CAPITOLO 2

Probabilità

In probabilità si considera un fenomeno osservabile esclusivamente dal punto di vista della possibilità o meno del suo verificarsi, prescindendo dalla sua natura. Tra due estremi, detti evento certo (ad esempio: lanciando un dado si ottiene un numero compreso tra 1 e 6) ed evento impossibile (ottenere 1 come somma dal lancio di due dadi), si collocano eventi più o meno probabili.

Gli oggetti della probabilità sono gli esperimenti casuali o aleatori. Ciò che contraddistingue gli esperimenti casuali, rispetto agli esperimenti scientifici, è che per loro natura non sono ripetibili. Se lascio cadere un corpo sferico di massa nota da un'altezza assegnata in ambiente neutro, so che esso cadrà per effetto della accelerazione di gravità e dunque posso prevedere con esattezza con quale velocità toccherà il suolo. Ogni volta che questo esperimento scientifico verrà effettuato, si otterrà lo stesso risultato. Se invece lascio cadere un dado, non posso prevedere con esattezza quale faccia mostrerà, e in generale lanci diversi (ma effettuati nelle stesse condizioni) daranno risultati diversi. Si parla in questo caso di esperimento casuale: non possiamo conoscere in anticipo il risultato dell'esperimento, ma sappiamo che, se effettuiamo l'esperimento molte volte, tutti e sei i possibili risultati si presenteranno con la stessa frequenza. Sappiamo prevedere l'andamento medio dell'esperimento, ma non il singolo esito.

1. Definizioni

Precisiamo il vocabolario che useremo d'ora in poi.

Definizione 1.1

Useremo indistintamente le parole **esperimento casuale**, **esperimento aleatorio** *e* **prova**. *Il risultato di un tale esperimento è detto* **esito** *o* **punto campione**. *L'insieme di tutti i possibili esiti di un esperimento aleatorio si dice* **spazio campione** *e viene solitamente indicato con la lettera greca* Ω.
Un **evento** *è un qualunque sottoinsieme dello spazio campione. Dopo avere effettuato una prova, diremo che l'evento si è* **verificato** *se l'esito appartiene all'evento.*

Esempio 1.1: Lancio di un dado

Il lancio di un dado a sei facce è un esperimento aleatorio. Gli esiti sono i numeri da 1 a 6, dunque $\Omega = \{1, 2, 3, 4, 5, 6\}$. *Sono esempi di eventi*

A: uscita di un numero pari. $A = \{2, 4, 6\}$.
B: uscita del numero 6. $B = \{6\}$.
C: uscita di un numero inferiore a 3. $C = \{1, 2\}$.
D: uscita di un numero negativo. $D = \emptyset$.
E: uscita di un numero compreso fra 1 e 6. $E = \{1, 2, 3, 4, 5, 6\}$.

Se lanciamo il dado ed esce 2, allora gli eventi A, C *ed* E *sono verificati, mentre gli eventi* B *e* D *non sono verificati.*

Esempio 1.2: Lancio di una moneta

Anche il lancio di una moneta è un esperimento aleatorio, i cui possibili esiti sono $T =$"testa" *e* $C =$"croce". *Dunque lo spazio campione è l'insieme* $\{T, C\}$.

Esempio 1.3: Numero di iscritti

Possiamo interpretare il numero delle immatricolazioni ad informatica come un esperimento aleatorio il cui esito è un numero naturale, cioè $\Omega = \mathbb{N}$. Sono esempi di eventi

A: nessun nuovo iscritto. $A = \{0\}$.
B: meno di 100 iscritti. $B = \{0, 1, 2 \ldots 99\}$.
C: almeno 200 iscritti. $C = \{200, 201, \ldots\}$.

Esempio 1.4: Altezza di un individuo

Possiamo interpretare come esperimento aleatorio la misura dell'altezza di un individuo scelto a caso, espressa in centimetri e con qualunque livello di precisione (cioè con un numero di decimali potenzialmente infinito). Ciò fa sí che ogni numero reale positivo è un possibile esito, sicché lo spazio campione è l'intervallo $(0, +\infty)$. Sono esempi di eventi

A: altezza compresa fra 170 e 180 cm. $A = [170, 180]$.
B: altezza inferiore a 165 cm. $B = (0, 165)$.
C: altezza superiore a 2 metri. $C = [200, +\infty)$.

Possiamo effettuare operazioni fra eventi facendo le corrispondenti operazioni fra insiemi.

Definizione 1.2: Operazioni tra eventi

$A \wedge B \quad = A \cap B$ — *evento intersezione o prodotto logico, si legge "A e B"*
l'evento che si verifica quando sia A che B si verificano,

$A \vee B \quad = A \cup B$ — *evento unione o somma logica, si legge "A o B"*
l'evento che si verifica quando almeno uno tra A e B si verifica,

$\neg A \quad = \Omega \setminus A$ — *evento negazione, si legge "non A"*
l'evento che si verifica quando A non è verificato.

- *L'evento che coincide con tutto lo spazio campione si dice evento certo.*
- *L'insieme vuoto è detto evento impossibile,*
- *Due eventi A e B sono mutuamente esclusivi o incompatibili se non si possono verificare contemporaneamente, cioè se l'evento $A \wedge B$ è impossibile, ovvero se $A \cap B = \emptyset$.*
- *L'evento A implica l'evento B se $A \subset B$. In questo caso, infatti, se l'esito dell'esperimento appartiene ad A, allora appartiene anche a B.*

Esempio 1.5: Mazzo di carte

Scegliamo casualmente una carta da un mazzo di carte francesi (senza Jolly). Il nostro spazio campione è costituito dalle 52 carte del mazzo. Consideriamo gli eventi

A: un asso,
B: una carta di cuori,
C: una carta di seme rosso,
D: una carta con un numero pari.

Si ha

$A \wedge B$ = *asso di cuori,* $\quad A \wedge D$ = *evento impossibile ,*
$B \wedge C$ = *una carta di cuori,* $\quad B \vee C$ = *una carta di seme rosso ,*
$\neg D$ = *una carta con un numero dispari oppure una figura.*

Gli eventi A e D sono mutuamente esclusivi. L'evento B implica l'evento C.

1.1. La misura di probabilità. Ci vogliamo ora occupare di misurare la probabilità di un dato evento. Precisamente, ad ogni evento vogliamo associare un numero che cresca all'aumentare della probabilità che tale evento si verifichi. Dal punto di vista matematico, dunque, la misura di probabilità sarà una funzione che ha come *dominio* l'insieme degli eventi e fornisce come output un *numero*. Nel seguito la indicheremo con la lettera P.

Definire la legge di questa funzione non è un problema di facile soluzione, e in effetti il concetto di misura di probabilità si è evoluto nel tempo. La prima idea, attribuita a Laplace, è nota come definizione classica di probabilità.

Definizione 1.3: Probabilità - definizione classica

La probabilità di un evento è il rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili.

Per scrivere una formula, introduciamo qualche notazione. Indichiamo con Ω lo spazio campione, con $A \subset \Omega$ un evento e con il simbolo $|\cdot|$ il *numero* di elementi distinti contenuti in un dato insieme. Abbiamo allora

$$P(A) \coloneqq \frac{|A|}{|\Omega|}.$$

Ad esempio, se consideriamo lo spazio campione associato al lancio di un dado e l'evento A : numero pari (Esempio 1.1), abbiamo Ora $|\Omega| = 6$ e $|A| = 3$, dunque $\mathrm{P}(A) = 3/6 = 1/2$.

Osservazione 1.4

Nella vita di tutti i giorni capita spesso di parlare di probabilità in termini di percentuali. In questo caso, ad esempio, diremmo che c'è una probabilità del 50% che esca un numero pari. Non dimentichiamo che le percentuali altro non sono che frazioni con denominatore 100, e infatti $50\% = \dfrac{50}{100} = \dfrac{1}{2}$.

Dalla definizione seguono tre regole:

Proposizione 1.5: Regole di calcolo della probabilità

i) *la probabilità di un evento è un numero compreso tra 0 e 1;*
ii) *la probabilità dell'evento certo è pari a 1;*
iii) *la probabilità del verificarsi di almeno uno tra due eventi incompatibili è pari alla somma delle probabilità dei due eventi, cioè*

$$A \cap B = \emptyset \;\Rightarrow\; \mathrm{P}(A \vee B) = \mathrm{P}(A) + \mathrm{P}(B).$$

Esercizio 1.6

Un'urna contiene 7 biglie rosse, 4 bianche e 9 verdi. Si consideri l'esperimento aleatorio di estrarre una biglia dall'urna.

1. *Qual è lo spazio campione corrispondente?*
2. *Si descrivano insiemisticamente gli eventi*
   A: estrazione di una biglia rossa,
   B: estrazione di una biglia verde.
3. *Si calcolino le probabilità di A e di $\neg B$.*

Esercizio 1.7

Calcolare la probabilità di fare terno giocando 3 numeri al lotto sulla ruota di Napoli.

