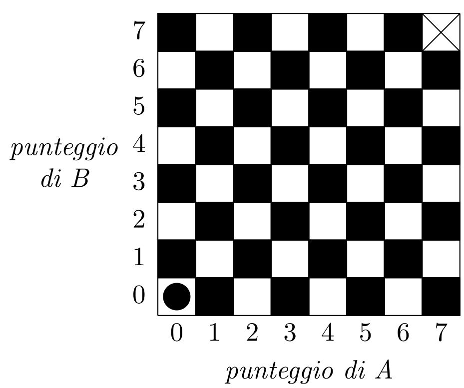
La suddivisione della posta è[10]

$$22 \cdot P_{5,3} \textit{ al giocatore } A \qquad e \qquad 22 \cdot (1 - P_{5,3}) \textit{ al giocatore } B.$$

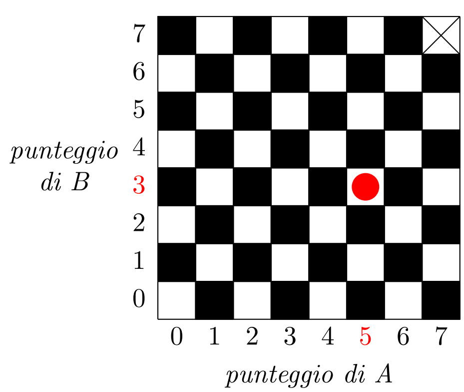
Resta dunque da calcolare $P_{5,3}$. In generale, poniamo

$$P_{i,j} \;=\; \text{“}\textit{probabilità che il giocatore } A \textit{ vinca quando il risultato parziale è } i \textit{ per } A \textit{ a } j \textit{ per } B\text{”}.$$

Per calcolare $P_{i,j}$ è utile servirsi di una scacchiera: una pedina è inizialmente collocata nella casella in basso a sinistra di una scacchiera, come in figura. Dopo ogni partita, se vince A la pedina viene spostata nella casella alla sua destra, altrimenti nella casella sopra di essa.



Ogni casella della scacchiera corrisponde ad una coppia di numeri che, a sua volta, corrisponde al risultato parziale in quell'istante del gioco. In base al testo dell'esercizio, il gioco viene interrotto quando il risultato è 5 a 3, quindi la pedina si trova nella posizione riportata nella seguente figura:



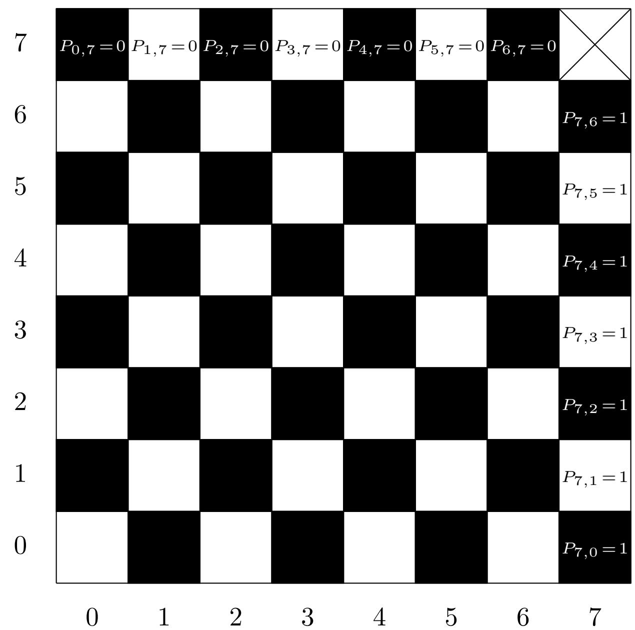
[10]Si tratta di due valori attesi relativi alle variabili aleatorie: $22 \cdot 1_{E_{5,3}}$ e $22 \cdot 1_{E^c_{5,3}}$, dove $E_{5,3}$ è l'evento "il giocatore A vince quando il risultato parziale è 5 a 3".

