SAPIENZA Università di Roma
Anno Accademico 2017-18

Facoltà di Scienze Matematiche Fisiche e Naturali
Corso di Laurea Triennale in Matematica

INTRODUZIONE AL
CALCOLO DELLE PROBABILITÀ

versione del 12 aprile 2018 (Lezioni 1-7)

Fabio Spizzichino e Giovanna Nappo

A.A. 2017-18

versione del 12 aprile 2018

Indice

Introduzione

L'introduzione sarà scritta in seguito.

Si tratta di una prima parte di Appunti attualmente in fase di revisione.

Per segnalazioni di eventuali sviste, errori di stampa, incoerenza delle notazioni, o anche per suggerimenti si prega di inviare un messaggio a
Giovanna Nappo: e-mail `nappo@mat.uniroma1.it`
e
Fabio Spizzichino: e-mail `spizzichino@mat.uniroma1.it`

1 Fenomeni aleatori; spazio dei risultati elementari di un esperimento

1.1 Operazioni logiche su eventi e interpretazione insiemistica

Iniziamo con una discussione euristica mirante a giustificare la definizione della nozione di *spazio finito di probabilità*, che verrà data nella prossima lezione.

Come punto di partenza, pensiamo ad un esperimento che possa dar luogo a diversi risultati possibili. I risultati verranno chiamati "eventi".

Possiamo vedere un evento come una proposizione relativa al modo di risultare di tale esperimento.

Esempio 1.1. *Consideriamo l'esperimento consistente nell'osservazione dei punti ottenuti dal lancio di una coppia di dadi a sei facce.*

Indichiamo tali punti con i simboli X_1, X_2.

Esempi di possibili eventi sono: $A \equiv \{X_1 \leq X_2\}, B \equiv \{X_1 + X_2 \text{ pari}\}, C \equiv \{X_1 > 3\}, \ldots$

Indichiamo, per il momento, con il simbolo $\mathcal{E}$ la famiglia dei possibili eventi distinti in un esperimento. Come è facile rendersi conto (e verificheremo presto), la famiglia $\mathcal{E}$ costituita da tutti gli eventi nel precedente Esempio 1.1 è una famiglia finita. In quanto immediatamente segue, ci limiteremo ancora a considerare esperimenti per cui $\mathcal{E}$ è una famiglia finita; successivamente, tale limitazione verrà eliminata.

È anche facile rendersi conto che, all'interno della famiglia $\mathcal{E}$, è naturale introdurre le operazioni di *somma logica* (oppure *or*), di *prodotto logico* (oppure *and*) e di *negazione* (oppure *not*), che verranno rispettivamente indicate (per il momento) con i simboli $\bigvee, \bigwedge, \widetilde{\cdot}$; siano E_1, E_2, E eventi appartenenti ad $\mathcal{E}$, allora la somma logica $E_1 \bigvee E_2$ coincide con l'evento:

$$\{\textit{si è verificato almeno uno dei due eventi } E_1 \textit{ ed} E_2\}$$

il prodotto logico $E_1 \bigwedge E_2$ coincide con l'evento:

$$\{\textit{si sono verificati entrambi i due eventi } E_1 \textit{ ed } E_2\}$$

la negazione $\widetilde{E}$ coincide con l'evento:

$$\{\textit{non si è verificato l'evento } E\}.$$

Definizione 1.1. *Un evento* $E \in \mathcal{E}$ *si dice* **composto** *se esistono almeno due eventi* $E_1, E_2 \in \mathcal{E}$*, tali che*

$$E = E_1 \bigvee E_2, E \neq E_1, E \neq E_2.$$

Un evento che non sia composto si dice **semplice** *o* **elementare**.

Esempio 1.2. *Nell'esperimento del lancio di un dado a sei facce, l'evento $E = \{X_1 > 3\}$ è un evento composto. In tale esperimento gli eventi semplici sono dati da*

$$\{X_1 = 1\},\ \{X_1 = 2\}, ..., \{X_1 = 6\},$$

e l'evento E si riscrive come $\{X_1 = 4\} \bigvee \{X_1 = 5\} \bigvee \{X_1 = 6\}$.

Nell'esperimento del lancio di una coppia di dadi gli eventi semplici sono invece quelli del tipo

$$\{X_1 = h,\, X_2 = k\} \qquad h = 1, 2, ..., 6;\ k = 1, 2, ..., 6$$

ed un evento del tipo $\{X_1 = h\}$ risulta essere un evento composto, in quanto possiamo scrivere

$$\{X_1 = h\} = \bigvee_{k=1}^{6} \{X_1 = h, X_2 = k\}.$$

Osservazione 1. È facile verificare che $E \in \mathcal{E}$ composto si decompone in uno ed in un sol modo (a meno dell'ordine) come somma logica di un numero finito di eventi elementari.

Indichiamo ora con i simboli $\omega_1, ..., \omega_N$ gli eventi elementari in un esperimento.

Definizione 1.2. *L'insieme $\Omega \equiv \{\omega_1, ..., \omega_N\}$ che ha come punti gli eventi elementari di un esperimento viene detto* **spazio campione***, per quell'esperimento.*

Indichiamo con il simbolo $\mathcal{P}(\Omega)$ la ***famiglia delle parti*** di Ω (ossia la famiglia dei sottoinsiemi di Ω) e, per $E \in \mathcal{P}(\Omega)$, indichiamo con $|E|$ la ***cardinalità*** di E.

Esempio 1.3. *Un'urna inizialmente contiene quattro oggetti numerati da 1 a 4. Vuotiamo l'urna facendo quattro successive estrazioni senza reinserimento, osservando di volta in volta il numero indicato sull'oggetto estratto.*

Si ha $\Omega = \big\{$permutazioni[1] di $\{1, 2, 3, 4\}\big\}$*, e* $|\Omega| = 24$.

Noi vogliamo analizzare quei casi in cui vi sia una *situazione di incertezza* (cioè di mancanza di completa informazione) circa il modo di risultare dell'esperimento stesso. Ciò significa che non sappiamo a priori quale effettivamente si realizzerà fra i diversi risultati elementari possibili. In tali casi parleremo di fenomeni aleatori o di *esperimenti aleatori*.