Svolgimento. In base alla nozione classica di probabilità, devo calcolare quanti sono gli esiti possibili e quanti quelli favorevoli all'evento. I possibili esiti dell'estrazione del lotto su un'unica ruota sono tutte le combinazioni semplici di 5 numeri su 90, dunque il nostro spazio campione ha $\mathcal{C}_{90,5} = \binom{90}{5}$ punti. Gli eventi a noi favorevoli sono tutte le possibili combinazioni di 5 numeri che contengono i

3 che abbiamo giocato. Esse sono tante quante le combinazioni di 2 numeri sui 97 che non abbiamo giocato, ovvero $\mathcal{C}_{87,2} = \binom{87}{2}$. Di conseguenza la probabilità di fare terno è

$$\frac{\mathcal{C}_{87,2}}{\mathcal{C}_{90,5}} = \frac{\binom{87}{2}}{\binom{90}{5}} = \frac{87!\,5!\,85!}{2!\,85!\,90!} = \frac{5 \cdot 4 \cdot 3}{88 \cdot 89 \cdot 90} = \frac{60}{704880} = 0.000085121 \approx 8 \cdot 10^{-5} \quad \text{o } 0.008\%.$$

□

Nel seguito (Esercizio 2.11) vedremo tencniche più efficaci per calcolare la probabilità in situazioni di questo tipo, e anche più complesse. Esse si poggiano sulla teoria astratta della probabilità.

La nozione classica di probabilità risulta utile nel caso di spazi campione che contengono un numero finito di esiti equiprobabili (cioè tutti con uguale probabilità di verificarsi), come ad esempio nel lancio di un dado o di una moneta, o nell'estrazione da un mazzo di carte. Se però prendiamo un dado truccato, il risultato non è più corretto poiché a esiti diversi corrisponderanno diverse probabilità. Inoltre, se i possibili esiti sono infiniti (come nell'Esempio 0.2) o addirittura continui (come nell'Esempio 1.4) la formula della probabilità classica (Definizione 1.3) non può essere applicata.

Per valutare la probabilità in queste situazioni ci è utile appellarci alla statistica: dalla analisi delle immatricolazioni negli anni passati possiamo dedurre previsioni sulle immatricolazioni future. In questo caso, stiamo usando una nozione di probabilità mutuata dalla statistica, la cosiddetta nozione frequentista attribuita a von Mises.

Definizione 1.6: Probabilità - definizione frequentista

La probabilità di un evento è il limite cui tende la frequenza relativa dell'evento al crescere del numero degli esperimenti.

Per scrivere una formula, introduciamo qualche notazione. Indichiamo con n il numero di esperimenti effettuati, con A un evento e con n_A il numero degli esperimenti in cui A si è verificato.
Il rapporto n_A/n rappresenta la frequenza di A.
Abbiamo allora

$$P(A) \coloneqq \lim_{n\to\infty} \frac{n_A}{n}.$$

Anche da questa definizione si ricavano le tre regole elencate nella Proposizione 1.5.

La nozione frequentista si applica ad esperimenti casuali i cui esiti non siano ritenuti ugualmente possibili, e siano ipoteticamente infiniti. Tuttavia, essa assume che l'esperimento sia ripetibile più volte, idealmente infinite, sotto le stesse condizioni. Inoltre il limite che appare nella formula non è "calcolabile": chi conosce la legge con cui varia la frequenza (al variare del numero di esperimenti)?

Per ovviare a queste problematiche, De Finetti e Savage hanno proposto una definizione di probabilità applicabile ad esperimenti casuali che non siano necessariamente ripetibili più volte sotto le stesse condizioni

Definizione 1.7: Probabilità - definizione soggettiva

La probabilità di un evento è il prezzo che un individuo razionale ritiene equo pagare per ricevere 1 se l'evento si verifica, 0 se l'evento non si verifica.

L'aggettivo "razionale" si usa per ribadire un criterio di coerenza: la misura di probabilità deve essere attribuita in modo tale che non sia possibile ottenere una vincita o una perdita certa, cioè il prezzo pagato deve essere compreso tra 0 e 1.

Grazie al criterio di coerenza si ricavano dalla definizione soggettiva le stesse tre regole di calcolo della Proposizione 1.5.

La definizione soggettiva consente quindi di calcolare la probabilità di eventi anche quando gli eventi elementari non sono equiprobabili e quando l'esperimento non può essere ripetuto. Rimane fondata, tuttavia, 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.

Un decisivo cambio di prospettiva viene introdotto negli anni '30 del secolo scorso da Kolmogorov. Egli non si occupa di dare una definizione operativa e fornire indicazioni su come calcolare la probabilità, bensì di dare un assetto formale alla "teoria della probabilità", nell'ambito del quale sia possibile dimostrare teoremi che poi potranno essere utilizzati sia nell'ambito di un approccio frequentista che nell'ambito di un approccio soggettivista.

La definizione assiomatica dà alla probabilità lo status di una branca della matematica ed è quella cui faremo riferimento d'ora in avanti. Per la sua importanza, le dedichiamo una sezione distinta.

1.2. Definizione assiomatica di probabilità e proprietà. Volendo dare una nozione di probabilità che si adatti ad ogni ambito applicativo (in particolare, a spazi campione infiniti e possibilmente non numerabili), dobbiamo innanzitutto chiarire *quali eventi sono misurabili*. Nel caso di spazi campione finiti o numerabili, non ci sono grossi problemi nel misurare la probabilità di un qualunque evento (cioè di un qualunque sottoinsieme dello spazio campione), ma ciò non è più vero se trattiamo con spazi campione che abbiano la potenza del continuo (come $\mathbb{R}$ o un intervallo).

Se Ω è un insieme qualunque, indichiamo con 2^Ω il suo *insieme delle parti*, ovvero la famiglia di tutti i suoi sottoinsiemi. Definiamo poi

Definizione 1.8: σ-algebra

Sia $\mathcal{E} \subset 2^\Omega$ una famiglia non vuota di sottoinsiemi di Ω. $\mathcal{E}$ si dice σ-algebra se soddisfa le proprietà

i) se un evento E appartiene a $\mathcal{E}$, vi appartiene anche la sua negazione:

$$E \in \mathcal{E} \;\Rightarrow\; \neg E \in \mathcal{E};$$

ii) se un'infinità numerabile di eventi, $E_1, E_2, \ldots E_n, \ldots$ appartiene a $\mathcal{E}$, vi appartiene anche l'evento costituito dalla loro unione:

$$E_n \in \mathcal{E}\ \forall n \in \mathbb{N} \;\Rightarrow\; \bigcup_{n\in\mathbb{N}} E_n \in \mathcal{E}.$$

Dalla definizione di σ-algebra discende che

Proposizione 1.9

Sia $\mathcal{E} \subset 2^\Omega$ una σ-algebra, allora

- *l'evento impossibile e l'evento certo appartengono alla σ-algebra, cioè $\emptyset, \Omega \in \mathcal{E}$;*
- *se un'infinità numerabile di eventi, $E_1, E_2, \ldots E_n, \ldots$ appartiene a $\mathcal{E}$, vi appartiene anche l'evento costituito dalla loro intersezione, cioè*

$$E_n \in \mathcal{E}\ \forall n \in \mathbb{N} \;\Rightarrow\; \bigcap_{n\in\mathbb{N}} E_n \in \mathcal{E}.$$

Dato un qualunque insieme Ω, la totalità del suo insieme delle parti è ovviamente una σ-algebra. Se prendiamo $\mathbb{R}$ come spazio campione, la famiglia degli intervalli non è una σ-algebra poiché, ad esempio, l'unione di due intervalli disgiunti non è un intervallo.

Esempio 1.8: algebra di Borel reale

Si dice algebra di Borel la più piccola σ-algebra che contiene tutti gli intervalli. Essa consiste, oltre che di tutti gli intervalli (siano essi aperti o chiusi), di tutte le possibili unioni e intersezioni (finite o numerabili) di intervalli.
Ogni insieme costituito da un singolo punto fa parte dell'algebra di Borel, perché l'insieme $\{x\}$ coincide con l'intersezione della famiglia numerabile di intervalli $A_n = (x - 1/n, x + 1/n)$.

Ora possiamo precisare il significato di **misura di probabilità**.

Definizione 1.10: probabilità - definizione assiomatica

Siano Ω uno spazio campione e $\mathcal{E} \subset 2^\Omega$ una σ-algebra di eventi. Una funzione $\mathrm{P} : \mathcal{E} \mapsto \mathbb{R}$ si dice misura di probabilità se soddisfa le proprietà

(i) la probabilità di ogni evento è nonnegativa, cioè
$$\mathrm{P}(A) \geq 0 \quad \forall A \in \mathcal{E};$$
(ii) l'evento certo ha probabilità uno, cioè
$$\mathrm{P}(\Omega) = 1;$$
(iii) se $A_1, \cdots A_n, \cdots \in \mathcal{E}$ è una famiglia finita o numerabile di eventi mutualmente esclusivi (ovvero tali che $A_n \wedge A_m = \emptyset$ se $n \neq m$), la probabilità della somma logica è la somma delle probabilità, cioè
$$\mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathrm{P}(A_n).$$

La terna $(\Omega, \mathcal{B}, \mathrm{P})$ viene detta **spazio di probabilità**.

In particolare l'assioma iii) si applica anche ad una coppia di eventi incompatibili, sicché
$$A \wedge B = \emptyset \implies \mathrm{P}(A \vee B) = \mathrm{P}(A) + \mathrm{P}(B).$$

Esempio 1.9: Dado classico

Se $\Omega = \{1, 2, 3, 4, 5, 6\}$ è lo spazio campione associato al lancio di un dado a sei facce, $\mathcal{E}$ la σ-algebra di tutti i suoi possibili sottoinsiemi e P la probabilità classica definita in 1.3, allora la terna $(\Omega, \mathcal{B}, \mathrm{P})$ è uno spazio di probabilità secondo la Definizione 1.10

Esempio 1.10: Dado truccato

Sia $\Omega = \{1, 2, 3, 4, 5, 6\}$ lo spazio campione associato al lancio di un dado a sei facce e $\mathcal{E}$ la σ-algebra di tutti i suoi possibili sottoinsiemi. Si assegni poi la misura di probabilità P come segue

- $\mathrm{P}(\{1\}) = 1/12$, $\mathrm{P}(\{6\}) = 1/3$,
- $\mathrm{P}(\{2\}) = \mathrm{P}(\{3\}) = \mathrm{P}(\{4\}) = \mathrm{P}(\{5\}) = 7/48$,
- *se $A \subset \{1, 2, 3, 4, 5, 6\}$ è un qualunque evento, $\mathrm{P}(A)$ si ottiene sommando le probabilità di tutti gli esiti che lo compongono.*

È facile verificare che la terna $(\Omega, \mathcal{B}, \mathrm{P})$ è uno spazio di probabilità. Essa corrisponde ad un dado "truccato", in cui l'esito 6 è più probabile.

