

# Psicometria

Corrado Caudek

Questo documento è stato realizzato con:

- $\text{\LaTeX}$  e la classe memoir (<http://www.ctan.org/pkg/memoir>);
- R (<http://www.r-project.org/>) e RStudio (<http://www.rstudio.com/>);
- bookdown (<http://bookdown.org/>) e memoir (<https://ericmarcon.github.io/memoir/>).



Nel blog della mia pagina personale sono forniti alcuni approfondimenti degli argomenti qui trattati.

<https://ccaudek.github.io/caudeklab/>

# Indice

<b>Indice</b>	<b>iii</b>
<b>Prefazione</b>	<b>vii</b>
La psicologia e la Data Science . . . . .	vii
Come studiare . . . . .	viii
Sviluppare un metodo di studio efficace . . . . .	viii
<b>1 Probabilità condizionata</b>	<b>1</b>
1.1 Probabilità condizionata su altri eventi . . . . .	1
La fallacia del pubblico ministero . . . . .	2
1.2 Legge della probabilità composta . . . . .	3
1.3 L'indipendenza stocastica . . . . .	3
Considerazioni conclusive . . . . .	4
<b>Bibliografia</b>	<b>5</b>
<b>Elenco delle figure</b>	<b>7</b>



Copyright © 2022.

Data della versione presente: Dicembre 21, 2021.



# Prefazione

**Data Science per psicologi** contiene il materiale delle lezioni dell'insegnamento di *Psicometria B000286* (A.A. 2021/2022) rivolto agli studenti del primo anno del Corso di Laurea in Scienze e Tecniche Psicologiche dell'Università degli Studi di Firenze.

L'insegnamento di Psicometria si propone di fornire agli studenti un'introduzione all'analisi dei dati in psicologia. Le conoscenze/competenze che verranno sviluppate in questo insegnamento sono quelle della *Data science*, ovvero le conoscenze/competenze che si pongono all'intersezione tra statistica (ovvero, richiedono la capacità di comprendere teoremi statistici) e informatica (ovvero, richiedono la capacità di sapere utilizzare un software).

## La psicologia e la Data Science

*It's worth noting, before getting started, that this material is hard. If you find yourself confused at any point, you are normal. Any sense of confusion you feel is just your brain correctly calibrating to the subject matter. Over time, confusion is replaced by comprehension [...]*

— Richard McElreath

Sembra sensato spendere due parole su un tema che è importante per gli studenti: quello indicato dal titolo di questo Capitolo. È ovvio che agli studenti di psicologia la statistica non piace. Se piacesse, forse studierebbero Data Science e non psicologia; ma non lo fanno. Di conseguenza, gli studenti di psicologia si chiedono: “perché dobbiamo perdere tanto tempo a studiare queste cose quando in realtà quello che ci interessa è tutt'altro?” Questa è una bella domanda.

C'è una ragione molto semplice che dovrebbe farci capire perché la Data Science è così importante per la psicologia. Infatti, a ben pensarci, la psicologia è una disciplina intrinsecamente statistica, se per statistica intendiamo quella disciplina che studia la variazione delle caratteristiche degli individui nella popolazione. La psicologia studia *gli individui* ed è proprio la variabilità inter- e intra-individuale ciò che vogliamo descrivere e, in certi casi, predire. In questo senso, la psicologia è molto diversa dall'ingegneria, per esempio. Le proprietà di un determinato ponte sotto certe condizioni, ad esempio, sono molto simili a quelle di un altro ponte, sotto le medesime condizioni. Quindi, per un ingegnere la statistica è poco importante: le proprietà dei materiali sono unicamente dipendenti dalla loro composizione e restano costanti. Ma lo stesso non può dirsi degli individui: ogni individuo è unico e cambia nel tempo. E le variazioni tra gli individui, e di un individuo nel tempo, sono l'oggetto di studio proprio della psicologia: è dunque chiaro che i problemi che la psicologia si pone sono molto diversi da quelli affrontati, per esempio, dagli ingegneri. Questa è la ragione per cui abbiamo tanto bisogno della *data science* in psicologia: perché la *data science* ci consente di descrivere la variazione e il cambiamento. E queste sono appunto le caratteristiche di base dei fenomeni psicologici.