Parleremo dunque di esperimento aleatorio quando non sappiamo quali eventi saranno verificati e quali risulteranno falsi.

In tale ambito, un evento composto E è verificato se e solo se si verifica un evento elementare ω_i che si presenti nella decomposizione di E.

Esempio 1.4. *Si lancia un dado a sei facce; si ha $\Omega = \{\omega_1, ..., \omega_6\}$. Supponiamo si verifichi ω_4, allora saranno anche verificati, ad esempio, gli eventi composti: $\{X \leq 5\}, \{X \text{ pari}\}, \{X > 2\}$ e non sono verificati, ad esempio, gli eventi $\{X > 5\}, \{X \text{ dispari}\}$, $\{X \leq 3\}$, $\{X \text{ numero primo}\}, ...$.*

[1] Per la definizione di permutazione vedere più avanti la Lezione 3.2

Dati due eventi $E_1, E_2 \in \mathcal{E}$, dunque,

(a) $E_1 \bigvee E_2$ si verifica se e solo se è verificato un evento elementare ω_i che si presenti nella decomposizione di E_1 e/oppure di E_2

(b) $E_1 \bigwedge E_2$ si verifica se e solo se è verificato un evento elementare ω_i che si presenti sia nella decomposizione di E_1 che in quella di E_2

(c) $\widetilde{E}_1$ si verifica se e solo se è verificato un evento elementare ω_i che non sia presente nella decomposizione di E_1

Osservazione 2 **(Eventi come sottoinsiemi di** Ω**).** Per definizione, i punti dello spazio Ω sono gli "eventi semplici" o "risultati elementari" dell'esperimento considerato.

Notiamo ora che sussiste una corrispondenza biunivoca fra sottoinsiemi di Ω, costituiti da più di un elemento, e gli eventi composti: basta infatti associare, ad un evento composto, l'insieme costituito dagli eventi semplici che lo compongono; viceversa ad un sottoinsieme di Ω possiamo associare l'evento composto che si ottiene come somma logica degli elementi (eventi semplici) in esso contenuti.

Ad un evento semplice $\omega_i \in \Omega$, facciamo corrispondere il singleton $\{\omega_i\} \in \mathcal{P}(\Omega)$.

Dato un evento $E \in \mathcal{E}$, indichiamo per comodità con $\mathcal{H}(E)$ il sottoinsieme di Ω individuato secondo quanto appena detto.

Osservazione 3 **(Operazioni su eventi e operazioni su sottoinsiemi).** Consideriamo di nuovo la corrispondenza biunivoca $\mathcal{H}$ fra eventi e sottoinsiemi di Ω, stabilita nella precedente ***Osservazione 2***:

$$\mathcal{E} \overset{\mathcal{H}}{\Leftrightarrow} \mathcal{P}(\Omega).$$

Ci si rende facilmente conto, da quanto detto sopra, che, in tale corrispondenza biunivoca fra eventi e sottoinsiemi, le operazioni $\bigvee, \bigwedge, \widetilde{\cdot}$ (definite su $\mathcal{E}$) vengono rispettivamente trasformate nelle operazioni booleane di unione $\cup$, di intersezione $\cap$, e di passaggio al complementare $\complement$ (definite su $\mathcal{P}(\Omega)$, la famiglia delle parti di Ω); infatti, traducendo "in formule" i precedenti punti (a), (b) e (c) potremo scrivere, per degli arbitrari $E_1,\ E_2,\ E \in \mathcal{E}$,

$$\mathcal{H}\left(E_1 \bigvee E_2\right) = \mathcal{H}(E_1) \cup \mathcal{H}(E_2),$$

$$\mathcal{H}\left(E_1 \bigwedge E_2\right) = \mathcal{H}(E_1) \cap \mathcal{H}(E_2),$$

$$\mathcal{H}\left(\widetilde{E}\right) = \complement(\mathcal{H}(E)).$$

Da questo momento in poi, quindi, potremo identificare "eventi" e sottoinsiemi di Ω e dunque lasceremo cadere l'uso dei simboli $\mathcal{E}, \bigvee, \bigwedge, \widetilde{\cdot}, \mathcal{H}(\cdot)$; continueremo la trattazione utilizzando solo le nozioni di sottoinsieme di Ω e di operazioni booleane fra sottoinsiemi.

Dovremo però continuare ad aver presente il significato di tipo "logico" che stiamo dando a tali nozioni, nel contesto dell'analisi di fenomeni aleatori. In tale ambito, risulterà naturale attribuire un'interpretazione di tipo "logico" a varie semplici nozioni di tipo insiemistico; a tale proposito vediamo intanto lo specchietto presentato qui di seguito.

Interpretazione "logica" di nozioni di tipo insiemistico:

$\cdot$ $A \subseteq B$ significa che ogni evento elementare che rende verificato A rende verificato anche B e dunque interpretiamo la relazione $A \subseteq B$ come "***A implica B***"

$\cdot\cdot$ Ω è un evento vero qualunque evento elementare si verifichi, in quanto esso contiene tutti gli eventi elementari e dunque interpretiamo Ω come l'***evento certo***

$\cdots$ $\emptyset$, l'insieme vuoto, non contenendo alcuno degli eventi elementari possibili, è un evento che non è mai verificato; dunque interpretiamo $\emptyset$ come l'***evento impossibile***

$\cdots\cdot\cdot$ $A \cup B = \Omega$ significa che l'evento costituito dal verificarsi di almeno uno dei due eventi A o B coincide con l'evento certo Ω; dunque interpretiamo tale condizione come ***A e B sono esaustivi*** *(è certo che se ne verifichi* **almeno** *uno dei due)*

$\cdots\cdot\cdot$ $A \cap B = \emptyset$ significa che l'evento costituito dal verificarsi di entrambi gli eventi A e B coincide con l'evento impossibile $\emptyset$; dunque interpretiamo la condizione $A \cap B = \emptyset$ come ***A e B sono incompatibili*** *(è certo che se ne verifichi* **al più** *uno dei due).*