Esempio 1.11: Integrale e probabilità

Siano $\Omega = \mathbb{R}$ *e* $\mathcal{E}$ *l'algebra di Borel (Esempio 1.8). Se* A *è un evento boreliano, definiamo*

$$\mathrm{P}(A) = \int_A \frac{1}{\pi(1+x^2)} dx.$$

Essa è una misura di probabilità in quanto

i) La monotonia dell'integrale (rispetto all'integrando) implica che $P(A) \geq 0$*, poichè*

$$\frac{1}{\pi(1+x^2)} \geq 0.$$

ii) $\mathrm{P}(\mathbb{R}) = \int_{-\infty}^{+\infty} \frac{1}{\pi(1+x^2)} dx = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b \frac{1}{\pi(1+x^2)} dx = \lim_{\substack{a\to-\infty \\ b\to+\infty}} \frac{1}{\pi}(\arctan b - \arctan a) = \frac{1}{\pi}(\frac{\pi}{2} - (-\frac{\pi}{2})) = 1.$

iii) La verifica del terzo assioma è un po' più complicata, ma fattibile... Nel caso di una famiglia finita di intervalli, discende dall'additività dell'integrale rispetto all'insieme di integrazione.

Si noti che ogni singolo esito ha probabilità nulla, cioè $\mathrm{P}(\{t\}) = 0$ *per ogni* $t \in \mathbb{R}$*, infatti*

$$\mathrm{P}(\{t\}) = \int_t^t \frac{1}{\pi(1+x^2)} dx = 0.$$

Dagli assiomi si ricavano immediatamente alcune proprietà elementari.

Proposizione 1.11: proprietà elementari della probabilità

Sia dato uno spazio di misura $(\Omega, \mathcal{E}, \mathrm{P})$ *e siano* $A, B \in \mathcal{E}$ *due eventi. Allora*

(1.1) $0 \leq \mathrm{P}(A) \leq 1, \quad \mathrm{P}(\neg A) = 1 - \mathrm{P}(A), \qquad \mathrm{P}(\emptyset) = 0,$

(1.2) $\mathrm{P}(A \vee B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \wedge B)$

(1.3) $\mathrm{P}(A \wedge B) \leq \mathrm{P}(A) \leq \mathrm{P}(A \vee B).$

Inoltre se A *implica* B*, allora*

(1.4) $\mathrm{P}(A) \leq \mathrm{P}(B).$

Dimostrazione. Dimostriamo solo la proprietà (1.2), nota anche come legge della probabilità totale.

Risulta utile decomporre gli eventi A, B e $A \wedge B$ nella somma logica di altri eventi fra loro incompatibili, come segue

$$A = (A \setminus B) \vee (A \wedge B), \qquad B = (B \setminus A) \vee (A \wedge B),$$
$$A \vee B = (A \setminus B) \vee (B \setminus A) \vee (A \wedge B).$$

Applicando ora la proprietà assiomatica iii) ad ognuna di queste decomposizioni si ottiene

$$\mathrm{P}(A) = \mathrm{P}(A \setminus B) + \mathrm{P}(A \wedge B), \qquad \mathrm{P}(B) = \mathrm{P}(B \setminus A) + \mathrm{P}(A \wedge B),$$
$$\mathrm{P}(A \vee B) = \mathrm{P}(A \setminus B) + \mathrm{P}(B \setminus A) + \mathrm{P}(A \wedge B).$$

Sommando le prime due uguaglianze si ottiene

$$\mathrm{P}(A) + \mathrm{P}(B) = \mathrm{P}(A \setminus B) + \mathrm{P}(B \setminus A) + 2\mathrm{P}(A \wedge B)$$

e ricordando la terza uguaglianza si ha

$$= \mathrm{P}(A \vee B) + \mathrm{P}(A \wedge B).$$

Sottraendo la quantità $\mathrm{P}(A \wedge B)$ al primo e all'ultimo membro dell'uguaglianza si ha la tesi. □

Gli assiomi (ii) e (iii), assieme alla proprietà (1.1), altro non sono che le regole di calcolo che abbiamo riconosciuto per la probabibilità classica, frequentista e soggettiva. Di fatto, una misura di probabilità è una qualunque funzione degli eventi che soddisfa le tre regole di calcolo empiriche.

Esercizio 1.12

A partire dalla definizione assiomatica di probabilità (Definizione 1.10), dimostrare le proprietà (1.1), (1.3) *e* (1.4).

Esercizio 1.13: lancio di due dadi

Nel lancio due dadi perfetti e distinguibili (immaginiamo che uno sia di colore rosso e l'altro verde), calcolare la probabilità degli eventi
A: somma pari a 8,
B: uscita di almeno un 1.

Svolgimento. Dobbiamo innanzitutto capire quali sono i possibili esiti dell'esperimento aleatorio, cioè individuare lo spazio campione. In linea di principio, lo spazio campione associato è il prodotto cartesiano dei due spazi campione associati ad ogni singolo dado, cioè

$$\Omega = \{1, \dots, 6\} \times \{1, \dots, 6\} = \{(i, j) \,:\, i = 1, \dots, 6;\ j = 1, \dots, 6\},$$

che possiamo equipaggiare con la probabilità classica. Usando le nostre conoscenze di calcolo combinatorio, possiamo dire che lo spazio campione è formato da tutti i possibili raggruppamenti di due oggetti scelti nell'insieme $\{1, 2, \dots 6\}$.
Per calcolare la probabilità dell'evento A dovremo prima valutare quali sono gli esiti favorevoli (cioè che danno somma 8): $(2, 6)$, $(6, 2)$, $(3, 5)$, $(5, 3)$, $(4, 4)$ e poi fare il rapporto fra il numero dei casi favorevoli (5) e il numero dei casi possibili ($6 \times 6 = 36$). Ne ricaviamo che la probabilità che la somma sia 8 è $5/36$.
Possiamo calcolare la probabilità dell'evento B in modo analogo: ora gli eventi favorevoli sono 11, precisamente $(1, 1), (1, 2), \cdots (1, 6), (2, 1), (3, 1), \cdots (6, 1)$ e dunque la probabilità è $11/36$. Questo calcolo è semplice ma laborioso: immaginate cosa succederebbe se dovessimo gestire contemporaneamente 10 dadi!
In alternativa, possiamo considerare i due dadi come due copie dello stesso spazio campione $\Omega = \{1, \cdots, 6\}$. In quest'ottica, il nostro evento B si verifica se almeno uno dei due dadi verifica l'evento "uscita dell'1". In simboli abbiamo

B_1 : uscita dell'1 sul dado rosso, $\mathrm{P}(B_1) = 1/6$,
B_2 : uscita dell'1 sul dado verde, $\mathrm{P}(B_2) = \mathrm{P}(B_1) = 1/6$,

e $B = B_1 \vee B_2$. Dunque la legge della probailità totale (1.2) afferma che

$$\mathrm{P}(B) = \mathrm{P}(B_1 \vee B_2) = \mathrm{P}(B_1) + \mathrm{P}(B_2) - \mathrm{P}(B_1 \wedge B_2) = 1/3 - \mathrm{P}(B_1 \wedge B_2).$$

D'altra parte $B_1 \wedge B_2$ corrisponde al solo esito $(1, 1)$, dunque $\mathrm{P}(B_1 \wedge B_2) = 1/36$ e infine $\mathrm{P}(B_1 \vee B_2) = 1/3 - 1/36 = 11/36$. □

2. Probabilità condizionata

Immaginate di giocare a dadi, scommettendo che uscirà il numero 3. Prima di effettuare il lancio, la vostra probabilità di vittoria è la probabilità dell'evento $A = \{3\}$, ovvero $P(A) = 1/6$. Poi il lancio viene effettuato: voi non vedete l'esito (che dunque per voi è ancora aleatorio), ma ricevete l'informazione "È uscito un numero dispari". Sapete dunque che si è verificato l'evento $B = \{1, 3, 5\}$.

A questo punto, la vostra confidenza nella vittoria cambia perché sono cambiate le informazioni in vostro possesso. Se prima valutavate uno spazio campione composto da 6 esiti possibili, ora sapete che in effetti si è verificato uno tra i 3 esiti in B. Dunque la probabilità di vittoria che vi attribuite non è più $1/6$, bensì $1/3$, ovvero il rapporto fra 1 (l'unico esito a voi favorevole) e 3 (gli esiti possibili, alla luce della nuova informazione).
Di fatto, avete calcolato la probabilità dell'evento A, condizionata al verificarsi dell'evento B.

