3 Calcolo delle probabilità

3.1 Probabilità: un concetto intuitivo

La nostra vita di tutti i giorni è costellata da considerazioni di natura probabilistica. La possibilità che piova o meno nel corso della giornata (per decidere se prendere o no l'ombrello), le previsioni del tipo la squadra X ha ormai vinto al 90% il campionato, le speranze di vincita in giochi e nelle lotterie ne sono classici esempi. In tutte le situazioni di incertezza, si tende in sostanza a dare una *misura* dell'incertezza, che esprime il significato intuitivo della probabilità.

Tuttavia, l'affidarsi completamente all'intuizione può portare a conclusioni scorrette. La pubblicità su quotidiani riporta spesso mirabolanti promesse di vincita al gioco del Lotto. Se il gioco non è truccato (come dobbiamo credere fino a prova contraria), cioè se, ad ogni estrazione, ogni numero ha la stessa possibilità di essere estratto e il meccanismo di estrazione non è influenzato da quanto avvenuto nelle estrazioni precedenti, allora il fatto che un numero sia ritardatario non aumenta la sua probabilità di essere estratto.

Per evitare di pervenire a conclusioni scorrette, è pertanto necessario formalizzare il *calcolo delle probabilità* stabilendone le regole e i concetti in modo logico e rigoroso.

3.2 Breve storia del calcolo delle probabilità

Il gioco d'azzardo, nel quale si scommettono somme più o meno grandi, era già noto agli antichi romani e quello dei dadi sicuramente era il gioco più diffuso ed è stato per molti secoli il passatempo preferito. I giochi andavano dai dadi, alla morra, alle carte, alla *zara*, gioco menzionato persino da Dante nel VI Canto del Purgatorio.
Il concetto di probabilità nasce nel Rinascimento con lo studio dei codici segreti e si sviluppa in modo sistematico nel 17° secolo proprio con i giochi d'azzardo. I primi studi conosciuti su questioni di probabilità si riferiscono al gioco dei dadi e compaiono nel libro *Il gioco dei dadi* di Girolamo Cardano (1501-1576), figura poliedrica del Rinascimento: era medico, filosofo, astrologo e matematico.

Anche Galileo Galilei, qualche anno più tardi, intorno al 1630, si occupò di probabilità ed il suo contributo fu notevole nell'opera *Sulla scoperta dei dadi.* Ed è proprio dal gioco d'azzardo che ha inizio lo studio sistematico del calcolo delle probabilità nel Seicento, per risolvere alcuni problemi sui dadi posti da un accanito giocatore del tempo, il Cavalier de Méré, a Blaise Pascal, del quale rimane, sull'argomento, un carteggio datato 1654, con il matematico Pierre de Fermat.

Il Cavalier de Méré riscontrava che le sue deduzioni probabilistiche non si accordavano con le sue fortune, o meglio sfortune, di gioco e si rivolse a Pascal chiedendo lumi al riguardo. Questa la sua domanda: *esiste la stessa probabilità di vincere scommettendo che esca almeno un 6 su 4 tiri consecutivi, lanciando un dado alla volta, oppure scommettendo che escano almeno due 6 su 24 tiri, lanciando due dadi alla volta?* (secondo i suoi calcoli le probabilità erano uguali e pari a 2/3). Pascal e Fermat fornirono la risposta giusta giungendo alla conclusione che il doppio 6 su 24 lanci è un evento meno probabile a realizzarsi di un singolo 6 su 4 lanci (alla fine di questo capitolo vedremo la soluzione).

Però né Pascal né Fermat diedero una stesura sistematica ai risultati a cui erano pervenuti, che tuttavia ispirarono un grande scienziato olandese, Christian Huygens, che nel 1657 pubblicò un trattato: *Sui ragionamenti nel gioco dei dadi*.

Nel nostro excursus storico sulla probabilità arriviamo al 1713, dove incontriamo Jakob Bernoulli che formulò il primo enunciato della legge dei grandi numeri, oggi fondamentale per le scienze statistiche, nel libro pubblicato postumo *Ars Conjectandi*.

La teoria della probabilità conobbe un grande sviluppo nel XX secolo, quando Andrej Nikolaevič Kolmogorov, nel 1933, ne introdusse l'approccio assiomatico che ancora oggi ne costituisce il fondamento.

Possiamo dire che il calcolo delle probabilità è lo studio delle proprietà quantitative (come la frequenza) che possono essere osservate per quegli eventi il cui verificarsi o meno (in seguito ad osservazioni o prove) non è prevedibile in modo deterministico.
Tali eventi vengono detti *casuali* o *aleatori*.
Matematicamente la probabilità viene descritta mediante una quantità scalare che caratterizza la frequenza di ricorrenza di un dato evento al ripetersi delle prove.

3.3 La costruzione di un modello probabilistico

Abbiamo definito il calcolo delle probabilità come la teoria matematica dell'incertezza. La teoria dice come si deve formulare un modello probabilistico. Come ogni modello matematico, anche quello probabilistico, da un lato, consente la trattazione di un problema di interesse in modo logico e rigoroso; dall'altro, rappresenta necessariamente un'astrazione della realtà e ne cattura solo alcuni aspetti. Un aspetto affascinante del calcolo delle probabilità è la possibilità di affrontare problemi interessanti con una matematica sostanzialmente elementare anche se questo non vuole dire che il calcolo delle probabilità sia una disciplina facile. Esistono diverse teorie della probabilità.