Passiamo ora ad analizzare il significato "logico" della nozione di ***partizione dell'evento certo***. Consideriamo una collezione di sottoinsiemi $H_1, ..., H_m$ dello spazio Ω ($H_l \in \mathcal{P}(\Omega), l = 1, ..., m$). Tale collezione costituisce una ***partizione di*** $\boldsymbol{\Omega}$ se e solo se

$$\bigcup_{l=1}^{m} H_l = \Omega; \quad H_{l_1} \cap H_{l_2} = \emptyset, \text{ per } l_1 \neq l_2.$$

Interpretando $H_1, ..., H_m$ come eventi, abbiamo che essi sono a ***due a due incompatibili*** (cioè è *impossibile che se ne possono verificare due contemporaneamente*) e, d'altra parte, essi sono ***esaustivi*** (*è certo che se ne verifichi almeno uno*). In altre parole è certo che si verifichi uno ed uno soltanto degli eventi $H_1, ..., H_m$.

Ricordiamo infine le seguenti proprietà basilari delle operazioni booleane su insiemi:

Proprietà distributiva dell'unione rispetto all'intersezione:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Proprietà distributiva dell'intersezione rispetto all'unione:

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

Legge di De Morgan:

$$\complement(A \cap B) = \complement(A) \cup \complement(B) \quad \text{o equivalemente} \quad A \cap B = \complement\big(\complement(A) \cup \complement(B)\big).$$

Per brevità d'ora in poi scriveremo $\overline{A}$ invece di $\complement(A)$, per indicare il complementare (o la "negazione") di un insieme $A \in \mathcal{P}(\Omega)$ per cui la Legge di De Morgan si riscrive

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \quad \text{o equivalentemente} \quad A \cap B = \overline{(\overline{A} \cup \overline{B})}. \tag{1}$$

Anche tali proprietà ammettono una semplice interpretazione logica, ad esempio, per quanto riguarda la legge di De Morgan, l'interpretazione è la seguente:
negare il verificarsi sia di A che di B equivale a richiedere il verificarsi di almeno una tra la negazione di A e la negazione di B.
Le proprietà precedenti si estendono per induzione a famiglie finite di insiemi $\{A_i, i \in I\}$ (ma valgono anche per famiglie infinite):

$$\big(\bigcup_{i \in I} A_i\big) \cap C = \bigcup_{i \in I} \big(A_i \cap C\big),$$

$$\big(\bigcap_{i \in I} A_i\big) \cup C = \bigcap_{i \in I} \big(A_i \cup C\big),$$

$$\complement\big(\bigcap_{i \in I} A_i\big) = \bigcup_{i \in I} \complement(A_i) \quad \text{o con l'altra notazione,} \quad \overline{\big(\bigcap_{i \in I} A_i\big)} = \bigcup_{i \in I} \overline{A_i}.$$

1.2 Esercizi di verifica

Esercizio 1.1. Consideriamo l'esperimento consistente nel lancio di una pallina nel gioco della roulette. In tale esperimento è naturale porre

$$\Omega \equiv \{0, 1, ..., 36\}$$

e vedere i risultati *manque, passe, noir, rouge, pair, unpair,* come altrettanti eventi composti Supponiamo che nell'esperimento si verifichi l'evento elementare $\{16\}$. Quale degli eventi composti sopra elencati è verificato e quale no?

Come funziona la roulette

Si ricorda che la Roulette è una ruota con trentasette settori numerati da zero a trentasei. Una pallina viene fatta girare e alla fine si ferma su uno di questi numeri. Inoltre puntare su *manque* significa puntare su un numero tra 1 e 18, puntare su *passe* significa puntare su un numero tra 19 e 36, puntare su *noir* significa puntare su un numero nero, puntare su *rouge* significa puntare su un numero rosso, ed analogamente per *pair*, ovvero pari e *unpair*, ovvero dispari. Ai fini della soluzione dell'esercizio, è importante sapere che il 16 è rosso.

Esercizio 1.2. Dati due eventi A e B, scrivete, in termini di operazioni booleane, l'espressione dell'evento:

$$\{\textit{si verifica esattamente un solo evento fra } A \text{ e } B\}.$$

Esercizio 1.3. Siano A, B e C eventi. Scrivete le espressioni degli eventi:
a) Almeno due tra questi si verificano;
b) Esattamente due tra questi si verificano;
c) Al più due tra questi si verificano;
d) Esattamente uno tra questi si verifica.

Esercizio 1.4. Un'urna contiene oggetti di tipo A ed oggetti di tipo B; si eseguono due successive estrazioni dall'urna e si definiscono, per $i = 1, 2$, gli eventi:

$$E_i = \{\textit{oggetto di tipo A alla i-esima estrazione}\}.$$

In termini di operazioni booleane su E_1, E_2, scrivete l'espressione per l'evento

$$\{\textit{gli oggetti risultanti dalle due successive estrazioni sono dello stesso tipo}\}.$$

Esercizio 1.5. Un'urna contiene esattamente quattro elementi di tipo A e tre elementi di tipo B; da tale urna si effettuano tre successive estrazioni senza reinserimento, registrando il tipo dell'elemento via via estratto.
a) Elencate gli eventi elementari in questo esperimento e contate quanti sono.
b) Quanti sono, fra tali eventi elementari, quelli che realizzano l'evento

$$\{\textit{almeno due elementi di tipo } B \textit{ fra i tre elementi estratti}\}?$$

c) Quali e quanti sono, fra tali eventi elementari, quelli che realizzano l'evento

$$\{\textit{almeno due elementi di tipo } B\} \cup \{\textit{l'elemento estratto alla seconda estrazione è di tipo } B\}?$$

Esercizio 1.6. Consideriamo di nuovo l'urna di cui nell'esercizio precedente. Se ne effettuano sette successive estrazioni senza reinserimento (cioè l'urna viene progressivamente svuotata), registrando anche in questo caso soltanto il tipo dell'elemento via via estratto (tutti gli elementi di tipo A sono indistinguibili fra di loro, e tutti gli elementi di tipo B sono indistinguibili fra di loro).
Elencate gli eventi elementari in questo esperimento e contate quanti sono.