Definizione 2.1: Probabilità condizionata

In uno spazio di misura $(\Omega, \mathcal{E}, \mathrm{P})$, siano $A, B \in \mathcal{E}$ due eventi. Si dice probabilità condizionata di A, dato B (o probabilità di A condizionata a B), la probabilità che l'evento A ha di verificarsi quando si sappia che B si è verificato.
Essa si indica con il simbolo $\mathrm{P}(A|B)$ e si definisce mediante la formula

$$P(A|B) \coloneqq \frac{P(A \wedge B)}{P(B)}.$$

Osservazione 2.2: Probabilità condizionata in senso classico

Quando il nostro spazio di misura è formato da un numero finito di esiti equiprobabili (cioè stiamo lavorando con la definizione classica di probabilità), abbiamo

$$P(A \wedge B) = \frac{\#\text{ esiti favorevoli sia ad } A \text{ che a } B}{\#\text{ esiti possibili}}, \qquad P(B) = \frac{\#\text{ esiti favorevoli a } B}{\#\text{ esiti possibili}},$$

quindi la Definizione 2.1 si semplifica in

$$P(A|B) := \frac{\#\text{ esiti favorevoli sia ad } A \text{ che a } B}{\#\text{ esiti favorevoli a } B}.$$

In questo caso, calcolare la probabilità condizionata a B è equivalente a calcolare la probabilità classica nello spazio campionario B.

Osserviamo esplicitamente che i due eventi che entrano in gioco nella definizione di probabilità condizionata (Definizione 2.1) hanno due ruoli diversi. Quando calcoliamo la probabilità di A condizionata a B, l'evento B è dato per certo (indipendentemente da quale fosse la sua probabilità a priori), lo chiamiamo *evento condizionante*. L'evento A, invece, si ritiene incerto e ci interessa proprio valutare la sua probabilità. Diremo che A è l'*evento condizionato*.

Attenzione dunque a non scambiare i ruoli di condizionante e condizionato.

Esercizio 2.1

Un gruppo di amici è formato da 8 uomini e 6 donne. Degli uomini, 4 sono castani, 3 biondi e uno calvo, mentre delle donne 2 sono castane e 4 bionde.

(i) *Qual è la probabilità che una donna sia bionda?*
(ii) *Qual è la probabilità che una persona coi capelli biondi sia donna?*

Svolgimento. Indichiamo

B: capelli biondi,
D: donna.

Nel gruppo di 14 amici, quelli coi capelli biondi sono in totale 7, mentre le donne sono 6. Quindi a priori ho $P(B) = 7/14 = 1/2$ e $P(D) = 6/14 = 3/7$.
Posso rispondere direttamente alla domanda *(i)* osservando che su 6 donne, 4 sono bionde e dunque la probabilità è $4/6 = 2/3$ (qui ho usato la formula semplificata dell'Osservazione 2.2). Oppure posso calcolare la probabilità di B condizionata a D secondo la Definizione 2.1. Ora $B \wedge D$ è il

gruppo delle donne bionde, che comprende 4 persone, dunque $P(B \wedge D) = 4/14 = 2/7$ e

$$P(B|D) = \frac{P(B \wedge D)}{P(D)} = \frac{\frac{2}{7}}{\frac{3}{7}} = \frac{2}{3}.$$

Per quanto riguarda la domanda *(ii)*, posso osservare che su 7 persone bionde, le donne sono 4 e dunque la probabilità è $4/7$. Oppure calcolo la probabilità di D condizionata a B, che da definizione è

$$P(D|B) = \frac{P(B \wedge D)}{P(B)} = \frac{\frac{2}{7}}{\frac{1}{2}} = \frac{4}{7}.$$

□

Partendo dalla definizione, è facile verificare le seguenti proprietà.

Proposizione 2.3: Proprietà elementari della probabilità condizionata

- *Se A e B sono mutuamente esclusivi, allora* $P(A|B) = 0$.
- *Se B implica A, allora* $\mathrm{P}(A|B) = 1$.
- *Se A implica B, allora* $\mathrm{P}(A|B) = \mathrm{P}(A)/\mathrm{P}(B) \geq \mathrm{P}(A)$.

Esercizio 2.2

Verificate le proprietà elencate nella Proposizione 2.3.

Esercizio 2.3

Mario pesca una carta da un mazzo di 52 carte francesi. Ada, Bernardo e Claudia scommettono rispettivamente su

A: qualsiasi figura (cioè J, Q o K) di seme rosso (cioè cuori o quadri),
B: qualsiasi figura di seme rosso, oppure una qualunque carta compresa tra 1 e 10 di seme nero,
C: qualunque asso.

Calcolare la probabilità di vittoria dei 3 giocatori.
Se poi Mario, senza svelare quale carta ha pescato, dice che è una carta di cuori, come cambia l'aspettativa di vittoria dei 3 giocatori?

Svolgimento. Gli eventi A, B e C contengono rispettivamente $3 \times 2 = 6$, $3 \times 2 + 10 \times 2 = 26$ e $1 \times 4 = 4$ esiti. Dunque le probabilità a priori dei 3 giocatori sono

$$P(A) = \frac{6}{52} = \frac{3}{26} \approx 0.115 \;\text{ o }\; 11\%, \qquad P(B) = \frac{6+20}{52} = \frac{1}{2} = 0.5 \;\text{ o }\; 50\%,$$
$$P(C) = \frac{4}{52} = \frac{1}{13} \approx 0.077 \;\text{ o }\; 8\%,$$

calcolate usando la nozione classica di probabilità.
Passiamo poi a calcolare la probabilità a posteriori, cioè dopo che abbiamo saputo che si è verificato l'evento rivelato da Mario, che indicheremo con la lettera M. M è costituito da 13 carte e dunque $P(M) = \frac{13}{52} = \frac{1}{4}$.
Ora $A \wedge M$ è rappresentato dalle 3 figure di cuori, quindi $P(A \wedge M) = \frac{3}{52}$ e da definizione

$$P(A|M) = \frac{\frac{3}{52}}{\frac{13}{52}} = \frac{3}{13} \approx 0.230 \;\text{ o }\; 23\%.$$

Anche $B \wedge M$ è rappresentato dalle 3 figure di cuori, quindi $P(B \wedge M) = \dfrac{3}{52}$ e

$$P(B|M) = \frac{\frac{3}{52}}{\frac{13}{52}} = \frac{3}{13} \approx 0.230 \ \text{ o } \ 23\%.$$

Infine $C \wedge M$ è costituito dal solo asso di cuori, quindi $P(C \wedge M) = \frac{1}{52}$ e

$$P(C|M) = \frac{\frac{1}{52}}{\frac{13}{52}} = \frac{1}{13} \approx 0.077 \ \text{ o } \ 8\%.$$

□

Notiamo come, in questo esercizio, la probabilità condizionata di A è maggiore della sua probabilità a priori, mentre quella di B è minore e quella di C è invariata. Questo fatto ci dà una misura dell'idea intuitiva che l'evento M è favorevole ad A, sfavorevole a B e indifferente a C. Diamo una definizione generale.

Definizione 2.4: Eventi indipendenti e dipendenti

In uno spazio di misura $(\Omega, \mathcal{E}, \mathrm{P})$, *siano* $A, B \in \mathcal{E}$ *due eventi*
Essi si dicono indipendenti se $\mathrm{P}(A|B) = \mathrm{P}(A)$. *Altrimenti si dicono dipendenti; in particolare diremo che sono positivamente correlati se* $\mathrm{P}(A|B) > \mathrm{P}(A)$, *e viceversa negativamente correlati se* $\mathrm{P}(A|B) < \mathrm{P}(A)$.

Nell'Esercizio 2.3 abbiamo visto che M è positivamente correlato con A, negativamente correlato con B e indipendente da C.

2.1. La legge della probabilità composta. Dalla definizione di probabilità condizionata discende il seguente teorema, che risulta utilissimo nel calcolo di esperimenti ripetuti o composti.

Teorema 2.5: Legge delle probabilità composte

In uno spazio di misura $(\Omega, \mathcal{E}, \mathrm{P})$, *siano* $A, B \in \mathcal{E}$ *due eventi. Allora*

$$P(A \wedge B) = P(A|B)\, P(B) = P(B|A)\, P(A) \tag{2.1}$$

e

$$\mathrm{P}(A \wedge B) = \mathrm{P}(A)\, \mathrm{P}(B) \tag{2.2}$$

se, e solo se, A *e* B *sono indipendenti.*
Inoltre se abbiamo qualsiasi numero di eventi $A_1, A_2, A_3 \dots A_n$ *si ha*

$$P(A_1 \wedge A_2 \cdots A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \wedge A_2) \cdots P(A_n|A_1 \wedge A_2 \dots A_{n-1}). \tag{2.3}$$

Dimostrazione. Per dimostrare la prima uguaglianza in (2.1), cioè $P(A \wedge B) = P(A|B)\, P(B)$ basta prendere l'equazione definitiva di probabilità condizionata $P(A|B) \coloneqq \dfrac{P(A \wedge B)}{P(B)}$, moltiplicare ambo i membri per la quantità $P(B)$ e semplificare.
Analogamente, per dimostrare $P(A \wedge B) = P(B|A)\, P(A)$, partiamo dalla definizione $P(B|A) \coloneqq \dfrac{P(A \wedge B)}{P(A)}$, moltiplichiamo ambo i membri per la quantità $P(A)$ e semplifichiamo.

Inoltre la formula (2.2) vale se, e solo se, $\mathrm{P}(A|B) = \mathrm{P}(B)$, cioè appunto se A e B sono indipendenti. □

Esercizio 2.4

Abbiamo due urne. L'urna $\mathcal{A}$ contiene 3 palline rosse, 3 bianche e 5 nere; l'urna $\mathcal{B}$ contiene 6 palline bianche e 10 nere. Si estrae una pallina da $\mathcal{A}$ e una da $\mathcal{B}$. Calcolare le probalilità che

a) siano entrambe nere,
b) siano una rossa e una bianca,
c) non siano entrambe bianche.