Teoria classica o "a priori"

Se l'esito delle prove può essere descritto da un numero finito n di casi possibili, allora la probabilità p di uno di tali casi viene definita "a priori" come:

$$p = \frac{f}{n} \in [0, 1]$$

dove f rappresenta il numero dei casi favorevoli.

Questa è essenzialmente la definizione di Laplace: "La probabilità di un evento è il rapporto tra il numero di casi favorevoli ed il numero di casi possibili, quando questi sono tutti equiprobabili". Tale definizione risente delle circostanze in cui è nato il calcolo delle probabilità (in relazione a problemi di gioco d'azzardo).

Esempi

- Lancio di due dadi (non truccati).
  Questo esempio presenta $n = 6^2 = 36$ casi possibili, che possono essere rappresentati dagli elementi di una matrice 6×6:

$$\begin{pmatrix} (1,1) & \dots & (1,6) \\ \vdots & \ddots & \vdots \\ (6,1) & \dots & (6,6) \end{pmatrix}$$

  - probabilità che esca un doppio 6:

$$f = 1, \qquad p = \frac{1}{36}$$

  - probabilità di ottenere 5 dalla somma dei punteggi dei due dadi:

$$f = 4, \quad (1,4),(2,3),(3,2),(4,1) \quad p = \frac{4}{36} = \frac{1}{9}$$

- Estrazione di una carta da un mazzo di carte francesi.
  Questo esempio presenta $n = 52$ casi ugualmente possibili, mutuamente esclusivi.

  – probabilità che esca una carta di fiori:

  $$f = 13, \qquad p = \frac{13}{52} = \frac{1}{4}$$

  – probabilità che esca un asso:

  $$f = 4, \quad p = \frac{4}{52} = \frac{1}{13}$$

  – probabilità che esca un asso o una carta di fiori:

  $$f = 4 + 13 - \underbrace{\boxed{1}}_{\text{asso contato due volte}}, \quad p = \frac{16}{52} = \frac{4}{13}$$

  – probabilità che non esca nè un asso, nè una carta di fiori:

  $$f = 52 - 16 = 36\,, \quad p = \frac{36}{52} = \frac{9}{13}$$

Difetti

- Occorre supporre che gli eventi possibili siano in numero finito.
- Occorre supporre che gli eventi siano mutuamente esclusivi o incompatibili.
- Occorre supporre che gli eventi siano tutti ugualmente probabili (equiprobabili).

Il primo problema si supera introducendo la definizione di probabilità geometrica, che rappresenta un'estensione di quella classica.
Il secondo e il terzo possono essere superati passando alla teoria empirica o frequentista.

Teoria empirica o frequentista

Può essere empiricamente riscontrato che, se si osservano gli esiti di successive repliche di un esperimento, la frequenza relativa di un evento associato a tali esiti (ad esempio l'uscita di una testa nel lancio di una moneta o, nel lancio di due dadi, l'uscita di un doppio sei), tenderà a stabilizzarsi su un certo valore che sarà pari a $1/2$ nel lancio della moneta e a $1/36$ nel lancio dei due dadi. La coincidenza di tale valore con il valore calcolato secondo la definizione classica di probabilità ha portato a definire la probabilità di un evento come il limite della frequenza relativa del verificarsi dell'evento quando il numero delle prove tende all'infinito.

È necessario concepire una serie di esperimenti o prove che avvengano tutte in condizioni "abbastanza" uniformi e in tal caso:

$$p = \lim_{n\to\infty} \frac{n_A}{n},$$

dove n è il numero delle prove, n_A è il numero delle volte che si verifica un certo evento A.
Si noti che non bisogna specificare nè l'equiprobabilità nè l'incompatibilità degli eventi.

Difetti

- Si applica ad esperimenti ripetibili e per i quali il limite per $n \to \infty$ abbia senso.

Teoria soggettiva

Secondo la definizione soggettiva di probabilità, dovuta all'italiano Bruno De Finetti, la probabilità è il prezzo che un individuo *coerente* ritiene equo pagare per ricevere 1 se l'evento si verifica e 0 altrimenti.

La definizione soggettiva consente di calcolare la probabilità di eventi anche quando questi non sono equiprobabili e quando l'esperimento non può essere ripetuto.

Difetto

- Tale definizione rimane fondata sull'opinione di singoli individui, che potrebbero presentare diverse propensioni al rischio. Basta pensare che molti sarebbero disposti a giocare 1 euro per vincerne 1000, ma pochi giocherebbero un milione di euro per vincerne un miliardo.

Pertanto, al di là delle diverse interpretazioni della probabilità, è possibile formulare una teoria che dica come costruire modelli probabilistici e analizzarne le implicazioni in modo rigoroso. Poichè le regole e i metodi di calcolo nelle diverse teorie esaminate non differiscono tra loro, è possibile seguire l'impostazione dovuta ad A.N. Kolmogorov, fondatore della Teoria assiomatica della probabilità. Il linguaggio utilizzato è quello della *teoria degli insiemi*.

3.4 Spazio campione ed eventi

Definizione: Lo spazio campione, indicato con Ω, è la totalità di tutti i possibili risultati di un esperimento concettuale.
Se lo spazio campione ha un numero finito di elementi, tali elementi si possono elencare separati da una virgola e racchiudere tra parentesi graffe $\{\ ,\ \}$.
Se lo spazio campione ha un numero infinito di elementi, può essere descritto tramite un'affermazione o una regola. Ad esempio:

$$\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$$

Definizione: Un evento è un sottoinsieme dello spazio campione e si indica con le lettere maiuscole A, B, C,...