Il gioco termina quando si raggiunge una casella a destra della scacchiera (in tal caso vince A), oppure una casella in alto (in tal caso vince B). La casella nell'angolo in alto a destra della scacchiera, contrassegnata con la "X", non viene mai utilizzata, in quanto il gioco termina prima di poterci arrivare.

All'interno di ciascuna casella è utile riportare la probabilità $P_{i,j}$. Partendo dalle caselle agli estremi della scacchiera, abbiamo che:

- *se $i = 7$ il gioco è appena terminato e ha vinto A, quindi $P_{7,j} = 1$;*
- *al contrario, se $j = 7$ ha vinto B, quindi $P_{i,7} = 0$.*

Abbiamo dunque la seguente figura:



Ora, come possiamo calcolare ad esempio $P_{6,6}$? Se il risultato parziale del gioco è 6 a 6, allora il gioco si conclude in una partita. Dato che i due giocatori, ad ogni partita, hanno la stessa probabilità di vincere, si ha che $P_{6,6}$ vale $1/2$. Notiamo allora che è verificata l'uguaglianza

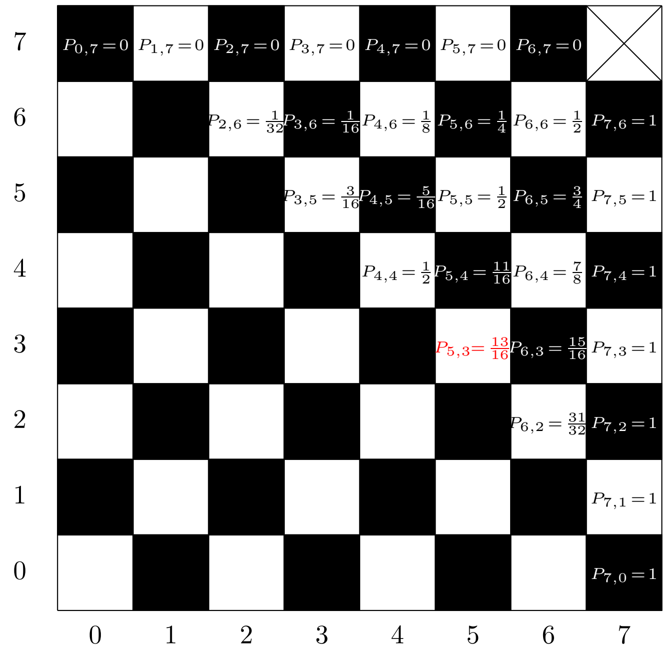
$$P_{6,6} \;=\; \frac{1}{2}P_{7,6} + \frac{1}{2}P_{6,7}.$$

In generale vale la seguente uguaglianza

$$P_{i,j} \;=\; \frac{1}{2}P_{i+1,j} + \frac{1}{2}P_{i,j+1}.$$

Possiamo dunque applicare questa formula per determinare in modo ricorsivo le probabilità $P_{i,j}$

fino ad arrivare a $P_{5,3}$*, come riportato nella figura seguente:*



Abbiamo dunque risolto il problema della posta.

Il problema della posta è stato affrontato da vari studiosi, tra cui Luca Pacioli, Girolamo Cardano, Niccolò Tartaglia, Pietro Cataneo, Blaise Pascal e Pierre de Fermat.

Luca Pacioli affronta per primo questo problema fornendo nel 1494 una soluzione, che tuttavia risulta essere sbagliata. Infatti suddivide la posta proporzionalmente in base ai punti raggiunti da ciascuno al momento della sospensione della partita. Per Pacioli contano dunque solo le partite già giocate. Questa soluzione non è corretta, come si vede ad esempio quando il gioco si interrompe dopo una sola partita: in tal caso il risultato è 1 a 0 oppure 0 e 1. In ogni caso, tutta la posta va al giocatore che ha vinto quell'unica partita. Al contrario, è chiaro che è molto più ragionevole spartire in parti quasi uguali la posta dato che aver fatto un'unica partita è quasi come non aver nemmeno giocato. Il ragionamento adottato da Luca Pacioli è dunque sbagliato ed è in effetti l'opposto del ragionamento probabilistico (in cui si considerano le partite ancora da giocare, non quelle già giocate).