Esercizio 2.5

Consideriamo il dado truccato dell'Esempio 1.10, in cui il numero sei ha probabilità $1/3$, il numero uno ha probabilità $1/12$ e tutti gli altri numeri hanno probabilità $7/48$. Luca, non sapendo che il dado è truccato, ha puntato sul numero cinque. Davide, consapevole che il dado è truccato, viene informato che il lancio ha avuto come esito un numero dispari. Che probabilità di vittoria attribuisce a Luca?

Esercizio 2.6

La domenica Susanna va al parco solo se non piove e la sua amica Cristina è libera. Per la prossima domenica sono previste piogge con probabilità del 20%. Cristina va a trovare i nonni una domenica su due, indipendentemente dal tempo. Con che probabilità Susanna andrà al parco domenica prossima?

Osservazione 2.6: Estrazioni ripetute

Quando facciamo estrazioni ripetute dalla stessa urna (o mazzo di carte), ogni evento che riguarda la successione di estrazioni può essere interpretato come l'intersezione di eventi che riguardano le singole estrazioni, dunque le formule (2.1) e (2.2) possono essere d'aiuto. Dobbiamo però chiarire chi sia la probabilità condizionata, in questo ambito. A tal fine è necessario precisare le regole del gioco; distinguiamo due diverse situazioni
Estrazioni ripetute con reintegro: *si effettua un'estrazione, si reintegra la situazione di partenza reinserendo nell'urna la biglia estratta, poi si procede con la seconda estrazione.*
In questo caso l'esito del primo turno non ha alcuna influenza sul secondo, sicché eventi riguardanti turni diversi sono indipendenti e vale la formula (2.2).
Estrazioni ripetute senza reintegro: *si effettua un'estrazione, si esclude la biglia estratta, poi si procede con la seconda estrazione.*
In questo caso l'esito della seconda estrazione dipende dall'esito della prima e dobbiamo utilizzare la formula generale (2.1). *I due turni, infatti, avvengono in due spazi campione diversi: l'esito del primo turno determina lo spazio campione in cui avviene il secondo.*

Esercizio 2.7: Estrazioni ripetute con e senza reintegro

Si effettuano di due estrazioni consecutive da un'urna contenente 7 biglie rosse e 3 nere. Calcolare le probabilità degli eventi

A: estrazione di due biglie rosse (con reintegro)
B: estrazione di due biglie rosse (senza reintegro)
C: estrazione di una biglia rossa e una nera (con reintegro)
D: estrazione di una biglia rossa e una nera (senza reintegro)

Svolgimento. Calcoliamo dapprima le probabilità degli eventi elementari

R : estrazione di una biglia rossa (su un'unica estrazione), $\mathrm{P}(R) = 7/10$,
N : estrazione di una biglia nera (su un'unica estrazione), $\mathrm{P}(N) = 3/10$.

Indichiamo poi con R_1/R_2, rispettivamente, gli eventi "estrazione di una biglia rossa al primo/secondo turno". Ora, tanto l'evento A quanto l'evento B corrispondono all'evento intersezione $R_1 \wedge R_2$ e dunque possono essere calcolati usando le formule della probabilità composta del Teorema 2.5. Tuttavia, a seconda che si estragga con o senza reintegro, cambia la correlazione tra i due eventi, e quindi il risultato finale
Nel caso con reintegro gli eventi R_1 e R_2 sono indipendenti, pertanto usiamo la formula (2.2)

$$\mathrm{P}(A) = \underbrace{\mathrm{P}(R_1 \wedge R_2)}_{\text{eventi indipendenti}} = \mathrm{P}(R_1)\,\mathrm{P}(R_2) = \mathrm{P}(R)^2 = \frac{49}{100} = 0.49 \quad \text{o} \quad 49\%.$$

Nel caso senza reintegro, invece, dobbiamo usare la formula (2.1)

$$\mathrm{P}(B) = \underbrace{\mathrm{P}(R_1 \wedge R_2)}_{\text{eventi dipendenti}} = \mathrm{P}(R_1)\,\mathrm{P}(R_2|R_1).$$

Ci serve calcolare $\mathrm{P}(R_2|R_1)$, cioè la probabilità di estrarre una biglia rossa dall'urna, avendone già estratta una rossa al primo turno. Alla seconda estrazione abbiamo un'urna che contiene solo 9 biglie, di cui 6 sono rosse e 3 nere, dunque la formula della probabilità "classica" ci dà $\mathrm{P}(R_2|R_1) = 6/9 = 2/3$ e infine

$$P(B) = \underbrace{\mathrm{P}(R_1 \wedge R_2)}_{\text{eventi dipendenti}} = \frac{7}{10}\,\frac{2}{3} = \frac{7}{15} = 0.4\bar{6} \quad \text{o} \quad 47\%.$$

Passiamo ora a valutare la probabilità di estrarre due biglie di colore diverso. Usando le notazioni appena introdotte, sia gli eventi C che D possono essere decomposti come $(R_1 \wedge N_2) \vee (N_1 \wedge R_2)$. Poiché i due eventi $R_1 \wedge N_2$ e $N_1 \wedge R_2$ sono mutuamente esclusivi, si ha senz'altro che

$$\mathrm{P}\left((R_1 \wedge N_2) \vee (N_1 \wedge R_2)\right) = \mathrm{P}\left(R_1 \wedge N_2\right) + \mathrm{P}\left(N_1 \wedge R_2\right).$$

Poi, nel valutare le due probabilità rimaste, dovremo distinguere il caso con reintegro (eventi indipendenti) da quello senza reintegro (eventi dipendenti).
Se c'è reintegro abbiamo

$$\underbrace{\mathrm{P}(R_1 \wedge N_2)}_{\text{eventi indipendenti}} = \mathrm{P}(R)\,\mathrm{P}(N) = \frac{21}{100} = 0.21, \qquad \underbrace{\mathrm{P}(N_1 \wedge R_2)}_{\text{eventi indipendenti}} = \mathrm{P}(N)\,\mathrm{P}(R) = \frac{21}{100} = 0.21,$$

quindi $\mathrm{P}(C) = 2 \cdot \underbrace{\mathrm{P}(R_1 \wedge N_2)}_{\text{eventi indipendenti}} = \frac{42}{100} = 0.42 \quad \text{o} \quad 42\%.$

Invece se non c'è reintegro

$$\underbrace{\mathrm{P}(R_1 \wedge N_2)}_{\text{eventi dipendenti}} = \mathrm{P}(R)\,\mathrm{P}(N|R) = \frac{7}{10}\,\frac{3}{9} = \frac{7}{30} = 0.2\bar{3}.$$

anche se è meno evidente, anche ora $\mathrm{P}\left(N_1 \wedge R_2\right) = \mathrm{P}(R_1 \wedge N_2)$, infatti

$$\underbrace{\mathrm{P}(N_1 \wedge R_2)}_{\text{eventi dipendenti}} = \mathrm{P}(N)\,\mathrm{P}(R|N) = \frac{3}{10}\,\frac{7}{9} = \frac{7}{30} = 0.2\bar{3},$$

dunque concludiamo $\mathrm{P}(D) = 2 \cdot \underbrace{\mathrm{P}(R_1 \wedge N_2)}_{\text{eventi dipendenti}} = \frac{7}{15} = 0.4\bar{6} \quad \text{o} \quad 47\%.$ □

Osservazione 2.7

Con questo esercizio ci siamo accorti che scambiando l'ordine degli esiti di due (o più) estrazioni otteniamo eventi che hanno uguale probabilità. Questo è vero sia nel caso con reintegro, che in quello senza reintegro!
Ripercorrendo i conti svolti, è facile convincersi che è una proprietà generale, in probabilità classica.

Esercizio 2.8

Da un'urna contenente 4 palline bianche, 3 nere e 3 rosse vengono estratte consecutivamente due palline, senza reinserire la prima. Calcolare le probalilità che

a) siano entrambe nere,
b) siano una rossa e l'altra no,
c) almeno una sia rossa,
d) nessuna sia bianca.

2.2. Esperimenti indipendenti e formula di Bernoulli. Le estrazioni ripetute con reintegro sono un esempio di una successione di esperimenti aleatori che avvengono tutti esattamente nelle stesse condizioni. In statistica queste successioni di eventi hanno grossa rilevanza e vi si riferisce come *prove ripetute* o *esperimenti indipendenti*.

Immaginiamo di fare un esperimento che ha successo quando si verifica un certo evento che indicheremo con la lettera E. Indichiamo poi con la lettera p la probabilità che l'esperimento abbia successo, cioè $p \coloneqq P(E)$, e con $q \coloneqq P(\neg E) = 1 - p$ la probabilità di insuccesso. Se ripetiamo la prova due volte (sempre nelle stesse condizioni), la legge degli eventi indipendenti (2.2) afferma che

· la probabilità di successo in entrambe le prove è p^2
· la probabilità di insuccesso in entrambe le prove è q^2
· per calcolare la probabilità di avere un successo e un insuccesso, dobbiamo fare un piccolo ragionamento. Certamente la probabilità di avere successo nel primo esperimento e non nel secondo è pq. Ma questo non esaurisce le possibilità, può anche succedere di avere successo nel secondo esperimento e non nel primo, e questo accade sempre con probabilità pq. Ne deduciamo dunque che la probabilità che E si verifichi una volta sì e l'altra no è complessivamente $2pq$ (grazie all'assioma *(iii)*).