Definizione: Ω è detto evento certo.

Definizione: $\varnothing$ è detto evento impossibile.

Definizione: Il complementare di un evento A rispetto ad Ω è il sottoinsieme di tutti gli elementi di Ω che non sono contenuti in A e viene indicato con $\overline{A}$.

Esempi

- Lancio di tre monete.
  $\Omega = \{\mathrm{TTT}, \mathrm{TTC}, \mathrm{TCT}, \mathrm{CTT}, \mathrm{TCC}, \mathrm{CTC}, \mathrm{CCT}, \mathrm{CCC}\}$
  "Escono due teste e una croce" è un possibile evento $\rightarrow A = \{\mathrm{TTC}, \mathrm{TCT}, \mathrm{CTT}\}$

- Tempo di vita di una lampadina.
  $\Omega = \{t : t \geq 0\}$, con t ad esempio misurato in ore.
  "La lampadina si brucia prima di 300 ore è un evento" $\rightarrow A = \{t : 0 \leq t < 300\}$

Definizione: Se $A, B \subset \Omega$ sono due eventi dello stesso spazio campione, l'intersezione di A e B, $A \cap B$, è l'evento che contiene tutti gli elementi comuni sia ad A che a B.

Definizione: Se $A, B \subset \Omega$ sono due eventi dello stesso spazio campione che non hanno elementi in comune, cioè $A \cap B = \varnothing$, i due eventi A e B si dicono disgiunti o mutuamente esclusivi o incompatibili.

Definizione: Se $A, B \subset \Omega$ sono due eventi dello stesso spazio campione, l'unione di A e B, $A \cup B$, è l'evento che contiene tutti gli elementi che appartengono ad A o a B.

La relazione tra gli eventi e il corrispondente spazio campione possono essere illustrati graficamente attraverso i *diagrammi di Venn*, dove Ω è rappresentato da un rettangolo e gli eventi da curve chiuse in Ω.

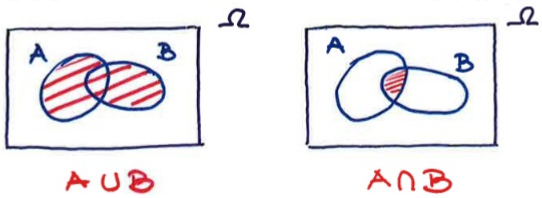

Figura 8: Diagrammi di Venn

3.5 Probabilità di un evento

Non siamo interessati agli eventi, ma alla probabilità che uno di questi eventi si verifichi o meno.
L'impostazione assiomatica parte dal concetto di $\sigma - algebra$ o classe additiva.
La probabilità viene vista come una misura, cioè come una funzione che associa ad ogni sottoinsieme di Ω un numero reale non negativo, tale che la somma delle probabilità di tutti gli eventi sia uguale ad 1.
Se la cardinalità di Ω è finita, diciamo $\text{card}(\Omega) = n$,

l'insieme di tutti i suoi sottoinsiemi, detto insieme delle parti, ha cardinalità 2^n.
Se Ω ha la cardinalità del continuo, il suo insieme delle parti è "troppo grande" perchè su di esso si possa definire una misura.
Si considerano perciò i soli sottoinsiemi di Ω che costituiscono un insieme non vuoto $\mathcal{A}$ (classe additiva) tale che:

- $A \in \mathcal{A} \to \overline{A} \in \mathcal{A}$
- $\forall i \in \mathbb{N}, A_i \in \mathcal{A} \to \bigcup\limits_{i \in \mathbb{N}} A_i \in \mathcal{A}$

Una classe additiva è quindi un sottoinsieme dell'insieme delle parti di Ω che risulta chiuso rispetto alle operazioni di complemento e di unione numerabile.
Inoltre, per le leggi di De Morgan

- $\forall i \in \mathbb{N}, A_i \in \mathcal{A} \to \bigcap\limits_{i \in \mathbb{N}} A_i = \overline{\bigcup\limits_{i \in \mathbb{N}} \overline{A}_i} \in \mathcal{A}$

Assiomi della probabilità

Definizione: Dati uno spazio campione Ω e una classe additiva (σ-algebra) $\mathcal{A}$ di eventi su Ω, una funzione

$$P : \mathcal{A} \to [0, 1]$$

è detta funzione di probabilità se valgono i seguenti assiomi:

1. $\forall A \in \mathcal{A} \quad P[A] \geq 0$
2. $P[\Omega] = 1$
3. Per ogni successione $A_1, ..., A_n, ... \in \mathcal{A}$ di eventi a due a due disgiunti (che quindi verificano $A_i \cap A_j = \varnothing \; \forall i, j \;\; i \neq j$) e tali che $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$ si ha:

$$P[\bigcup_{i \in \mathbb{N}} A_i] = \sum_{i \in \mathbb{N}} P[A_i]$$

La funzione P è una funzione d'insieme perchè gli elementi della funzione sono insiemi di punti anzichè punti singoli. La terna $(\Omega, \mathcal{A}, P)$ è detta *spazio di probabilità*.
Come conseguenza degli assiomi è possibile verificare le seguenti

Proprietà di $P[\,\cdot\,]$

1. $P[\varnothing] = 0$
2. se $A_1, ..., A_n \in \mathcal{A}$, a due a due disgiunti, allora $P[\bigcup_{i=1}^{n} A_i] = \sum_{i=1}^{n} P[A_i]$