Sono sicuro che, leggendo queste righe, a molti studenti sarà venuta in mente la seguente domanda: perché non chiediamo a qualche esperto di fare il “lavoro sporco” (ovvero le analisi statistiche) per noi, mentre noi (gli psicologi) ci occupiamo solo di ciò

che ci interessa, ovvero dei problemi psicologici slegati dai dettagli “tecnici” della *data science*? La risposta a questa domanda è che non è possibile progettare uno studio psicologico sensato senza avere almeno una comprensione rudimentale della *data science*. Le tematiche della *data science* non possono essere ignorate né dai ricercatori in psicologia né da coloro che svolgono la professione di psicologo al di fuori dell’Università. Infatti, anche i professionisti al di fuori dall’università non possono fare a meno di leggere la letteratura psicologica più recente: il continuo aggiornamento delle conoscenze è infatti richiesto dalla deontologia della professione. Ma per potere fare questo è necessario conoscere un bel po’ di *data science*! Basta aprire a caso una rivista specialistica di psicologia per rendersi conto di quanto ciò sia vero: gli articoli che riportano i risultati delle ricerche psicologiche sono zeppi di analisi statistiche e di modelli formali. E la comprensione della letteratura psicologica rappresenta un requisito minimo nel bagaglio professionale dello psicologo.

Le considerazioni precedenti cercano di chiarire il seguente punto: la *data science* non è qualcosa da studiare a malincuore, in un singolo insegnamento universitario, per poi poterla tranquillamente dimenticare. Nel bene e nel male, gli psicologi usano gli strumenti della *data science* in tantissimi ambiti della loro attività professionale: in particolare quando costruiscono, somministrano e interpretano i test psicometrici. È dunque chiaro che possedere delle solide basi di *data science* è un tassello imprescindibile del bagaglio professionale dello psicologo. In questo insegnamento verranno trattati i temi base della *data science* e verrà adottato un punto di vista bayesiano, che corrisponde all’approccio più recente e sempre più diffuso in psicologia.

## Come studiare

*I know quite certainly that I myself have no special talent. Curiosity, obsession and dogged endurance, combined with self-criticism, have brought me to my ideas.*

— Albert Einstein

Il giusto metodo di studio per prepararsi all’esame di Psicometria è quello di seguire attivamente le lezioni, assimilare i concetti via via che essi vengono presentati e verificare in autonomia le procedure presentate a lezione. Incoraggio gli studenti a farmi domande per chiarire ciò che non è stato capito appieno. Incoraggio gli studenti a utilizzare i forum attivi su Moodle e, soprattutto, a svolgere gli esercizi proposti su Moodle. I problemi forniti su Moodle rappresentano il livello di difficoltà richiesto per superare l’esame e consentono allo studente di comprendere se le competenze sviluppate fino a quel punto sono sufficienti rispetto alle richieste dell’esame.

La prima fase dello studio, che è sicuramente individuale, è quella in cui è necessario acquisire le conoscenze teoriche relative ai problemi che saranno presentati all’esame. La seconda fase di studio, che può essere facilitata da scambi con altri e da incontri di gruppo, porta ad acquisire la capacità di applicare le conoscenze: è necessario capire come usare un software (R) per applicare i concetti statistici alla specifica situazione del problema che si vuole risolvere. Le due fasi non sono però separate: il saper fare molto spesso ci aiuta a capire meglio.