Vale la pena di osservare che sommando le probabilità di questi tre eventi otteniamo $p^2+q^2+2pq = (p+q)^2 = 1^2 = 1$, come deve essere perchè essi esauriscono tutti i casi possibili e sono a due a due incompatibili.

Cosa succede se ripetiamo l'esperimento più volte? Ci risponde il cosiddetto "schema di Bernoulli", che enunciamo nel prossimo teorema.

Teorema 2.8: Formula di Bernoulli per le prove ripetute

Dato un esperimento aleatorio ripetibile n volte nelle stesse condizioni, siano p la probabilità di successo e $q = 1-p$ la probabilità di insuccesso. Allora la probabilità di ottenere esattamente k successi su n esperimenti è

$$\binom{n}{k} p^k q^{n-k}.$$

Dimostrazione. Si possono verificare tutti e soli gli eventi del tipo

$E_{k,n}$: l'esperimento ha successo esattamente k volte

con k che varia da 0 a n.
Come nell'esempio introduttivo ($n = 2$), è immediato calcolare che

$$P(E_{n,n}) = p^n, \qquad P(E_{0,n}) = q^n.$$

Cosa possiamo dire delle situazioni intermedie?
La probabilità che E si verifichi nei primi k esperimenti e non si verifichi in tutti gli $n-k$ esperimenti successivi è facilmente calcolabile con la formula (2.2) ed è

$$\underbrace{p\,p\ldots p}_{k \text{ volte}}\underbrace{q\,q\ldots q}_{n-k \text{ volte}} = p^k q^{n-k}.$$

In quanti altri modi può succedere che E si verifichi k volte su n? Ce lo dice il calcolo combinatorio: dobbiamo cambiare l'ordine delle lettere p (successo) e q (insuccesso) nella "parola" composta da

k lettere p e $n-k$ lettere q, quindi usiamo la formula per le permutazioni con ripetizioni della Proposizione 3.4 e otteniamo

$$P'_{k,n-k} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

Riassumendo, per ottenere la probabilità di $E_{k,n}$ dobbiamo sommare la probabilità di questi $P'_{k,n-k}$ eventi e otteniamo

$$P(E_{k,n}) = \binom{n}{k} p^k q^{n-k}.$$

□

Osservazione 2.9

Se ripetiamo n volte un esperimento, certamente si deve verificare uno tra gli eventi $E_{k,n}$ con $k = 0, 1, \dots n$. Quindi l'evento somma logica $E_{0,n} \vee E_{1,n} \vee \cdots E_{n,n}$ è certo, cioè $P(E_{0,n} \vee E_{1,n} \vee \cdots E_{n,n}) = 1$. D'altra parte gli $E_{k,n}$ si escludono a vicenda, dunque l'assioma (iii) in Definizione 1.10) assicura che

$$P(E_{0,n} \vee E_{1,n} \vee \cdots E_{n,n}) = \sum_{k=0}^{n} P(E_{k,n}).$$

Se ora sostituiamo il valore di $P(E_{k,n})$ che abbiamo appena calcolato si ottiene

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = 1. \tag{2.4}$$

Ritroviamo in senso probabilistico la formula di Newton per la potenza n-esima di un binomio, infatti

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1,$$

perché per definizione $p + q = 1$.

In effetti il Teorema 2.8 fornisce una dimostrazione della formula di Newton.

Corollario 2.10: Dimostrazione probabilistica della formula di Newton per le potenze di un binomio

Per ogni numero naturale $n \in \mathbb{N}$ e $a, b \in \mathbb{R}$, con $a, b \geq 0$, si ha

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

Dimostrazione. Introduciamo la notazione

$$p = \frac{a}{a+b}, \quad q = \frac{b}{a+b}.$$

Poichè $p, q \geq 0$ e $p + q = 1$, vale lo schema di Bernoulli e dunque la formula (2.4). Sostituendo i valori che abbiamo scelto per p e q otteniamo

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{a}{a+b}\right)^k \left(\frac{b}{a+b}\right)^{n-k} = 1,$$

cioè $\displaystyle\sum_{k=0}^{n} \binom{n}{k} \frac{a^k b^{n-k}}{(a+b)^{k+n-k}} = 1$ e raccogliendo $\displaystyle\frac{1}{(a+b)^n} \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = 1.$

Moltiplicando entrambi i termini per $(a+b)^n$ otteniamo infine

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

Esercizio 2.9

Si fanno cinque estrazioni ripetute con reintegro da un'urna contenente 5 biglie rosse, 7 nere e 8 verdi. calcolare le probabilità degli eventi:

A: estrazione di esattamente 3 biglie nere,
B: estrazione di almeno 3 biglie nere.

Svolgimento. Poiché le estrazioni si ripetono dopo aver reinserito le biglie estratte nell'urna, tutte le estrazioni avvengono nelle stesse condizioni e possiamo usare la formula di Bernoulli. Calcoliamo prima le probabilità degli eventi elementari

successo: estrazione di una biglia nera (su un'unica estrazione), $p = \frac{7}{5+7+8} = \frac{7}{20}$,
insuccesso: estrazione di una biglia non nera (su un'unica estrazione), $q = 1 - \frac{7}{20} = \frac{13}{20}$.

L'evento A corrisponde ad avere 3 successi su 5 tentativi ($E_{3,5}$ se usiamo le notazioni che abbiamo introdotto nella dimostrazione del Teorema 2.8), quindi

$$\mathrm{P}(A) = P'_{3,2} \left(\frac{7}{20}\right)^3 \left(\frac{13}{20}\right)^2 = \frac{10 \cdot 7^3 \cdot 13^2}{20^5} \approx 0.18 \quad \text{o} \quad 18\%.$$

L'evento B corrisponde alla somma logica dell'evento $A = E_{3,5}$ con gli eventi $E_{4,5}$ ed $E_{5,5}$. Poiché sono mutuamente esclusivi, si ha $\mathrm{P}(B) = \mathrm{P}(A) + \mathrm{P}(E_{4,5}) + \mathrm{P}(E_{5,5})$ e infine

$$\mathrm{P}(B) = 10\frac{7^3 \cdot 13^2}{20^5} + 5\frac{7^4 \cdot 13}{20^5} + \frac{7^5}{20^5} = \frac{7^3}{20^5}\left(10 \cdot 13^2 + 5 \cdot 7 \cdot 13 + 7^2\right) \approx 0.23 \quad \text{o} \quad 23\%.$$

□

Attenzione! Possiamo usare la formula di Bernoulli solo quando gli eventi sono indipendenti (in questo senso parliamo di esperimento ripetibile). Nel caso di estrazioni ripetute senza reintegro, mancano le condizioni per applicare il Teorema 2.8. Tuttavia il ragionamento che abbiamo usato per la sua dimostrazione ci fa da guida nello svolgimento di esercizi di questo tipo.

Esercizio 2.10

Facendo cinque estrazioni ripetute senza reintegro dalla stessa urna dell'Esercizio 2.9, qual è la probabilità dell'evento

A: estrazione di esattamente 3 biglie nere.

Svolgimento. Come prima, il "successo" nel nostro esperimento corrisponde all'estrazione di una biglia nera. Tuttavia a differenza dell'Esercizio 2.9 la probabilità di successo cambia da un'estrazione all'altra poiché le estrazioni non avvengono tutte nelle stesse condizioni, dato che le palline estratte in precedenza vengono tolte dall'urna. Per comodità schematizziamo con la lettera S il successo (cioè l'estrazione di una biglia nera) e con I l'insuccesso (cioè l'estrazione di una biglia di un altro colore). Calcoliamo, per cominciare, la probabilità dell'evento $SSSII$ (cioè successo nei primi tre turni di estrazione e insuccesso nel quarto e quinto). Ora dobbiamo usare la formula della probabilità composta (2.3)

$$P\{SSSII\} = P(S) \cdot P(S|S) \cdot P(S|SS) \cdot P(I|SSS) \cdot P(I|SSSI) = \frac{7}{20} \cdot \frac{6}{19} \cdot \frac{5}{18} \cdot \frac{13}{17} \cdot \frac{12}{16} \approx 0.0176.$$

Anche qui l'evento A si decompone nella somma logica di tutte quelle "catene di eventi" che corrispondono ad anagrammi della parola $SSSII$, cioè in cui abbiamo sempre 3 successi su 5, ma cambia l'ordine.

Già sappiamo che, complessivamente, sono $P'_{3,2} = \dfrac{5!}{3!2!} = 10$.

Ora però potrebbe sembrare che, cambiando l'ordine dei turni, la probabilità cambi a sua volta. Così non è: a titolo di esempio calcoliamo

$$P\{SISSI\} = P(S)\cdot P(I|S)\cdot P(S|SI)\cdot P(S|SIS)\cdot P(I|SISS) = \frac{7}{20}\cdot\frac{13}{19}\cdot\frac{6}{18}\cdot\frac{5}{17}\cdot\frac{12}{16} = \mathrm{P}\{SSSII\}.$$

Quindi possiamo colncludere che

$$P(A) = P'_{3,2}\cdot \mathrm{P}\{SSSII\} = 10\frac{7\cdot 6\cdot 5\cdot 13\cdot 12}{20\cdot 19\cdot 18\cdot 17\cdot 16} \approx 0.17 \quad \text{o} \quad 17\%.$$

□

Esercizio 2.11: Rifacciamo l'Esercizio 1.7

Calcolare la probabilità di fare ambo e quella di fare terno, giocando tre numeri al lotto sulla ruota di Napoli.