Girolamo Cardano fornisce una differente soluzione intorno al 1539. La sua idea si avvicina al ragionamento probabilistico, infatti consiste nel suddividere la posta tenendo conto del numero di partite che ciascuno deve ancora giocare. Tuttavia la formula che egli fornisce non è corretta.

Anche Niccolò Tartaglia e Pietro Cataneo affrontano il problema della posta. Il primo fornisce una soluzione di poco migliore, ma comunque simile, a quella di Luca Pacioli; Cataneo invece si avvicina al ragionamento di Girolamo Cardano. In ogni caso, nessuno di loro arriva alla soluzione corretta e, in particolare, non si nota ancora un ragionamento di tipo probabilistico (come ad esempio quello utilizzato nella soluzione presentata qui sopra).

La risoluzione del problema della posta e la nascita della Probabilità (1654). Il Cavaliere di Méré, famoso giocatore d'azzardo, pone a Blaise Pascal il problema di ripartizione della posta in gioco. Tale problema diventa motivo di scambio epistolare tra Pascal e Fermat. In questo carteggio del 1654 si va via via definendo un ragionamento nuovo, di tipo probabilistico, che porta ad una risoluzione corretta del problema.

Christiaan Huygens. Qualche anno più tardi, nel 1657, il grande fisico olandese Christiaan Huygens, ispirandosi anche ai problemi studiati da Pascal e Fermat, pubblica il *De ratiociniis in ludo aleae*, considerato il primo trattato di Probabilità, dove appare il concetto di *valore atteso* (in latino *expectatio*), già menzionato nel carteggio tra Pascal e Fermat.

Il Settecento. Giacomo Bernoulli, il più anziano tra i matematici della famiglia Bernoulli, nasce il 27 dicembre del 1654, l'anno del carteggio tra Pascal e Fermat. Egli studia a fondo l'opera di Huygens. La sua opera *Ars conjectandi*, pubblicata postuma nel 1713, inizia con il trattato stesso di Huygens, a cui Bernoulli aggiunge commenti che, per ampiezza e profondità di contenuti concettuali, costituiscono un nuovo trattato intramezzato con quello di Huygens e di più ampia mole. Tale opera contiene anche il famoso teorema limite che oggi porta il suo nome ed è più spesso chiamato *Legge dei grandi numeri*; vengono inoltre riportate le cosiddette *prove di Bernoulli* e l'espressione della *distribuzione binomiale*.

Qualche anno più tardi, nel 1733, Abraham de Moivre, francese fuggito in Inghilterra per la repressione religiosa degli Ugonotti, formula una prima versione del *Teorema centrale del limite*, che si trova nel suo trattato *Doctrine of chances*. Tale teorema è riportato nel caso particolare di variabili aleatorie di Bernoulli simmetriche ($p = 1/2$). La dimostrazione di de Moivre utilizza la formula di Stirling, appena pubblicata nel 1730.

Il trattato di Abraham de Moivre sembra aver ispirato Thomas Bayes, che negli ultimi della sua vita si interessa molto di probabilità. Nell'articolo *An Essay towards solving a Problem in the Doctrine of Chances*, pubblicato postumo nel 1763, compare un caso speciale di quella che oggi è chiamata la *formula di Bayes*. Nonostante la prima trattazione generale su questo argomento sia dovuta a Laplace, il nome di Bayes è comunque legato indissolubilmente a questo risultato, che è alla base della *Statistica bayesiana*.

Pierre Simon de Laplace. Pierre Simon de Laplace è una delle figure principali della storia della Probabilità. Si è occupato di questa disciplina da tutti i punti vista (filosofico, statistico e matematico) e ha contribuito notevolmente nel fare in modo che venisse riconosciuta come una vera disciplina scientifica, evidenziandone le importanti potenzialità che vanno ben oltre i giochi d'azzardo.

Nel 1812 viene pubblicato il suo trattato principale in ambito probabilistico, intitolato *Théorie analitique des probabilités*, in cui si trova la dimostrazione del *Teorema centrale del limite* per p generico e viene riconosciuta la *distribuzione normale* come distribuzione limite. Laplace riscopre la *formula di Bayes* ed è grazie ai suoi lavori se tale formula diventa così importante in Probabilità e Statistica.