## Sviluppare un metodo di studio efficace

*Memorization is not learning.*

— Richard Phillips Feynman

Avendo insegnato molte volte in passato un corso introduttivo di analisi dei dati ho notato nel corso degli anni che gli studenti con l’atteggiamento mentale che descriverò qui sotto generalmente ottengono ottimi risultati. Alcuni studenti sviluppano naturalmente



---

questo approccio allo studio, ma altri hanno bisogno di fare uno sforzo per maturarlo. Fornisco qui sotto una breve descrizione del “metodo di studio” che, nella mia esperienza, è il più efficace per affrontare le richieste di questo insegnamento (Burger & Starbird, 2012).

- Dedicate un tempo sufficiente al materiale di base, apparentemente facile; assicuratevi di averlo capito bene. Cercate le lacune nella vostra comprensione. Leggere presentazioni diverse dello stesso materiale (in libri o articoli diversi) può fornire nuove intuizioni.
- Gli errori che facciamo sono i nostri migliori maestri. Istintivamente cerchiamo di dimenticare subito i nostri errori. Ma il miglior modo di imparare è apprendere dagli errori che commettiamo. In questo senso, una soluzione corretta è meno utile di una soluzione sbagliata. Quando commettiamo un errore questo ci fornisce un’informazione importante: ci fa capire qual è il materiale di studio sul quale dobbiamo ritornare e che dobbiamo capire meglio.
- C’è ovviamente un aspetto “psicologico” nello studio. Quando un esercizio o problema ci sembra incomprensibile, la cosa migliore da fare è dire: “mi arrendo”, “non ho idea di cosa fare!”. Questo ci rilassa: ci siamo già arresi, quindi non abbiamo niente da perdere, non dobbiamo più preoccuparci. Ma non dobbiamo fermarci qui. Le cose “migliori” che faccio (se ci sono) le faccio quando non ho voglia di lavorare. Alle volte, quando c’è qualcosa che non so fare e non ho idea di come affrontare, mi dico: “oggi non ho proprio voglia di fare fatica”, non ho voglia di mettermi nello stato mentale per cui “in 10 minuti devo risolvere il problema perché dopo devo fare altre cose”. Però ho voglia di *divertirmi* con quel problema e allora mi dedico a qualche aspetto “marginale” del problema, che so come affrontare, oppure considero l’aspetto più difficile del problema, quello che non so come risolvere, ma invece di cercare di risolverlo, guardo come altre persone hanno affrontato problemi simili, oppure lo stesso problema in un altro contesto. Non mi pongo l’obiettivo “risolvi il problema in 10 minuti”, ma invece quello di farmi un’idea “generale” del problema, o quello di capire un caso più specifico e più semplice del problema. Senza nessuna pressione. Infatti, in quel momento ho deciso di non lavorare (ovvero, di non fare fatica). Va benissimo se “parto per la tangente”, ovvero se mi metto a leggere del materiale che sembra avere poco a che fare con il problema centrale (le nostre intuizioni e la nostra curiosità solitamente ci indirizzano sulla strada giusta). Quando faccio così, molto spesso trovo la soluzione del problema che mi ero posto e, paradossalmente, la trovo in un tempo minore di quello che, in precedenza, avevo dedicato a “lavorare” al problema. Allora perché non faccio sempre così? C’è ovviamente l’aspetto dei “10 minuti” che non è sempre facile da dimenticare. Sotto pressione, possiamo solo agire in maniera automatica, ovvero possiamo solo applicare qualcosa che già sappiamo fare. Ma se dobbiamo imparare qualcosa di nuovo, la pressione è un impedimento.
- È utile farsi da soli delle domande sugli argomenti trattati, senza limitarsi a cercare di risolvere gli esercizi che vengono assegnati. Quando studio qualcosa mi viene in mente: “se questo è vero, allora deve succedere quest’altra cosa”. Allora verifico se questo è vero, di solito con una simulazione. Se i risultati della simulazione sono quelli che mi aspetto, allora vuol dire che ho capito. Se i risultati sono diversi da quelli che mi aspettavo, allora mi rendo conto di non avere capito e ritorno indietro a studiare con più attenzione la teoria che pensavo di avere capito – e ovviamente mi rendo conto che c’era un aspetto che avevo frainteso. Questo tipo di verifica è qualcosa che dobbiamo fare da soli, in prima persona: nessun altro può fare questo al posto nostro.