Svolgimento: Per fare un ambo, si deve verificare una stringa di eventi di tipo SSIII. Gli eventi di questo tipo sono $\mathcal{P}_{2,3} = \frac{5!}{3!\,2!} = 10$. Poiché le estrazioni del lotto avvengono senza reintegro, dovremo usare la formula (2.3), dunque

$$\mathrm{P}\{SSIII\} = \frac{3}{90}\frac{2}{89}\frac{87}{88}\frac{86}{87}\frac{85}{86} = \frac{3\cdot 2\cdot 85}{90\cdot 89\cdot 88\cdot} \approx 0.0007.$$

Come nell'Esercizio 2.10 cambiando l'ordine tra successi e insuccesi ottengo sempre la stessa probabilità, quindi concludo che la probabilità di fare ambo è

$$\mathcal{P}_{2,3}\cdot \mathrm{P}\{SSIII\} = 10\frac{17}{23496} = 0.0072\cdots \approx 7\cdot 10^{-3} \quad \text{o} \quad 0.7\%.$$

Per fare terno, invece, si deve verificare una stringa di eventi di tipo SSSII. Gli eventi di questo tipo sono $\mathcal{P}_{3,2} = \frac{5!}{2!\,3!} = 10$ ed ognuno di essi ha probabilità

$$\mathrm{P}\{SSSII\} = \frac{3}{90}\frac{2}{89}\frac{1}{88}1\,1 \approx 0.000008.$$

Dunque la probabilità di fare terno è

$$\mathcal{P}_{3,2}\cdot \mathrm{P}\{SSSII\} \approx 0.000085 \approx 8\cdot 10^{-5} \quad \text{o} \quad 0.008\%.$$

Esercizio 2.12

Tre persone pescano rispettivamente una carta da un mazzo di dieci carte numerate. Ognuno, dopo aver pescato la carta, la rimette nel mazzo.

(i) Qual è la probabilità che peschino tutti la stessa carta?
(ii) Qual è la probabilità che peschino tutti la carta numero 3?
(iii) Qual è la probabilità che due persone peschino la stessa carta e la persona rimanente ne peschi una differente?

2.3. Causa ed effetto: il Teorema di Bayes. Per cominciare illustriamo un'utile applicazione della legge della probabilità composta (2.1): la cosiddetta *legge di disintegrazione o delle probabilità totali*. Premettiamo una definizione

Definizione 2.11: Partizione

Dato uno spazio di probabilità $(\Omega, \mathcal{E}, P)$, una sua partizione è una famiglia di eventi $E_1, E_2, \dots E_n \in \mathcal{E}$ con le proprietà

- *gli eventi sono mutuamente esclusivi, cioè $E_i \wedge E_k = \emptyset$ se $i \neq k$,*
- *la loro somma logica esaurisce lo spazio campione, cioè $E_1 \vee E_2 \vee \cdots E_n = \Omega$.*

Ovviamente se $E_1, E_2, \dots E_n$ formano una partizione allora

$$P(E_1) + P(E_2) \dots + P(E_n) = P(E_1 \vee E_2 \vee \dots E_n) = 1.$$

Proposizione 2.12: Legge di disintegrazione o delle probabilità totali

Sia $E_1, E_2, \dots E_n$ una partizione in uno spazio di probabilità $(\Omega, \mathcal{E}, P)$. Allora per ogni evento $A \in \mathcal{E}$ si ha

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) \dots + P(A|E_n)P(E_n). \tag{2.5}$$

Dimostrazione. Poiché $E_1, E_2, \dots E_n$ esauriscono lo spazio campione, certamente

$$A = (A \wedge E_1) \vee (A \wedge E_2) \dots \vee (A \wedge E_n).$$

D'altra parte, poiché $E_1, E_2, \dots E_n$ sono mutuamente esclusivi, lo sono anche $A \wedge E_1, A \wedge E_2, \dots A \wedge E_n$. Quindi l'assioma *(iii)* ci dice che

$$P(A) = P(A \wedge E_1) + (A \wedge E_2) \dots + P(A \wedge E_n).$$

Ora ci viene in aiuto la legge della probabilità composta (2.1), che afferma che $P(A \wedge E_1) = P(A|E_1)P(E_1)$, e analogamente per tutti gli altri. Sostituendo nella formula precedente otteniamo infine (2.5). □

Osservazione 2.13

Ripercorrendo i passaggi della dimostrazione ci si accorge he non è necessario che $E_1, E_2, \dots E_n$ siano una partizone. In effetti basta

- *$E_1, E_2, \dots E_n$ sono mutuamente esclusivi,*
- *$A \subset E_1 \vee E_2 \dots \vee E_n$.*

Esercizio 2.13

Marco deve fare l'esame di Matematica 2 e vuole sapere che probabilità ha di ricevere una domanda sugli argomenti di probabilità. I professori che interrogano sono due: prof. Tizio chiede probabilità a due studenti su tre, mentre prof. Caio la chiede a uno su due. Marco non sa con quale professore capiterà, però sa che prof. Tizio interroga tre studenti in un'ora mentre prof. Caio ne interroga solo due.

Svolgimento. Chiamiamo X l'evento di cui vogliamo conoscere la probabilità, cioè ricevere la domanda sugli argomenti di probabilità. In questo caso la partizione è data dai due professori Tizio e Caio, che indicheremo con le loro iniziali T e C. Poiché essi interrogano rispettivamente 3 e 2 studenti su 5, a priori sappiamo che $P(T) = \frac{3}{5}$ e $P(C) = \frac{2}{5}$. D'altra parte, la probabilità di ricevere la domanda di probabilità è $\frac{2}{3}$ se si viene interrogati da Tizio e $\frac{1}{2}$ se invece si fa l'esame con Caio. In simboli $P(X|T) = \frac{2}{3}$, $P(X|C) = \frac{1}{2}$. Finalmente possiamo usare la legge di disintegrazione (2.5) e calcolare

$$P(X) = P(X|T)P(T) + P(X|C)P(C) = \frac{2}{3}\frac{3}{5} + \frac{1}{2}\frac{2}{5} = \frac{3}{5} = 0.6 \text{ oppure } 60\%.$$

□

Esercizio 2.14

Se estraggo in successione senza reintegro due biglie da un'urna che contiene 3 biglie rosse, 3 nere e 4 bianche, qual è la probabilità che la seconda biglia estratta sia bianca?

Esercizio 2.15

Abbiamo due sacchetti, il primo contiene 7 biglie nere e 3 bianche, il secondo 4 biglie nere e 6 bianche. Scegliamo casualmente uno dei due sacchetti e peschiamo una biglia. Con quale probabilità è bianca?

Esercizio 2.16

In un sacchetto ci sono 7 gettoni numerati da 1 a 7. Si estrae un primo gettone e se reca un numero dispari viene lasciato fuori, mentre se è pari viene rimesso dentro. Qual è la probabilità che nella seconda estrazione esca un numero pari?

Introduciamo il Teorema di Bayes con un esempio. Abbiamo due urne: un'urna $\mathcal{U}_A$ che contiene 3 palline rosse e 2 verdi, e un'urna $\mathcal{U}_B$ che contiene 4 palline rosse e 5 verdi. Lasciamo al caso anche la scelta dell'urna da cui estrarre la pallina: lanciamo un dato e se esce un numero minore di 3 peschiamo una pallina dall'urna $\mathcal{U}_A$, mentre se esce un numero uguale o maggiore a 3 peschiamo una pallina dall'urna $\mathcal{U}_B$.

Ora mi chiedo: se è stata pescata una pallina rossa, qual è la probabilità che provenga dall'urna $\mathcal{U}_A$? In altri termini, conosciamo l'*effetto* (la pallina è rossa) e vogliamo risalire alla *causa* (l'urna da cui è stata pescata).

Cominciamo a fissare un po' di notazioni. Chiamiamo

A: estrazione dall'urna $\mathcal{U}_A$,
B: estrazione dall'urna $\mathcal{U}_B$,
R: estrazione di una pallina rossa,

Si noti che A e B costituiscono una partizione. Poiché la scelta dell'urna dipende dal lancio del dado, *a priori* (cioè prima che l'esperimento abbia inizio) abbiamo

$$P(A) = \frac{2}{6} = \frac{1}{3} \quad \text{e} \quad P(B) = \frac{4}{6} = \frac{2}{3}.$$

I dati del problema ci forniscono anche la probabilità di R condizionata all'urna, infatti sappiamo che abbiamo 3 possibilità su 5 di avere rosso se estraiamo da A, e 4 su 9 se invece estraiamo da B, in simboli

$$P(R|A) = \frac{3}{5}, \quad \text{e} \quad P(R|B) = \frac{4}{9}.$$

Da questo possiamo dedurre la probabilità a priori di R, grazie alla legge di disintegrazione (2.5)

$$P(R) = P(R|A)P(A) + P(R|B)P(B) = \frac{3}{5}\frac{1}{3} + \frac{4}{9}\frac{2}{3} = \frac{67}{135}.$$

Noi però ora ci stiamo interrogando su tutt'altro quesito: vogliamo conoscere $P(A|R)$ e $P(B|R)$. Facciamo un passo indietro e torniamo alla definizione di probabilità condizionata

$$P(A|R) = \frac{P(A \wedge R)}{P(R)}.$$

Ricaviamo $P(A \wedge R)$ dalla legge della probabilità composta

$$P(A \wedge R) = P(A)P(R|A)$$

e sostituiamo

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

L'equazione che abbiamo appena ricavato è davvero notevole: ci permette di scambiare il ruolo di condizionante (causa) e condizionato (effetto), infatti nel termine di sinistra l'evento R ha il ruolo di condizionante, mentre nel termine di destra compare come evento condizionato. È, nella sostanza,

il Teorema di Bayes.
Concludiamo l'esercizio facendo i calcoli

$$P(A|R) = \frac{\frac{3}{5}\frac{1}{3}}{\frac{3}{5}\frac{1}{3} + \frac{4}{9}\frac{2}{3}} = \frac{\frac{1}{5}}{\frac{67}{135}} = \frac{27}{67} \approx 0.40 \quad \text{o } 40\%.$$

Passiamo ora ad enunciare in piena generalità il Teorema di Bayes.