3. $\forall A \in \mathcal{A} \quad P[\,\overline{A}\,] = 1 - P[A]$
   (conseguenza di $\Omega = A \cup \overline{A},\ A \cap \overline{A} = \varnothing$)

4. Se $A, B \in \mathcal{A}$

$$P[A] = P[A \cap B] + P[A \cap \overline{B}]$$

$$P[A - B] = P[A \cap \overline{B}] = P[A] - P[A \cap B]$$

   (conseguenza di $A = (A \cap B) \cup (A \cap \overline{B})$, $(A \cap B) \cap (A \cap \overline{B}) = \varnothing$)

5. Se $A, B \in \mathcal{A}$ e $A \subseteq B$,

$$P[A] \leq P[B]$$

6. Se $A, B \in \mathcal{A}$

$$\underline{P[A \cup B] = P[A] + P[B] - P[A \cap B]}$$

   regola di addizione

7. $A_1, ..., A_n \in \mathcal{A}$

$$P[\bigcup_{i=1}^{n} A_i] \leq \sum_{i=1}^{n} P[A_i]$$

Questa definizione assiomatica di probabilità ci dice quali funzioni di insieme sono accettabili come funzioni di probabilità, ma non ci dice quali valori la funzione $P[\cdot]$ attribuisce ad un dato evento.

3.6 Spazi campionari finiti

Definizione: Sia $\Omega = \{e_1, e_2, ..., e_n, ...\}$ uno spazio campione. Un qualsiasi sottoinsieme $A_i = \{e_i\}_{i \in \mathbb{N}}$ è chiamato evento semplice o elementare.
Se Ω è costituito da un numero finito di elementi distinti, cioè $\Omega = \{e_1, ..., e_n\}$ è un insieme finito, allora:

$\Omega = \bigcup_{i=1}^{n} \{e_i\}$, $\{e_i\} \cap \{e_j\} = \varnothing \quad \forall i \neq j$, e si ha:

- $\forall A \in \mathcal{A} \quad P[A] = \sum_{e_i \in A} P[\{e_i\}]$
- $P[\{e_i\}] \doteq p_i \, , \, i = 1, ..., n \; : \; \sum_{i=1}^{n} p_i = 1$

Se i punti dello spazio campionario sono anche equiprobabili allora:

- $P[\{e_i\}] = p_i = \dfrac{1}{n}, \quad i = 1, ..., n$
- $\forall A \in \mathcal{A} \quad P[A] = \dfrac{\text{card}(A)}{\text{card}(\Omega)} = \dfrac{\text{card}(A)}{n}$

e la funzione di probabilità P è detta uniforme.

Esempi

- Lancio di due dadi (non truccati).
  $\Omega = \{(i, j) : i = 1, ..., 6; j = 1, ..., 6\}$,
  $\text{card}(\Omega) = 6^2 = 36$,
  $\{e_k\} = (i, j) =$ evento elementare.
  $P[\{e_k\}] = 1/36$.
  Scegliamo l'evento "esce 7 come punteggio":
  $A_7 = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$
  $\text{card}(A_7) = 6$.
  $$P[A_7] = \frac{\text{card}(A_7)}{\text{card}(\Omega)} = \frac{6}{36} = \frac{1}{6}$$
  $$\begin{pmatrix} 11 & 12 & 13 & 14 & 15 & 16 \\ 21 & \cdot & & & 25 & \cdot \\ 31 & \cdot & & 34 & & \cdot \\ 41 & & 43 & & & \cdot \\ 51 & 52 & & & & \cdot \\ 61 & & & & & \cdot \end{pmatrix}$$

- Lancio di un dado truccato.
  Supponiamo che la probabilità che esca la faccia j, $j = 1, ..., 6$ sia direttamente proporzionale al numero j della faccia. Qual è la probabilità di ottenere una faccia con il numero pari?

Sia $P[j] = \alpha j$ dove α è il coefficiente di proporzionalità. Bisogna determinare il valore di α.

$$\sum_{j=1}^{6} P[j] = \sum_{j=1}^{6} \alpha j = \alpha \frac{6 \cdot 7}{2} = 21\alpha,$$

(ricorda che $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$)

e poichè $\sum_{j=1}^{6} P[j] = 1$, $\alpha = 1/21$.

$$P[A_2] + P[A_4] + P[A_6] = \frac{2}{21} + \frac{4}{21} + \frac{6}{21} = \frac{4}{7}$$

La probabilità di ottenere una faccia con un numero dispari è $3/7$.

Possiamo dire che nel caso di spazi campionari finiti, il calcolo della probabilità di un evento si riduce ad un problema di conteggio del numero degli elementi dell'evento.

I problemi del Cavalier De Méré

1. Trovare il più piccolo numero intero n tale che lanciando n volte 1 dado, la probabilità di avere almeno un 6 sia maggiore di $1/2$.

2. Trovare il più piccolo numero intero m tale che lanciando m volte 2 dadi, la probabilità di avere almeno un $(6,6)$ sia maggiore di $1/2$.