- Non aspettatevi di capire tutto la prima volta che incontrate un argomento nuovo.<sup>1</sup> È utile farsi una nota mentalmente delle lacune nella vostra comprensione e tornare su di esse in seguito per cercare di colmarle. L'atteggiamento naturale, quando non capiamo i dettagli di qualcosa, è quello di pensare: "non importa, ho capito in maniera approssimativa questo punto, non devo preoccuparmi del resto". Ma in realtà non è vero: se la nostra comprensione è superficiale, quando il problema verrà presentato in una nuova forma, non riusciremo a risolverlo. Per cui i dubbi che ci vengono quando studiamo qualcosa sono il nostro alleato più prezioso: ci dicono esattamente quali sono gli aspetti che dobbiamo approfondire per potere migliorare la nostra preparazione.
- È utile sviluppare una visione d'insieme degli argomenti trattati, capire l'obiettivo generale che si vuole raggiungere e avere chiaro il contributo che i vari pezzi di informazione forniscono al raggiungimento di tale obiettivo. Questa organizzazione mentale del materiale di studio facilita la comprensione. È estremamente utile creare degli schemi di ciò che si sta studiando. Non aspettate che sia io a fornirvi un riepilogo di ciò che dovete imparare: sviluppate da soli tali schemi e tali riassunti.
- Tutti noi dobbiamo imparare l'arte di trovare le informazioni, non solo nel caso di questo insegnamento. Quando vi trovate di fronte a qualcosa che non capite, o ottenete un oscuro messaggio di errore da un software, ricordatevi: "Google is your friend".

Corrado Caudek

Febbraio 2022

---

<sup>1</sup>Ricordatevi inoltre che gli individui tendono a sottostimare la propria capacità di apprendere (Horn & Loewenstein, 2021).

## Capitolo 1

# Probabilità condizionata



L'attribuzione di una probabilità ad un evento è sempre condizionata dalle conoscenze che abbiamo a disposizione. Per un determinato stato di conoscenze, attribuiamo ad un dato evento una certa probabilità di verificarsi; ma se il nostro stato di conoscenze cambia, allora cambierà anche la probabilità che attribuiamo all'evento in questione.

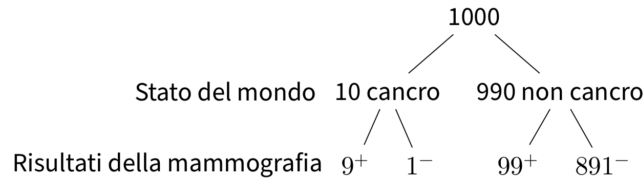
### 1.1 Probabilità condizionata su altri eventi

La probabilità condizionata è una componente essenziale del ragionamento scientifico dato che chiarisce come sia possibile incorporare le evidenze disponibili, in maniera logica e coerente, nella nostra conoscenza del mondo. Infatti, si può pensare che tutte le probabilità siano probabilità condizionate, anche se l'evento condizionante non è sempre esplicitamente menzionato. Consideriamo il seguente problema.

**Esempio 1.1.** Supponiamo che lo screening per la diagnosi precoce del tumore mammario si avvalga di test che sono accurati al 90%, nel senso che il 90% delle donne con cancro e il 90% delle donne senza cancro saranno classificate correttamente. Supponiamo che l'1% delle donne sottoposte allo screening abbia effettivamente il cancro al seno. Ci chiediamo: qual è la probabilità che una donna scelta casualmente abbia una mammografia positiva e, se ce l'ha, qual è la probabilità che abbia davvero il cancro?