Esercizio 1.7. Considerate di nuovo l'urna come nel precedente Esercizio 1.5. Questa volta però le tre estrazioni sono effettuate con reinserimento.
a) Elencate anche in questo caso gli eventi elementari.
b) Dove risiede la differenza fra le due situazioni di estrazioni con e senza reinserimento?
(Per rispondere a tale domanda servono degli elementi non ancora studiati in questa prima lezione[2]).

Esercizio 1.8. Consideriamo ora il caso in cui vengono effettuate sette estrazioni con reinserimento dalla stessa urna di cui nei precedenti esercizi.
Quanti sono gli elementi elementari?

Esercizio 1.9. Una moneta viene lanciata due volte, registrando ogni volta se il risultato sia stato testa o croce.
a) Elencate gli eventi elementari possibili, in questo esperimento, e contate quanto vale $|\Omega|$, la cardinalità di Ω.
b) Qual è la cardinalità di $\mathcal{P}(\Omega)$, l'insieme delle parti di Ω? Cioè, quanti sono in tutto gli eventi, contando sia quelli semplici, quelli composti e quelli "banali" $\emptyset$ e Ω?

[2] Questo tipo di problemi sarà esaminato in generale nella Lezione 6

2 Spazi finiti di probabilità

2.1 Prime definizioni e proprietà

Introduciamo ora il concetto di **probabilità**. Come vedremo, tale concetto permette di formalizzare il problema di esprimere uno stato di incertezza circa il modo di risultare di un esperimento aleatorio

Sia dato uno spazio campione Ω e sia $\mathcal{P}(\Omega)$ la famiglia delle sue parti.

Definizione 2.1. (provvisoria) *Una* **misura di probabilità** *o, più semplicemente, una* **probabilità** *su* $(\Omega, \mathcal{P}(\Omega))$ *è una funzione che soddisfa i seguenti assiomi*[3]

i) $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$

ii) $\mathbb{P}(\Omega) = 1$ *(***condizione di normalizzazione***)*

iii) $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$*, per* $E_1 \cap E_2 = \emptyset$ *(***proprietà di additività***).*

Definizione 2.2. *Uno* **spazio finito di probabilità** *è una terna* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *dove* Ω *è uno insieme finito,* $\mathcal{P}(\Omega)$ *è la famiglia delle parti di* Ω *e* P *è una misura di probabilità su* $(\Omega, \mathcal{P}(\Omega))$ *.*

Osservazione 1. Prima di proseguire è opportuno citare il fatto che esistono diverse possibili *interpretazioni* del termine probabilità: ad esempio probabilità *classiche*, *frequentistiche*, *soggettivistiche, etc...*

Non rientra nei nostri scopi soffermarci sul significato e la portata di tali interpretazioni; per quanto ci riguarda ci basta accennare al fatto che, all'interno di ciascuna di dette interpretazioni, è giustificato imporre che la probabilità soddisfi le condizioni *i)*, *ii)*, *iii)* della Definizione 2.1.

Su tale base possiamo imporre tale condizioni come assiomi e procedere in modo appunto assiomatico; e di tali assiomi vedremo fra poco alcune conseguenze immediate.

Esercizio proposto 2.1. *Pensiamo all'esperimento del lancio di un dado a sei facce con*

$$\Omega = \{\omega_1, ..., \omega_6\}.$$

e poniamo

$$\mathbb{P}(E) = \frac{|E|}{6}.$$

Verificate che $\mathbb{P}(\cdot)$ *soddisfa gli assiomi i), ii), iii) e calcolare*

$$\mathbb{P}(\{X \leq 2\} \cup \{X \geq 5\}), \quad \mathbb{P}(\{X \geq 3\} \cap \{X \leq 4\}).$$

In quanto segue consideriamo ancora il caso di spazio campione finito:

$$\Omega = \{\omega_1, ..., \omega_N\}.$$

Elenchiamo ora alcune proprietà della probabilità, che risultano conseguenze immediate degli assiomi *i)*, *ii)*, *iii)* della Definizione 2.1.

[3]Su alcuni testi la proprietà i) è sostituita dalla **proprietà di non negatività**, ossia:
i') $\mathbb{P} : \mathcal{P}(\Omega) \to \mathbb{R}$, con la proprietà che $\mathbb{P}(E) \geq 0$ per ogni $E \in \mathcal{P}(\Omega)$.
È facile vedere che le proprietà *i')*, *ii)* e *iii)* implicano che $\mathbb{P}(E) \in [0, 1]$.

Prima di tutto notiamo che la proprietà *iii)* di additività si generalizza al caso di n eventi disgiunti a due a due

iii') Siano $E_1, ..., E_n \in \mathcal{P}(\Omega)$ *disgiunti (o incompatibili) a due a due, ovvero tali che*

$$E_i \cap E_j = \emptyset, \quad \textit{per } i, j \in \{1, 2, \cdots, n\}, \textit{ con } i \neq j; \tag{2}$$

allora si ha la condizione

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} E_i\Big) = \sum_{i=1}^{n} \mathbb{P}(E_i). \tag{3}$$

(la dimostrazione si ottiene facilmente per induzione su n*)*

Le relazioni elencate qui di seguito costituiscono delle immediate conseguenze degli assiomi della probabilità. Si invita il lettore a verificarle per esercizio.

(a) Per $E \in \mathcal{P}(\Omega)$, ponendo

$$p(\omega_i) = \mathbb{P}(\{\omega_i\}), \qquad i = 1, ..., N,$$

risulta

$$\mathbb{P}(E) = \sum_{\substack{i=1 \\ i:\,\omega_i \in E}}^{N} p(\omega_i) = \sum_{\omega \in E} p(\omega). \tag{4}$$

(b) Per ogni $E \in \mathcal{P}(\Omega)$ risulta

$$\mathbb{P}\left(\overline{E}\right) = 1 - \mathbb{P}(E), \tag{5}$$

dove $\overline{E}$ (invece di $\complement(E)$) indica il complementare (o la "negazione") di un evento $E \in \mathcal{P}(\Omega)$.