1. In un singolo lancio, chiamiamo "A" l'evento "fare 6" e "B" l'evento "fare almeno un 6 in n lanci". Abbiamo:

$$P[A] = \frac{1}{6}, \quad P[\overline{A}] = \frac{5}{6}.$$

La strategia è quella di calcolare la probabilità dell'evento $\overline{B}$ = "nessun 6 in n lanci", in modo poi da poter calcolare $P[B] = 1 - P[\overline{B}]$.

$$P[\overline{B}] = \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdots \frac{5}{6}}_{\text{n volte}} = \left(\frac{5}{6}\right)^n \Rightarrow P[B] = 1 - \left(\frac{5}{6}\right)^n$$

Il problema si è ridotto allora a trovare il più piccolo numero intero n tale che $1 - \left(\frac{5}{6}\right)^n > \frac{1}{2}$.

$$1 - \left(\frac{5}{6}\right)^n > \frac{1}{2} \Rightarrow \left(\frac{5}{6}\right)^n < \frac{1}{2}$$

$$n = 1 \qquad \left(\frac{5}{6}\right)^1 = 0.8\overline{3}$$

$$n = 2 \quad \left(\frac{5}{6}\right)^2 \sim 0.694$$

$$n = 3 \quad \left(\frac{5}{6}\right)^3 \sim 0.578$$

$$n = 4 \quad \left(\frac{5}{6}\right)^4 \sim 0.482 \;\Rightarrow P[B]|_{n=4} \sim 0.518$$

2. In un lancio di due dadi, chiamiamo "A" l'evento "fare (6,6)" e "B" l'evento "fare almeno un (6,6) in m lanci". Quindi $\overline{B}$ è l'evento "ottenere nessun (6,6) in m lanci ". Si ha:

$$P[\overline{B}] = \left(\frac{35}{36}\right)^m \Rightarrow P[B] = 1 - \left(\frac{35}{36}\right)^m$$

Bisogna trovare il più piccolo intero m tale che $\left(\frac{35}{36}\right)^m < \frac{1}{2}$.

$$\vdots$$

$$m = 24 \quad \left(\frac{35}{36}\right)^{24} \sim 0.508596$$

$$m = 25 \quad \left(\frac{35}{36}\right)^{25} \sim 0.494468$$

$$\Rightarrow P[B]|_{m=25} \sim 0.50553$$

3.7 Probabilità condizionata e indipendenza

Il concetto di probabilità condizionata è molto importante poichè in molte applicazioni, alla struttura probabilistica vengono imposte determinate restrizioni. Tale concetto permette di rivalutare l'idea di probabilità di un evento alla luce di nuove informazioni, cioè quando un altro evento si è verificato.
Perciò indichiamo col simbolo $P[A|B]$ la probabilità che A si verifichi sapendo che B si è verificato, leggendo semplicemente "probabilità di A dato B".

Definizione: Dati due eventi $A, B \in \mathcal{A}$, definiamo probabilità condizionata di A dato B, cioè $P[A|B]$, la quantità scalare:

$$P[A|B] = \begin{cases} \frac{P[A \cap B]}{P[B]} & \text{se } P[B] > 0\,, \\ 0 & \text{altrove.} \end{cases}$$

Nota bene: la probabilità condizionata $P[\,\cdot\,|B]$ con $P[B] > 0$ è una funzione di probabilità poichè verifica i 3 assiomi e pertanto per essa valgono le proprietà della probabilità già enunciate.

Esempi

- Lancio di due dadi.
  Calcolare la probabilità di ottenere 9, ammesso che su un dado esca 3.
  Chiamiamo:
  "A"= "ottenere 9" $= \{(3,6);(6,3);(4,5);(5,4)\}$
  "B"= "su un dado esce 3" $= \{(1,3);(3,1);(2,3);(3,2);(3,3);(3,4);(4,3);(3,5);(5,3);(3,6);(6,3)\}$

$$\text{card}(\Omega) = 36,\ \text{card}(A) = 4,\ \text{card}(B) = 11;$$

$$A \cap B = \{(3,6);(6,3)\} \Rightarrow \text{card}(A \cap B) = 2,$$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{\frac{\text{card}(A\cap B)}{\text{card}(\Omega)}}{\frac{\text{card}(B)}{\text{card}(\Omega)}} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}$$

Osservazioni
1) Non confondere la nozione di probabilità condizionata con le prove ripetute dello stesso esperimento. Ad esempio, se lanciamo due volte un dado supponiamo che esca il numero 3 al primo lancio. Qual è la probabilità di "ottenere 9" con il secondo lancio? (cioè esce 6 al secondo lancio). La probabilità è $\frac{1}{6}$.
2) Se lo spazio campione Ω è finito e la probabilità è

uniforme (punti equiprobabili), allora:

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{\text{card}(A \cap B)}{\text{card}(B)}, \quad (B \neq \varnothing)$$

3) Dato lo spazio campione Ω e dati due eventi A e $B \in \mathcal{A}$, poichè

$$A = (A \cap B) \cup (A \cap \overline{B}) \quad \text{e} \quad (A \cap B) \cap (A \cap \overline{B}) = \varnothing$$