Per risolvere questo problema, supponiamo che il test in questione venga somministrato ad un grande campione di donne, diciamo a 1000 donne. Di queste 1000 donne, 10 (ovvero, l'1%) hanno il cancro al seno. Per queste 10 donne, il test darà un risultato positivo in 9 casi (ovvero, nel 90% dei casi). Per le rimanenti 990 donne che non hanno il cancro al seno, il test darà un risultato positivo in 99 casi (se la probabilità di un vero positivo è del 90%, la probabilità di un falso positivo è del 10%). Questa situazione è rappresentata nella figura 1.1. Combinando questi due risultati, vediamo che il test dà un risultato positivo per 9 donne che hanno effettivamente il cancro al seno e per 99 donne che non ce l'hanno, per un totale di 108 risultati positivi. Dunque, la probabilità di ottenere un risultato positivo al test è  $\frac{108}{1000} = 11\%$ . Ma delle 108 donne che hanno ottenuto un risultato positivo al test, solo 9 hanno il cancro al seno. Dunque, la probabilità di avere il cancro, dato un risultato positivo al test, è pari a  $\frac{9}{108} = 8\%$ .

Nell'esercizio precedente, la probabilità dell'evento "ottenere un risultato positivo al test" è una probabilità non condizionata, mentre la probabilità dell'evento "avere il cancro al seno, dato che il test ha dato un risultato positivo" è una probabilità condizionata.



**Figura 1.1:** Rappresentazione ad albero che riporta le frequenze attese dei risultati di una mammografia in un campione di 1,000 donne.

In termini generali, la probabilità condizionata  $P(A | B)$  rappresenta la probabilità che si verifichi l'evento  $A$  sapendo che si è verificato l'evento  $B$  (oppure: la probabilità di  $A$  in una prova valida solo se si verifica anche  $B$ ). Ciò ci conduce alla seguente definizione.

**Definizione 1.1.** Dato un qualsiasi evento  $A$ , si chiama *probabilità condizionata* di  $A$  dato  $B$  il numero

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad \text{con } P(B) > 0, \quad (1.1)$$

dove  $P(A \cap B)$  è la probabilità congiunta dei due eventi, ovvero la probabilità che si verifichino entrambi.

In alcuni casi può essere conveniente leggere al contrario la 1.1 e utilizzarla per calcolare la probabilità dell'intersezione di due eventi. Per esempio se conosciamo la probabilità dell'evento  $B$  e la probabilità condizionata di  $A$  su  $B$ , otteniamo

$$P(A \cap B) = P(B)P(A | B), \quad (1.2)$$

mentre se conosciamo la probabilità dell'evento  $A$  e la probabilità condizionata di  $B$  su  $A$ , otteniamo

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

**Esempio 1.2.** Da un mazzo di 52 carte (13 carte per ciascuno dei 4 semi) ne viene estratta 1 in modo casuale. Qual è la probabilità che esca una figura di cuori? Sapendo che la carta estratta ha il seme di cuori, qual è la probabilità che il valore numerico della carta sia 7, 8 o 9?

Ci sono 13 carte di cuori, dunque la risposta alla prima domanda è  $1/4$ . Per rispondere alla seconda domanda consideriamo solo le 13 carte di cuori; la probabilità cercata è dunque  $3/13$ .

## La fallacia del pubblico ministero

Un errore comune che si commette è quello di credere che  $P(A | B)$  sia uguale a  $P(B | A)$ . Tale fallacia ha particolare risalto in ambito forense tanto che è conosciuta con il nome di “fallacia del procuratore” (*prosecutor's fallacy*). In essa, una piccola probabilità dell'evidenza, data l'innocenza, viene erroneamente interpretata come la probabilità dell'innocenza, data l'evidenza.

Consideriamo il caso di un esame del DNA. Un esperto forense potrebbe affermare, ad esempio, che “se l'imputato è innocente, c'è solo una possibilità su un miliardo che vi sia una corrispondenza tra il suo DNA e il DNA trovato sulla scena del crimine”. Ma talvolta questa probabilità è erroneamente interpretata come avesse il seguente significato: “date le prove del DNA, c'è solo una possibilità su un miliardo che l'imputato sia innocente”.