Teorema 2.14: Teorema di Bayes

Sono dati uno spazio di misura $(\Omega, \mathcal{E}, \mathrm{P})$ *e degli eventi* $E_1, E_2, \ldots, E_n, A \in \mathcal{E}$. *Supponiamo che*

I. gli eventi $E_1, E_2, \ldots, E_n$ *sono mutuamente esclusivi (ovvero* $E_i \wedge E_k = \emptyset$ *se* $i \neq k$*),*
II. l'evento A *è possibile (ovvero* $\mathrm{P}(A) \neq 0$*),*
III. l'evento A *implica almeno uno tra* $E_1, E_2, \ldots, E_n$*, nel senso che* $A \subset \bigcup_{i=1}^{n} E_i$.

Allora per ogni $k = 1 \cdots n$ *si ha*

$$\mathrm{P}(E_k|A) = \frac{\mathrm{P}(A|E_k)\,\mathrm{P}(E_k)}{\sum_{i=1}^{n} \mathrm{P}(A|E_i)\,\mathrm{P}(E_i)}.$$

Dimostrazione. Cominciamo con l'osservare che, per definizione,

$$\mathrm{P}(E_k|A) = \frac{\mathrm{P}(A \wedge E_k)}{\mathrm{P}(A)}. \tag{2.6}$$

Per la legge delle probabilità composte, possiamo riscrivere il numeratore come

$$\mathrm{P}(A \wedge E_k) = \mathrm{P}(A|E_k) \cdot \mathrm{P}(E_k). \tag{2.7}$$

D'altra parte, grazie alle ipotesi *I.* e *III.* e all'Osservazione 2.13 vale la legge di disintegrazione

$$\mathrm{P}(A) = \sum_{i=1}^{n} \mathrm{P}(A|E_i) \cdot \mathrm{P}(E_i). \tag{2.8}$$

Infine, sosituendo (2.7) al posto del numeratore e (2.8) al posto del denominatore in (2.6) otteniamo la tesi. □

L'importanza del Teorema di Bayes non sta tanto nell'elaborazione formale (peraltro, la sua dimostrazione è molto semplice), quanto nelle sue applicazioni alle leggi di induzione statistica. Proviamo ad interpretarlo in questa chiave: la famiglia $E_1, \ldots E_n$ rappresenta un insieme di cause che non si possono mai verificare insieme (sono mutuamente esclusive). Sappiamo che l'effetto A si è verificato: ciò è avvenuto come conseguenza di una (e solo una) delle possibili cause $E_1, \ldots E_n$, ma non sappiamo di quale. Per scoprirlo (in senso probabilistico), dobbiamo calcolare $\mathrm{P}(E_k|A)$, detta anche *verosimiglianza*. Il Teorema di Bayes ci dice che possiamo scoprire quale tra le cause è più verosimile se conosciamo

- le probabilità a priori delle varie cause (cioè tutte le $\mathrm{P}(E_i)$),
- le probabilità a posteriori dell'effetto A, condizionate a ciascuna delle cause (cioè $\mathrm{P}(A|E_i)$)).

Esempio 2.17: Diagnosi medica

Indichiamo con A l'insieme dei sintomi manifestati da un paziente e con $E_1, \ldots E_n$ una famiglia di malattie. Il medico osserva i sintomi A e deve scoprire quale malattia li ha provocati con maggiore probabilità. Gli serve allora calcolare le probabilità condizionate $\mathrm{P}(E_k|A)$, per ogni k. Attraverso il teorema di Bayes, egli può farlo a partire da

- *$\mathrm{P}(E_i)$, le probabilità a priori delle varie malattie, che ricaverà dai dati statistici;*
- *$\mathrm{P}(A|E_i)$, le probabilità a posteriori dei sintomi A, che ricaverà a sua volta dai dati statistici cercando la frazione dei malati di E_i che hanno manifestato i sintomi A.*

Esempio 2.18: Il problema di Monty Hall

Il concorrente di un gioco a premi può scegliere tra tre porte: dietro una di esse c'è un'automobile, dietro le altre, due capre. Sceglie una porta, diciamo la numero 1, e il conduttore del gioco a premi, che sa cosa si nasconde dietro ciascuna porta, ne apre un'altra, diciamo la 3, rivelando una capra. Quindi domanda: "Vorresti scegliere la numero 2?" Conviene cambiare la tua scelta originale?

Si potrebbe pensare che, con due porte chiuse, si abbia una probabilità 1/2 per ognuna, e che quindi non ci sia motivo di cambiare porta. Non è questo il caso.

Chiamiamo l'evento che l'automobile si trovi dietro una certa porta rispettivamente E_1, E_2, e E_3. All'inizio, le nostre probabilità a priori sono

$$\mathrm{P}(E_1) = \mathrm{P}(E_2) = \mathrm{P}(E_3) = 1/3.$$

Come detto prima, la porta scelta è la numero 1. Chiamiamo poi A l'evento "il presentatore apre la porta 3". La sua probabilità priori sarà 1/2 (perché il presentatore può scegliere solo tra le porte 2 e 3). Ora noi sappiamo che si è verificato l'evento A, e vogliamo usare il Teorema di Bayes per valutare a posteriori le probabilità degli eventi E_1, E_2, E_3 (cioè la probabilità che l'automobile sia dietro alla porta 1, 2, 3).

- Se l'auto è dietro la porta 1, il presentatore è libero di scegliere la porta 2 o 3 casualmente. Pertanto, $\mathrm{P}(A|E_1) = 1/2$.
- Se l'auto è dietro la porta 2, il presentatore è obbligato ad aprire la porta 3. Pertanto $\mathrm{P}(A|E_2) = 1$.
- Se l'auto è dietro la porta 3, il presentatore sarà obbligato ad aprire la porta 2. Pertanto $\mathrm{P}(A|E_3) = 0$.

Verifichiamo, tra l' altro, che $\displaystyle\sum_{i=1}^{3} \mathrm{P}(A \wedge E_i) \cdot \mathrm{P}(E_i) = \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} = \frac{1}{2} = \mathrm{P}(A)$, cioè vale la legge di disintegrazione. Il teorema di Bayes afferma che

$$\mathrm{P}(E_1|A) = \frac{\mathrm{P}(A|E_1)\mathrm{P}(E_1)}{\sum_{i=1}^{3} \mathrm{P}(A \wedge E_i) \cdot \mathrm{P}(E_i)} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3},$$

$$\mathrm{P}(E_2|A) = \frac{\mathrm{P}(A|E_2)\mathrm{P}(E_2)}{\sum_{i=1}^{3} \mathrm{P}(A \wedge E_i) \cdot \mathrm{P}(E_i)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3},$$

$$\mathrm{P}(E_3|A) = \frac{\mathrm{P}(A|E_3)\mathrm{P}(E_3)}{\sum_{i=1}^{3} \mathrm{P}(A \wedge E_i) \cdot \mathrm{P}(E_i)} = \frac{0 \cdot \frac{1}{3}}{\frac{1}{2}} = 0.$$

Da ciò è evidente che, a posteriori, la probabilità che l'automobile sia dietro la porta 2 è maggiore di quella che sia dietro alla porta 1. Scegliere la porta 2 dopo che il presentatore ha aperto la porta

3 è equivalente a scegliere inizialmente entrambe le porte 2 e 3.
Se il conduttore non avesse saputo dov'è l'automobile, cosa sarebbe cambiato?

Esercizio 2.19

In un sacchetto ci sono 7 gettoni numerati da 1 a 7. Si estrae un primo gettone e se reca un numero dispari viene lasciato fuori, mentre se è pari viene rimesso dentro. Se nella seconda estrazione esce un numero pari, qual è la probabilità che il primo estratto fosse dispari?

Esercizio 2.20

In un gruppo di 30 persone, 10 sono donne. L'80% delle donne conosce la lingua inglese, mentre gli uomini che la conoscono sono 8. Calcolare la probabilità che una persona scelta a caso che conosce l'inglese sia uomo.

Esercizio 2.21

Un automobilista arriva ad un bivio e chiede indicazione ad una persona a caso. Ci sono due persone a cui può chiedere: A dice la verità 4 volte su 10, B invece 7 volte su 10.

a) Che probabilità ha di percorrere la strada giusta?
b) Se percorre la strada giusta, con che probabilità ha chiesto l'indicazione a B?

Esercizio 2.22

Abbiamo tre urne: la prima contiene 2 palline bianche e 3 rosse, la seconda 5 bianche e 3 rosse e la terza 4 bianche e 2 rosse. Scegliamo a caso un'urna ed estraiamo una pallina. Se la pallina è bianca, che probabilità c'è che provenga dalla seconda urna?

Esercizio 2.23

Un pezzo meccanico viene prodotto da due fabbriche. La fabbrica A produce il 40% del quantitativo totale, e il 98% della sua produzione è senza difetti. Il 7% dei pezzi prodotti dalla fabbrica B è difettoso.

a) Qual è la probabilità totale che un pezzo sia difettoso?
b) Se abbiamo ricevuto un pezzo difettoso, qual è la probabilità che provenga dalla fabbrica B?