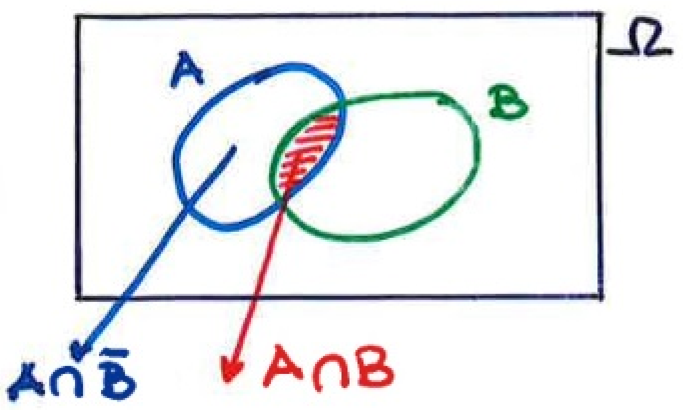

Figura 9

allora

$$P[A] = P[A \cap B] + P[A \cap \overline{B}]$$

ed anche

$$P[A] = P[A|B]P[B] + P[A|\overline{B}](1 - P[B])$$

- Lancio di un dado truccato.
  Il dado è stato modificato in modo che i numeri pari abbiano una probabilità doppia di uscire rispetto ai numeri dispari.
  Sia A l'evento di ottenere un quadrato perfetto nel lancio di un dado.
  $A = \{1, 4\}$
  Se indico con p la probabilità di ottenere un numero dispari, allora:
  $P[\text{dispari}] = P[1] + P[3] + P[5] = 3p$
  La probabilità di ottenere un numero pari è doppia
  $P[\text{pari}] = P[2] + P[4] + P[6] = 3 \cdot 2p = 6p$.
  Poichè $\Omega = D \cup P$ e $P[\Omega] = 1 \rightarrow$
  $\rightarrow 3p + 6p = 1 \rightarrow p = \frac{1}{9}$.
  Quindi ogni numero dispari ha probabilità $\frac{1}{9}$ ed ogni numero pari $\frac{2}{9}$.

$$P[A] = \frac{1}{9} + \frac{2}{9} = \frac{3}{9} = \frac{1}{3}$$

Ora supponiamo che nel lancio del dado sia uscito un numero maggiore di 3. Quanto vale ora la probabilità dell'evento A?

$$A = \{4\}\,,\; B = \{4;5;6\}\,,\;\; \{A \cap B\} = \{4\},$$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]}{P[B]} = \frac{\frac{2}{9}}{\frac{2+1+2}{9}} = \frac{2}{5}.$$

Il teorema delle probabilità totali

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e $B_1, ..., B_n$ una collezione finita di eventi di $\mathcal{A}$, a due a due incompatibili, tali che:

- $P[B_i] > 0\,, \quad i = 1, ..., n$
- $\Omega = \bigcup\limits_{i=1}^{n} B_i$

Allora $\forall A \in \mathcal{A}$ si ha:

$$P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i]$$

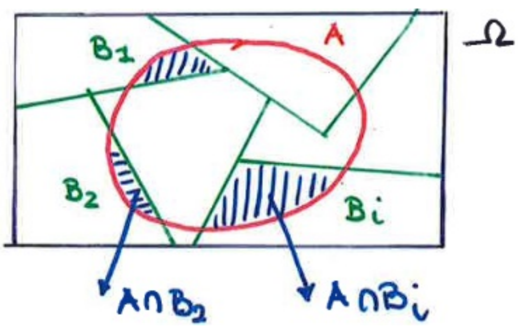

Figura 10

Prima formula di Bayes

Una forma equivalente della definizione di probabilità condizionata è la seguente espressione chiamata regola di moltiplicazione:

$$P[A \cap B] = P[A|B] \cdot P[B], \quad P[B] > 0$$

Poichè $P[A \cap B] = P[B \cap A]$, vale anche

$$P[B \cap A] = P[B|A] \cdot P[A], \quad P[A] > 0$$

ed uguagliando le due espressioni si ottiene:

$$P[A|B] = \frac{P[B|A] \cdot P[A]}{P[B]}$$

Esempi

1. Un sacchetto contiene 4 palline bianche e 3 palline nere. Un altro sacchetto contiene 3 palline bianche e 5 palline nere. Dal primo sacchetto si estrae una pallina e la si inserisce, senza guardarla, nel secondo sacchetto.
   Qual è la probabilità che una pallina estratta dal secondo sacchetto sia nera?

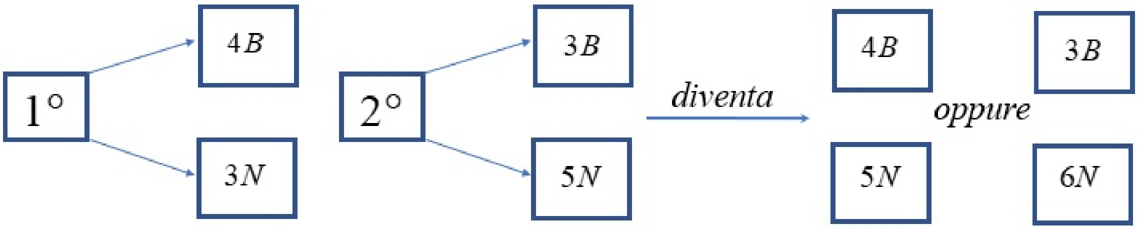

Figura 11

Siano N_i/B_i gli eventi "si estrae una pallina nera/bianca dall'i-esimo sacchetto". Utilizzando il teorema delle probabilità totali, abbiamo:

$$P[N_2] = P[N_2|N_1]P[N_1] + P[N_2|B_1]P[B_1],$$

e quindi

$$P[N_2] = \frac{6}{9} \cdot \frac{3}{7} + \frac{5}{9} \cdot \frac{4}{7} = \frac{18 + 20}{63} = \frac{38}{63}.$$

2. In un impianto, 3 macchine B_1, B_2 e B_3 assemblano una quantità di prodotti pari rispettivamente al 30%, 45% e 25% del totale.
   È noto che le macchine producono pezzi difettosi rispettivamente nella percentuale del 2%, 3% e 2%.
   Scelto a caso un pezzo, calcolare la probabilità che esso sia difettoso.