Le considerazioni precedenti risultano più chiare se facciamo nuovamente riferimento all'esercizio sul tumore mammario descritto sopra. In tale esercizio abbiamo visto come la probabilità di cancro dato un risultato positivo al test sia uguale a 0.08. Tale probabilità è molto diversa dalla probabilità di un risultato positivo al test data la presenza del cancro. Infatti, questa seconda probabilità è uguale a 0.90 ed è descritta nel problema come una delle caratteristiche del test in questione.

## 1.2 Legge della probabilità composta

Il teorema della probabilità composta deriva dal concetto di probabilità condizionata per cui la probabilità che si verifichino due eventi  $A_i$  e  $A_j$  è pari alla probabilità di uno dei due eventi moltiplicato con la probabilità dell'altro evento condizionato al verificarsi del primo.

L'equazione (1.2) si estende al caso di  $n$  eventi  $A_1, \dots, A_n$  nella forma seguente:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) \quad (1.3)$$

la quale esprime in forma generale la legge della probabilità composta.

**Esempio 1.3.** Da un'urna contenente 6 palline bianche e 4 nere si estrae una pallina per volta, senza reintrodurla nell'urna. Indichiamo con  $B_i$  l'evento: "esce una pallina bianca alla  $i$ -esima estrazione" e con  $N_i$  l'estrazione di una pallina nera. L'evento: "escono due palline bianche nelle prime due estrazioni" è rappresentato dalla intersezione  $\{B_1 \cap B_2\}$  e la sua probabilità vale, per la (1.2)

$$P(B_1 \cap B_2) = P(B_1)P(B_2 | B_1).$$

$P(B_1)$  vale 6/10, perché nella prima estrazione  $\Omega$  è costituito da 10 elementi: 6 palline bianche e 4 nere. La probabilità condizionata  $P(B_2 | B_1)$  vale 5/9, perché nella seconda estrazione, se è verificato l'evento  $B_1$ , lo spazio campionario consiste di 5 palline bianche e 4 nere. Si ricava pertanto:

$$P(B_1 \cap B_2) = \frac{6}{10} \cdot \frac{5}{9} = \frac{1}{3}.$$

In modo analogo si ha che

$$P(N_1 \cap N_2) = P(N_1)P(N_2 | N_1) = \frac{4}{10} \cdot \frac{3}{9} = \frac{4}{30}.$$

Se l'esperimento consiste nell'estrazione successiva di 3 palline, la probabilità che queste siano tutte bianche vale, per la (1.3):

$$P(B_1 \cap B_2 \cap B_3) = P(B_1)P(B_2 | B_1)P(B_3 | B_1 \cap B_2),$$

dove la probabilità  $P(B_3 | B_1 \cap B_2)$  si calcola supponendo che si sia verificato l'evento condizionante  $\{B_1 \cap B_2\}$ . Lo spazio campionario per questa probabilità condizionata è costituito da 4 palline bianche e 4 nere, per cui  $P(B_3 | B_1 \cap B_2) = 1/2$  e quindi:

$$P(B_1 \cap B_2 \cap B_3) = \frac{6}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} = \frac{1}{6}.$$

La probabilità dell'estrazione di tre palline nere è invece:

$$\begin{aligned} P(N_1 \cap N_2 \cap N_3) &= P(N_1)P(N_2 | N_1)P(N_3 | N_1 \cap N_2) \\ &= \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{1}{30}. \end{aligned}$$

## 1.3 L'indipendenza stocastica

Un concetto molto importante per le applicazioni statistiche della probabilità è quello dell'indipendenza stocastica. La definizione (1.1) esprime il concetto intuitivo di indipendenza di un evento da un altro, nel senso che il verificarsi di  $A$  non influisce sulla probabilità del verificarsi di  $B$ , ovvero non la condiziona. Infatti, per la definizione (1.1) di probabilità condizionata, si ha che, se  $A$  e  $B$  sono due eventi indipendenti, risulta:

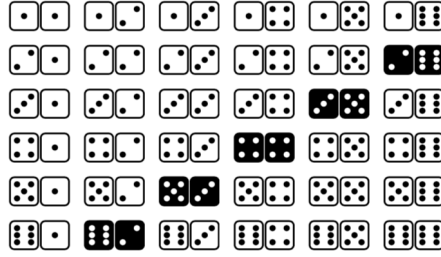
$$P(A | B) = \frac{P(A)P(B)}{P(B)} = P(A).$$

Possiamo dunque dire che due eventi  $A$  e  $B$  sono indipendenti se

$$\begin{aligned} P(A | B) &= P(A), \\ P(B | A) &= P(B). \end{aligned}$$

**Esempio 1.4.** Nel lancio di due dadi non truccati, si considerino gli eventi:  $A = \{\text{esce un 1 o un 2 nel primo lancio}\}$  e  $B = \{\text{il punteggio totale è 8}\}$ . Gli eventi  $A$  e  $B$  sono indipendenti?

Rappresentiamo qui sotto lo spazio campionario dell'esperimento casuale.



**Figura 1.2:** Rappresentazione dello spazio campionario dei risultati dell'esperimento casuale corrispondente al lancio di due dadi bilanciati. Sono evidenziati gli eventi elementari che costituiscono l'evento  $A$ : esce un 1 o un 2 nel primo lancio.

Gli eventi  $A$  e  $B$  non sono statisticamente indipendenti. Infatti, le loro probabilità valgono  $P(A) = 12/36$  e  $P(B) = 5/36$  e la probabilità della loro intersezione è

$$P(A \cap B) = 1/36 = 3/108 \neq P(A)P(B) = 5/108.$$

Si noti che il concetto di indipendenza è del tutto differente da quello di incompatibilità. Due eventi  $A$  e  $B$  incompatibili (per i quali si ha  $A \cap B = \emptyset$ ) sono statisticamente dipendenti, poiché il verificarsi dell'uno esclude il verificarsi dell'altro:  $P(A \cap B) = 0 \neq P(A)P(B)$ .

Si noti inoltre che, se due eventi con probabilità non nulla sono statisticamente indipendenti, la legge delle probabilità totali espressa dalla (1.4)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.4)$$

si modifica nella relazione seguente:

$$P(A \cup B) = P(A) + P(B) - P(A)P(B). \quad (1.5)$$

## Considerazioni conclusive

La probabilità condizionata è importante perché ci fornisce uno strumento per precisare il concetto di indipendenza statistica. Una delle domande più importanti delle analisi statistiche è infatti quella che si chiede se due variabili sono associate tra loro oppure no. In questo Capitolo abbiamo discusso il concetto di indipendenza (come contrapposto al concetto di associazione – si veda il Capitolo @ref(#chapter-descript)). In seguito vedremo come sia possibile fare inferenza sull'associazione tra variabili.

# Bibliografia

- Burger, E. B. & Starbird, M. (2012). *The 5 elements of effective thinking*. Princeton University Press. (Cit. a p. ix).
- Horn, S. & Loewenstein, G. (2021). Underestimating Learning by Doing. *Available at SSRN 3941441* (cit. a p. x).





## Elenco delle figure

1.1	Rappresentazione ad albero che riporta le frequenze attese dei risultati di una mammografia in un campione di 1,000 donne. . . . .	2
1.2	Rappresentazione dello spazio campionario dei risultati dell'esperimento casuale corrispondente al lancio di due dadi bilanciati. Sono evidenziati gli eventi elementari che costituiscono l'evento A: esce un 1 o un 2 nel primo lancio. . . . .	4

**Abstract** This document contains the material of the lessons of Psicometria B000286 (2021/2022) aimed at students of the first year of the Degree Course in Psychological Sciences and Techniques of the University of Florence, Italy.

**Keywords** Data science, Bayesian statistics.