(c) L'evento impossibile ha probabilità nulla, ovvero

$$\mathbb{P}(\emptyset) = 0 \tag{6}$$

(d) **(proprietà di monotonia)** Siano $A, B \in \mathcal{P}(\Omega)$ tali che $A \subseteq B$; allora risulta

$$\mathbb{P}(A) \leq \mathbb{P}(B). \tag{7}$$

(e) Per arbitrari $A, B \in \mathcal{P}(\Omega)$ risulta

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{8}$$

(f) Siano $H_1, ..., H_n \in \mathcal{P}(\Omega)$ tali che

$$\bigcup_{i=1}^{n} H_i = \Omega, \quad H_i \cap H_j = \emptyset \quad \text{per } i \neq j, \tag{9}$$

ossia la collezione $H_1, ..., H_n$ è una partizione dell'evento certo; allora si ha la condizione

$$\sum_{i=1}^{n} \mathbb{P}(H_i) = 1. \tag{10}$$

In particolare, ricordando che $p(\omega_i) = \mathbb{P}(\{\omega_i\})$, $i = 1, ..., N$, e prendendo $H_i = \{\omega_i\}$, deve risultare

$$p(\omega_i) \geq 0 \qquad e \qquad \sum_{i=1}^{N} p(\omega_i) = 1. \tag{11}$$

Ulteriori conseguenze degli assiomi della probabilità verranno viste in seguito, dopo aver introdotto la probabilità condizionata (si veda la Lezione 4).

Le precedenti proprietà $(a) - -(f)$ si possono dimostrare a partire dalla definizione (4) di probabilità in uno spazio finito.

Alcune dimostrazioni con la definizione di probabilità in uno spazio finito

A titolo di esempio dimostriamo la proprietà (a):
Chiaramente $E = \cup_{\omega \in E}\{\omega\}$, ovvero, se indichiamo con $n = |E|$, e con $\{j_1, j_2, \ldots, j_n\}$ il sottoinsieme degli indici $\{1, 2, \ldots, N\}$, tali che

$$E = \{\omega_{j_1}, \omega_{j_2}, \ldots, \omega_{j_n}\},$$

si ha

$$E = \bigcup_{i=1}^{n} \{\omega_{j_i}\}.$$

Gli insiemi $\{\omega_{j_i}\}$ sono chiaramente incompatibili a due a due e quindi dalla proprietà di additività *iii')* si ottiene che

$$\mathbb{P}(E) = \sum_{i=1}^{n} P\big(\{\omega_{j_i}\}\big).$$

Tuttavia molte delle altre proprietà possono essere dimostrate sia a partire dalla (4) che dalla seguente relazione: per ogni A e B si ha

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B}). \tag{12}$$

A sua volta (12) deriva dall'osservazione che

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

che gli eventi $A \cap B$ e $A \cap \overline{B}$ sono incompatibili e infine dalla proprietà *iii)*.

Suggerimenti per le dimostrazioni dei punti (b) – (f)

Ad esempio la proprietà (b) si ottiene prendendo $A = \Omega$ e $B = E$, per cui $A \cap B = \Omega \cap E = E$ e $A \cap \overline{B} = \Omega \cap \overline{E} = \overline{E}$, da cui

$$\mathbb{P}(\Omega) = \mathbb{P}(E) + \mathbb{P}(\overline{E}) \overset{ii)}{\Longleftrightarrow} 1 = \mathbb{P}(E) + \mathbb{P}(\overline{E})$$

che è equivalente alla (5).
La proprietà (c) a sua volta si può derivare dalla (b) osservando che l'evento impossibile è il complementare dell'evento certo, ossia $\emptyset = \overline{\Omega}$, oppure osservando che $E = E \cup \emptyset$ e che, ovviamente, $E \cap \emptyset = \emptyset$, per cui dalla proprietà $ii)$,

$$\mathbb{P}(E) = \mathbb{P}(E \cup \emptyset) = \mathbb{P}(E) + \mathbb{P}(\emptyset).$$

La proprietà (d) (monotonia della probabilità) invece deriva dall'osservare che, se $A \subset B$, allora $A \cap B = A$, $A \cap \overline{B} = \emptyset$, la proprietà di base (12) diviene

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A \cap B),$$

e infine dall-osservare che, scambiando il ruolo di A e B nella (12), si ha

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap \overline{A}) \geq \mathbb{P}(A \cap B) = \mathbb{P}(A).$$

La proprietà (e) si può dimostrare a partire sempre dalla (12) (utilizzata anche scambiando il ruolo di A e B) e osservando che

$$A \cup B = (A \cap B) \cup (A \cap \overline{B}) \cup (B \cap \overline{A}).$$

Infine la proprietà (f) è immediata conseguenza degli assiomi, e precisamente delle proprietà $i)$ e $iii')$.

Osservazione 2. Nel caso in cui Ω è un insieme finito, possiamo guardare alla probabilità nei due modi, apparentemente diversi ma sostanzialmente equivalenti, che verranno illustrati qui di seguito (teniamo presente il fatto che ciascun punto di Ω può essere visto come un particolare sottoinsieme, cioè come un sottoinsieme composto da un solo elemento) :

1) Prima definiamo $\mathbb{P}$ come una funzione di insieme, cioè

$$\mathbb{P}: \ \mathcal{P}(\Omega) \to [0,1]; \quad E \mapsto \mathbb{P}(E)$$

che soddisfi gli assiomi *i)*, *ii)*, *iii)* della Definizione 2.1, e poi definiamo la funzione di punto

$$p: \ \Omega \to [0,1]; \quad \omega_i \mapsto p(\omega_i) := P(\{\omega_i\}).$$

Questa funzione dovrà soddisfare le condizioni

$$p(\omega_i) \geq 0, \qquad \sum_{\omega \in \Omega} p(\omega) = 1. \quad \text{ovvero} \quad \sum_{i=1}^{N} p(\omega_i) = 1. \tag{13}$$

2) Prima definiamo una funzione di punto

$$p: \ \Omega \to [0,1]; \ \omega_i \to p(\omega_i),$$

che soddisfi le condizioni (13) e poi definiamo una funzione di insieme

$$E \mapsto \mathbb{P}(E) := \sum_{\omega \in E} p(\omega),$$

ossia attraverso la precedente formula (4).