Sia "A" l'evento "il pezzo è difettoso" e sia B_i, $i = \{1; 2; 3\}$, l'evento "il pezzo è prodotto dalla macchina B_i". I dati del problema ci dicono che:

$$P[B_1] = 0.3,\ P[B_2] = 0.45,\ P[B_3] = 0.25;$$

$$P[A|B_1] = 0.02,\ P[A|B_2] = 0.03,\ P[A|B_3] = 0.02.$$

Quindi, per il teorema delle probabilità totali:

$$P[A] = \sum_{i=1}^{3} P[A|B_i]P[B_i] = 0.0245$$

Nota bene: sottintesa c'è l'ipotesi di indipendenza nella produzione dei pezzi difettosi, per ogni macchina.

3. Un segnale può essere trasmesso da due canali, A e B, con la stessa probabilità .
Il canale A trasmette sempre correttamente.
Il canale B trasmette correttamente con probabilità $\frac{3}{4}$. Si chiede:
   - Qual è la probabilità di ricevere un segnale corretto?
   - Avendo ricevuto un segnale corretto, qual è la probabilità che esso provenga dal canale B?

Chiamiamo "A"/ "B" l'evento "il segnale proviene dal canale A/B "
Chiamiamo C l'evento "il segnale è corretto".

- Per il teorema delle probabilità totali:

$$P[C] = P[C|A]P[A] + P[C|B]P[B] =$$
$$= 1 \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} = \frac{7}{8}$$

- Dalla 1^a formula di Bayes:

$$P[B|C] = \frac{P[C|B]P[B]}{P[C]} = \frac{\frac{3}{4} \cdot \frac{1}{2}}{\frac{7}{8}} = \frac{3}{7}.$$

Nota bene: c'è equiprobabilità nella scelta del canale, cioè $P[A] = P[B] = \frac{1}{2}$.

Data una serie di eventi che ricoprono tutto lo spazio campione (ricoprimento finito) e dato un evento A, ci si può chiedere quale sia la probabilità condizionata di ognuno di questi eventi del ricoprimento, sapendo che si è verificato l'evento A. Non è altro che una generalizzazione della 1^a formula di Bayes.

Il teorema di Bayes

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e $B_1, ..., B_n$ una collezione finita di eventi di $\mathcal{A}$, a due a due incompatibili, tali che:

- $P[B_i] > 0\,, \quad i = 1, ..., n$
- $\Omega = \bigcup_{i=1}^{n} B_i$

Allora $\forall A \in \mathcal{A}$ con $P[A] > 0$ si ha:

$$P[B_i|A] = \frac{P[A|B_i]P[B_i]}{\sum_{k=1}^{n} P[A|B_k]P[B_k]}$$

Esempio

Sono date 5 urne numerate contenenti ciascuna 10 palline. Dentro la i-esima urna ci sono i palline difettose. Scelta un'urna a caso ed estratta una pallina,

- calcolare la probabilità che la pallina sia difettosa.
- estratta la pallina e visto che è difettosa, calcolare la probabilità che essa provenga dall'urna numero 5.

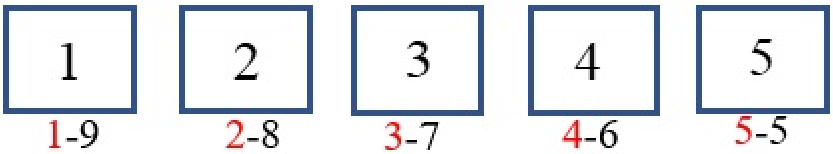

Figura 12

Con la definizione di probabilità a priori $p = \frac{f}{n}$ si ha immediatamente $p = \frac{15}{50} = \frac{3}{10}$.

Vogliamo applicare il teorema delle probabilità totali. Chiamiamo:

"A" l'evento "la pallina è difettosa".

"B_i" l'evento "la pallina viene scelta dalla i-esima urna".

Ogni urna ha la stessa probabilità di essere scelta, quindi

$$P[B_i] = \frac{1}{5}, \text{ i=1,...,5} \qquad P[A|B_i] = \frac{i}{10}$$

$$P[A] = \sum_{i=1}^{5} P[A|B_i]P[B_i] = \frac{1}{5} \cdot \sum_{i=1}^{5} \frac{i}{10} = \frac{3}{10}$$

Per trovare $P[B_5|A]$ applichiamo il teorema di Bayes:

$$P[B_5|A] = \frac{P[A|B_5]P[B_5]}{P[A]} = \frac{\frac{5}{10} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{1}{3}$$

In generale avremo

$$P[B_k|A] = \frac{P[A|B_k]P[B_k]}{P[A]} = \frac{\frac{k}{10} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{k}{15}\,, k = 1, .., 5$$

Quindi in termini non condizionati tutti i B_i sono equiprobabili ($P[B_i] = \frac{1}{5},\ i = 1, ..., 5$), condizionati al verificarsi dell'evento A non lo sono più poichè $P[B_i|A] = \frac{i}{15}$.

Osservazioni

1) Il teorema delle probabilità totali ed il teorema di Bayes sono tra le regole più importanti della teoria della probabilità . Si applicano in quegli esperimenti cosiddetti a più fasi.