Riportiamo infine la seguente citazione estratta da *Essai philosophique sur les probabilités* (1814):

La teoria della probabilità è, in fondo, semplice buon senso tradotto in calcolo; ci fa valutare con esattezza ciò che una mente ragionevole sente per una sorta di istinto [...] *È degno di nota che questa scienza, nata a servizio dei giochi d'azzardo, sia diventata il più importante oggetto della conoscenza umana* [...] *Le più importanti questioni della vita sono, per la maggior parte, solo dei problemi di probabilità.*

La scuola russa. Nella seconda metà dell'Ottocento si sviluppò la scuola russa che restituì un rinnovato impianto scientifico al Calcolo delle probabilità con i contributi di studiosi quali Pfanuty Lvovich Chebyshev o del suo allievo Andrei Andreyevich Markov, che lavorò sino agli inizi del Novecento costruendo le basi della *Teoria delle catene di Markov*, un primo importante tassello della *Teoria dei processi stocastici.* Chebyshev introdusse per primo la nozione di *variabile aleatoria* e formulò il *Teorema centrale del limite* per la prima volta in termini moderni.

La nascita della Probabilità come disciplina matematica. Nonostante i progressi raggiunti, alla fine dell'Ottocento la Probabilità è ancora una disciplina poco rigorosa, costituita da problemi non solo matematici ma anche statistici e filosofici che si intrecciano tra loro. Fino a questo momento infatti non esiste una suddivisione tra *Calcolo delle probabilità*, *Statistica* e *Filosofia della probabilità.* Non stupisce allora che in occasione del II Congresso Internazionale dei Matematici che si tiene a Parigi nel 1900, David Hilbert, nella sua famosa enumerazione dei problemi ritenuti fondamentali per lo sviluppo della Matematica del XX secolo, pone al sesto posto la questione di rendere la Probabilità una disciplina matematica, ovvero, più precisamente, la questione di assiomatizzare la Probabilità.

In realtà, già nel 1854 il matematico e logico inglese George Boole auspica un trattamento assiomatico della Probabilità, anche se né lui né altri affrontano questo problema fino all'inizio del XX secolo. Solo agli inizi del Novecento, su impulso di Hilbert, si registrano vari tentativi di assiomatizzazione della Probabilità, tra cui ricordiamo: Georg Bohlmann (1900), Ugo Broggi (1907, studente di dottorato di Hilbert), Richard von Mises (1919), Bruno de Finetti (1931), Francesco Paolo Cantelli (1932).

Nel 1933, il matematico russo Andrej Nikolaevič Kolmogorov pubblica le sue ricerche sulla formulazione assiomatica della Probabilità in un'opera, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (*Foundations of the Theory of Probability*), destinata a lasciare un segno indelebile negli sviluppi successivi della disciplina. Grazie a quest'opera Kolmogorov risolve il sesto problema di Hilbert, nel senso che l'assiomatizzazione proposta da Kolmogorov è oggi universalmente accettata. Tuttavia, ciò non avviene senza resistenza, e la maggioranza dei probabilisti dell'epoca (in particolare negli Stati Uniti) per lungo tempo hanno considerato l'integrale di Lebesgue una perdita di tempo e un'offesa nei confronti dell'intuizione probabilistica. Tra gli oppositori all'assiomatizzazione proposta da Kolmogorov troviamo anche l'italiano Bruno de Finetti.

Nonostante ciò, possiamo dire che, a partire dagli anni Cinquanta del Novecento, i matematici hanno scelto definitivamente il loro modello assiomatico, lasciando ai filosofi il compito di mettere in relazione questo modello con la realtà. Sono infatti ormai finite le discussioni sui fondamenti della Probabilità che hanno caratterizzato la generazione precedente. La risoluzione della questione dei fondamenti permette finalmente un rigoglioso sviluppo della Probabilità come disciplina matematica, oltre che della Statistica e della Filosofia della probabilità.