***Proposizione* 2.1.** *La funzione di insieme definita come sopra, ossia dalla formula (4), è una probabilità sull'insieme finito Ω, cioè soddisfa gli assiomi i), ii), iii) della Definizione 2.1.*

Esercizio proposto 2.2. *Dimostrare la precedente Proposizione 2.1.*

Esempio 2.1 (Un esempio concreto di spazio di probabilità finito)**.** *Sia* $\Omega = \{a, b, c, d\}$ *siano* $p(a) = 1/8$, $p(b) = 1/4$, $p(c) = 1/2$ $p(d) = 1/8$. *Chiaramente* $p(a)$, $p(b)$, $p(c)$, $p(d) \geq 0$ *e* $p(a) + p(b) + p(c) + p(d) = 1$. *Allora la probabilità definita dalla precedente formula sull'insieme delle parti di* $\Omega = \{a, b, c, d\}$ *è data dalla funzione che è specificata nella seguente tabella.*

Tabella della Probabilità definita nell'Esempio 2.1
(un esempio concreto di spazio di probabilità finito)

$$
\begin{aligned}
\emptyset &\mapsto \mathbb{P}(\emptyset) = 0\\
\{a\} &\mapsto \mathbb{P}(\{a\}) = p(a) = 1/8\\
\{b\} &\mapsto \mathbb{P}(\{b\}) = p(b) = 1/4\\
\{c\} &\mapsto \mathbb{P}(\{c\}) = p(c) = 1/2\\
\{d\} &\mapsto \mathbb{P}(\{d\}) = p(d) = 1/8\\
\{a,b\} &\mapsto \mathbb{P}(\{a,b\}) = p(a) + p(b) = 1/8 + 1/4 = 3/8\\
\{a,c\} &\mapsto \mathbb{P}(\{a,c\}) = p(a) + p(c) = 1/8 + 1/2 = 5/8\\
\{a,d\} &\mapsto \mathbb{P}(\{a,d\}) = p(a) + p(d) = 1/8 + 1/8 = 1/4\\
\{a,b\} &\mapsto \mathbb{P}(\{b,c\}) = p(b) + p(c) = 1/4 + 1/2 = 3/4\\
\{b,d\} &\mapsto \mathbb{P}(\{b,d\}) = p(b) + p(d) = 1/4 + 1/8 = 3/8\\
\{c,d\} &\mapsto \mathbb{P}(\{c,d\}) = p(c) + p(d) = 1/2 + 1/8 = 5/8\\
\{a,b,c\} &\mapsto \mathbb{P}(\{a,b,c\}) = p(a) + p(b) + p(c) = 1/8 + 1/4 + 1/2 = 7/8\\
\{a,b,d\} &\mapsto \mathbb{P}(\{a,b,d\}) = p(a) + p(b) + p(d) = 1/8 + 1/4 + 1/8 = 1/2\\
\{a,c,d\} &\mapsto \mathbb{P}(\{a,c,d\}) = p(a) + p(c) + p(d) = 1/8 + 1/2 + 1/8 = 3/4\\
\{b,c,d\} &\mapsto \mathbb{P}(\{b,c,d\}) = p(b) + p(c) + p(d) = 1/4 + 1/2 + 1/8 = 7/8\\
\{a,b,c,d\} &\mapsto \mathbb{P}(\{a,b,c,d\}) = p(a) + p(b) + p(c) + p(d) = 1/8 + 1/4 + 1/2 + 1/8 = 1
\end{aligned}
$$

***Osservazione* 3** (**Probabilità definite a meno di un fattore di proporzionalità; normalizzazione**). Una misura di probabilità sullo spazio $\Omega = \{\omega_1, ..., \omega_N\}$ è individuata quando vengano assegnati i numeri $p(\omega_i) = \mathbb{P}(\{\omega_i\})$, $(i = 1, 2, ..., N)$ soddisfacenti le condizioni (13). Supponiamo ora che $p(\omega_i)$, $(i = 1, 2, ..., N)$ siano assegnati a meno di una costante di proporzionalità; supponiamo cioè che siano assegnati dei numeri g_i $(i = 1, 2, ..., N)$, tali che

$$p(\omega_i) = K \cdot g_i \tag{14}$$

essendo K un'opportuna costante positiva. Dalla condizione di normalizzazione (11), si ricava

$$K = \frac{1}{\sum_{j=1}^{N} g_j}; \qquad p(\omega_i) = \frac{g_i}{\sum_{j=1}^{N} g_j}.$$

Notiamo che si usa esprimere brevemente la condizione (14) usando il seguente simbolismo:

$$p(\omega_i) \propto g_i.$$

Esempio 2.2 (dado non equilibrato)**.** *Un dado ha sei facce numerate da 1 a 6; esso è pesato in modo tale che ciascuna faccia abbia una probabilità di presentarsi (in un singolo lancio) proporzionale al suo valore. Sia*

$$A \equiv \{\textit{si presenta un numero pari}\}.$$

Trovare $\mathbb{P}(A)$.

Soluzione. Si ha $\Omega = \{\omega_1, ..., \omega_6\}$ e vogliamo imporre

$$p(\omega_i) = K \cdot i, \qquad i = 1, ..., 6,$$

essendo K una costante positiva da determinare, imponendo la condizione di normalizzazione (11); si ottiene dunque

$$p(\omega_i) = \frac{i}{21}, \quad i = 1, ..., 6$$

e

$$\mathbb{P}(A) = p(\omega_2) + p(\omega_4) + p(\omega_6) = \frac{12}{21}.$$

2.2 Esercizi di verifica

Esercizio 2.1. Un dado è pesato in modo tale che la probabilità di avere un punto pari è il doppio della probabilità di avere un punto dispari. Qual è la probabilità di avere punto pari?