2) È possibile estendere la regola di moltiplicazione enunciata per due eventi a situazioni in cui sono presenti più di due eventi.

Regola del prodotto

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e $B_1, ..., B_n$ una collezione finita di eventi di $\mathcal{A}$ tali che:

$$P[B_1 \cap B_2 \cap \ldots \cap B_{n-1}] > 0.$$

Allora vale:

$$P[B_1 \cap \ldots \cap B_n] = P[B_1]P[B_2|B_1]P[B_3|B_1 \cap B_2] \cdots$$
$$\cdots P[B_n|B_1 \cap \ldots \cap B_{n-1}]$$

Esempio

Un'urna contiene 10 palline, di cui 3 nere e 7 bianche. Estratta una pallina e registratone il colore, la si rimette nell'urna aggiungendone altre due dello stesso colore. Calcolare la probabilità di estrarre una pallina nera in ognuno dei primi tre tentativi.

Chiamiamo B_i l'evento "viene estratta una pallina nera all'i-esimo tentativo". Per la regola del prodotto

$$P[B_1 \cap B_2 \cap B_3] = P[B_1]P[B_2|B_1]P[B_3|B_1 \cap B_2]$$

$P[B_1] = \frac{3}{10}$, se aggiungo 2 palline nere $P[B_2|B_1] = \frac{5}{12}$, aggiungo altre due palline nere, $P[B_3|B_1 \cap B_2] = \frac{7}{14} = \frac{1}{2}$. Quindi

$$P[B_1 \cap B_2 \cap B_3] = \frac{3}{10} \cdot \frac{5}{12} \cdot \frac{1}{2} = \frac{1}{16}.$$

Sebbene la probabilità condizionata permetta di modificare la probabilità di un evento alla luce di nuove informazioni, è molto importante anche il concetto di indipendenza tra eventi.

Dati due eventi A, B se il verificarsi di B non ha alcun impatto sulla probabilità di A, allora A e B sono indipendenti.

Definizione: Due eventi A e B si dicono indipendenti se

$$P[A \cap B] = P[A]P[B]$$

Le seguenti condizioni sono equivalenti:

1. $P[A \cap B] = P[A]P[B]$
2. Se $P[B] > 0 \Rightarrow P[A|B] = P[A]$
3. Se $P[A] > 0 \Rightarrow P[B|A] = P[B]$

Se A e B sono indipendenti allora sono indipendenti A e $\overline{B}$, $\overline{A}$ e B, $\overline{A}$ e $\overline{B}$.

Esempio Lancio di due dadi.

Chiamiamo:

A l'evento "ottenere un numero dispari come punteggio".
B l'evento "ottenere 1 sul primo dado".
C l'evento "ottenere 7 come punteggio".
Stabilire se A e B, A e C, B e C sono indipendenti.

$\Omega = \{(i, j) : i = 1, ..., 6; j = 1, ..., 6\}$, $\quad \mathrm{card}(\Omega) = 36.$

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

$\text{card}(A) = 18$
$B = \{(1,1); (1,2); (1,3); (1,4); (1,5); (1,6)\}$
$\text{card}(B) = 6$
$C = \{(1,6); (6,1); (2,5); (5,2); (3,4); (4,3)\}$
$\text{card}(C) = 6$

$$P[A] = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{18}{36} = \frac{1}{2}$$

$$P[B] = \frac{6}{36} = \frac{1}{6}, \quad P[C] = \frac{6}{36} = \frac{1}{6}$$

$A \cap B = \{(1,2); (1,4); (1,6)\} \quad \text{card}(A \cap B) = 3$
$A \cap C = C \quad \text{card}(A \cap C) = 6$
$B \cap C = \{(1,6)\} \quad \text{card}(B \cap C) = 1$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{\text{card}(A \cap B)}{\text{card}(B)} = \frac{1}{2} = P[A]$$

$\Rightarrow$ A e B sono indipendenti.

$$P[A|C] = \frac{P[A \cap C]}{P[C]} = \frac{\text{card}(A \cap C)}{\text{card}(C)} = 1 \neq P[A]$$

$\Rightarrow$ A e C sono dipendenti.

$$P[B|C] = \frac{P[B \cap C]}{P[C]} = \frac{\text{card}(B \cap C)}{\text{card}(C)} = \frac{1}{6} = P[B]$$

$\Rightarrow B$ e C sono indipendenti.

È possibile generalizzare la nozione di eventi indipendenti al caso di più eventi.

Definizione: Dati n eventi $A_1, ..., A_n$, essi si dicono indipendenti se

$$P[A_i \cap A_j] = P[A_i]P[A_j], \quad i \neq j$$
$$P[A_i \cap A_j \cap A_k] = P[A_i]P[A_j]P[A_k] \; i \neq j \neq k \neq i$$
$$\vdots$$
$$\vdots$$
$$P[\bigcap_{i=1}^{n} A_i] = \prod_{i=1}^{n} P[A_i]$$

Osservazione

Eventi incompatibili (o disgiunti o mutuamente esclusivi) non sono eventi indipendenti.

Infatti dati A, B incompatibili $\Rightarrow A \cap B = \varnothing \Rightarrow P[A \cap B] = 0$. Se A e B fossero indipendenti allora avremmo $P[A \cap B] = P[A]P[B]$ e le due espressioni coincidono se e solo se $P[A]P[B] = 0$, cioè una fra $P[A]$ e $P[B]$, o entrambe, sono nulle, e quindi se uno fra A e B, o entrambi, è l'evento impossibile.