Esercizio 2.2. Siano A e B due eventi tali che

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap \overline{B}) = \mathbb{P}(\overline{A} \cap B) = \mathbb{P}(\overline{A} \cap \overline{B}).$$

a) Quanto vale $\mathbb{P}(A \cap B)$?
b) Qual è la probabilità che, fra A e B, se ne verifichi almeno uno?
c) Qual è la probabilità che se ne verifichi esattamente uno?

Esercizio 2.3. Una moneta viene lanciata due volte e poniamo

$$E_i \equiv \{\textit{testa all'i-esimo lancio}\}, \quad i = 1, 2.$$

Mostrare che la condizione $\mathbb{P}(E_1 \cap \overline{E}_2) = \mathbb{P}(\overline{E}_1 \cap E_2)$ implica $\mathbb{P}(E_1) = \mathbb{P}(E_2)$.

Esercizio 2.4. Siano A, B, e C tre eventi. Dimostrate che vale la formula

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

$$= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

(questa formula costituisce un caso particolare della formula di inclusione - esclusione: si veda la nota nella pagina seguente)

Esercizio 2.5. Siano A, B, e C tre eventi tali che

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}\left(A \cap \overline{B} \cap C\right) = 0.1,$$

$$\mathbb{P}\left(\overline{A} \cap B \cap C\right) = \mathbb{P}\left(A \cap B \cap \overline{C}\right) = 0.15,$$

$$\mathbb{P}\left(A \cap \overline{B} \cap \overline{C}\right) = \mathbb{P}\left(\overline{A} \cap B \cap \overline{C}\right) = \mathbb{P}\left(\overline{A} \cap \overline{B} \cap C\right) = 0.05.$$

Calcolare
a) $\mathbb{P}(A), \mathbb{P}(B), \mathbb{P}(C)$
b) $\mathbb{P}(A \cup B), \mathbb{P}(A \cup C)$
c) $\mathbb{P}(A \cup B \cup C)$.

FORMULA DI INCLUSIONE ED ESCLUSIONE

Dati n eventi A_1, A_2, A_n, la probabilità che se ne verifichi almeno uno tra tali eventi è data dalla seguente formula:

$$\begin{aligned}
&\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n)\\
=&\mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n)\\
-&[\mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \cdots + \mathbb{P}(A_i \cap A_j) + \cdots + \mathbb{P}(A_{n-1} \cap A_n)]\\
+&\cdots\\
+&(-1)^{k-1} \sum_{\{i_1, i_2, \cdots, i_k\}} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})\\
+&\cdots\\
+&(-1)^{n-1}\,\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)\\
=&\sum_{k=1}^{n} (-1)^{k-1} \sum_{\{i_1, i_2, \cdots, i_k\}} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}),
\end{aligned}$$

dove, per k generico, le somme sono estese a tutti i sottoinsiemi $\{i_1, i_2, \cdots, i_k\}$ di $\{1, 2, \cdots, n\}$, composti da k elementi (nella prossima Lezione 3, calcoleremo il numero complessivo di tali sottoinsiemi, al variare di n e di k).

La dimostrazione di tale formula si ottiene per induzione osservando che

$$A_1 \cup A_2 \cup \cdots \cup A_{n-1} \cup A_n = (A_1 \cup A_2 \cup \cdots \cup A_{n-1}) \cup A_n$$

e che quindi

$$\begin{aligned}
&\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n)\\
=&\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_{n-1}) + \mathbb{P}(A_n) - \mathbb{P}\big((A_1 \cup A_2 \cup \cdots \cup A_{n-1}) \cap A_n\big)\\
=&\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_{n-1}) + \mathbb{P}(A_n)\\
-&\,\mathbb{P}\big((A_1 \cap A_n) \cup (A_2 \cap A_n) \cup \cdots \cup (A_{n-1} \cap A_n)\big)
\end{aligned}$$

ed infine utilizzando il passo induttivo per $n-1$, anche per calcolare la probabilità di $(A_1 \cap A_n) \cup (A_2 \cap A_n) \cup \cdots \cup (A_{n-1} \cap A_n)$.

3 Probabilità “classiche” e calcolo combinatorio

3.1 Probabilità “classiche”

Qui ci soffermiamo a trattare alcuni casi particolari, ma molto rilevanti, di spazi di probabilità finiti. Sia dunque

$$\Omega = \{\omega_1, ..., \omega_N\},$$

e supponiamo che si voglia porre

$$p(\omega_i) = K, \qquad \forall \omega_i \in \Omega \tag{15}$$

per un'opportuna costante positiva K. Si vuole cioè imporre che tutti i risultati elementari siano, fra di loro, equiprobabili.

Ci riferiremo a tale caso dicendo che si ha una ***distribuzione di probabilità uniforme sugli eventi elementari***.

Confrontando la posizione (15) con la condizione di normalizzazione (11), otteniamo immediatamente

$$p(\omega_i) = \frac{1}{N}, \qquad i = 1, ..., N$$

e da ciò segue, ricordando la formula (4) del precedente paragrafo,

$$\mathbb{P}(E) = \frac{|E|}{N}, \qquad \forall E \in \mathcal{P}(\Omega). \tag{16}$$

Esempio 3.1. *L'addetto ad un guardaroba restituisce a caso n ombrelli che gli sono stati consegnati; qual è la probabilità che il secondo cliente abbia indietro il suo proprio ombrello?*

Soluzione. Si ha che Ω è costituito dalle permutazioni[4] di n elementi; dunque $|\Omega| = n!$ L'evento

$$E \equiv \{\textit{Il secondo cliente riceve indietro il suo ombrello}\}$$

è un evento composto, costituito da tutte le permutazioni che tengono fisso il secondo elemento; tali permutazioni sono in numero di $(n-1)!$, corrispondente al numero delle possibili permutazioni dei restanti $(n-1)$ elementi. L'espressione “a caso” vuole significare che tutte le permutazioni sono da considerare equiprobabili fra di loro. Dunque

$$\mathbb{P}(E) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Osservazione 1. La formula (16) esprime il fatto che, nel caso in cui tutti gli eventi elementari di uno spazio finito sono equiprobabili, la probabilità di un generico evento composto si calcola quale ***rapporto fra casi favorevoli e casi possibili***.

Si faccia attenzione al fatto che la (16) *non costituisce una definizione del concetto di probabilità, ma soltanto una formula per il suo calcolo in un caso particolare:*
nel precedente paragrafo abbiamo già introdotto tale concetto in modo assiomatico;
la suddetta formula è stata ottenuta come immediata conseguenza degli assiomi stessi, nel caso particolare di eventi elementari equiprobabili.

Osservazione 2. Nel caso in cui si imponga la condizione (15), il calcolo della probabilità di un evento composto E si riduce al problema, ***combinatorio***, di individuare $N = |\Omega|$ e $|E|$.

[4] Per la definizione di permutazione vedere più avanti la Lezione 3.2

3.2 Calcolo combinatorio: primi elementi

Facendo seguito alle precedenti ***Osservazione 1*** e ***Osservazione 2***, ci rivolgiamo ora a richiamare succintamente alcune nozioni basilari di calcolo combinatorio, che risultano indispensabili per affrontare i primi problemi di calcolo delle probabilità.

Le formule che verranno presentate si ricavano facilmente tramite applicazione del ***principio di induzione finita***.

Iniziamo innanzitutto ricordando due fatti fondamentali:

a) Due insiemi finiti hanno la stessa cardinalità se e solo se fra essi è possibile stabilire una ***corrispondenza biunivoca***.

b) Dati due arbitrari insiemi A e B, si definisce ***prodotto cartesiano*** di A per B l'insieme costituito dalle coppie ordinate (a, b) dove $a \in A$ e $b \in B$; indichiamo tale insieme con il simbolo $A \times B$. Nel caso in cui A e B sono insiemi finiti, risulta

$$|A \times B| = |A| \cdot |B|.$$

In quanto immediatamente segue supponiamo di aver fissato un arbitrario insieme A costituito da n elementi:

$$A \equiv \{a_1, ..., a_n\}.$$

Disposizioni con ripetizione di classe k di n elementi.

Una disposizione con ripetizione di classe k degli n elementi di A non è altro che una k-upla ordinata degli elementi stessi.

Tali disposizioni costituiscono dunque l'insieme $A^k = \overbrace{A \times A \times \cdots \times A}^{k\,volte}$, e si ha $|A^k| = n^k$.

Disposizioni senza ripetizione di classe k di n elementi e permutazioni di n elementi

Le disposizioni senza ripetizione di classe k degli n elementi sono le k-uple costituite da elementi di A, ***tutti diversi fra loro***.

Tali disposizioni costituiscono un sottoinsieme, dell'insieme A^k, di cardinalità[5]

$$\overbrace{n(n-1)...\big(n-(k-1)\big)}^{k\,fattori} = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!},$$

dove si è usata la notazione *n fattoriale*, ovvero $n! = n(n-1)\cdots 3 \cdot 2 \cdot 1$.

Nel caso in cui si ponga $k = n$, si ottengono le ***permutazioni*** degli elementi di A. Di conseguenza il numero delle permutazioni di n elementi è $n!$.

Combinazioni di classe k di n elementi

Si tratta di classi di equivalenza di disposizioni senza ripetizione di classe k di n elementi, modulo la relazione di equivalenza costituita dal considerare equivalenti due disposizioni che contengono gli stessi elementi, eventualmente in ordine diverso.

Alternativamente una combinazione di classe k di n elementi si può definire come ***un sottoinsieme di cardinalità k di un insieme di cardinalità n***.

Se C_k^n indica il numero delle combinazioni di classe k di n elementi, D_k^n indica il numero delle disposizioni senza ripetizione di classe k di n elementi, e P_k indica il numero delle permutazioni di k elementi, è facile convincersi[6] che

$$D_k^n = C_k^n \cdot P_k,$$

[5]Si veda la nota di approfondimento a pagina 31

[6]Si veda la nota di approfondimento pagina 32

da cui, tenendo conto che $P_k = k!$ e $D_k^n = \frac{n!}{(n-k)!}$, si ricava immediatamente $\frac{n!}{(n-k)!} = C_k^n \cdot k!$.
Il numero complessivo di tali combinazioni è dunque dato da

$$C_k^n = \frac{1}{k!} \cdot \frac{n!}{(n-k)!}.$$

Si pone

$$\binom{n}{k} \equiv \frac{n!}{k!(n-k)!}.$$

Il numero $\binom{n}{k}$ prende il nome di ***coefficiente binomiale n sopra k*** (o anche n su k).

Esempio 3.2. *Consideriamo un circolo costituito da n persone e supponiamo di dover eleggere un* presidente*, un* segretario *e un* tesoriere*.*

Se pensiamo di scegliere tre persone diverse, ognuna con la sua specifica carica, ciascuna scelta coincide con una disposizione senza ripetizione di classe 3 degli n elementi; abbiamo $n(n-1)(n-2)$ possibili scelte.

Se pensiamo che ogni carica è assegnata con una votazione indipendente dalle altre, si possono avere anche delle ripetizioni (cioè è ammesso un cumulo delle cariche); in tal caso ciascuna possibile scelta coincide con una disposizione di classe 3, con ripetizione, degli n elementi; abbiamo n^3 possibili scelte.

Se pensiamo di eleggere complessivamente una terna di persone diverse, senza attribuire una specifica carica a ciascuna di loro, ma incaricandoli complessivamente dei compiti di presidente, di segretario e di tesoriere, ciascuna possibile scelta coincide con una combinazione di classe 3 degli n elementi; abbiamo, in tal caso $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$ possibili scelte.

A proposito di coefficienti binomiali è utile introdurre la seguente convenzione: per ogni numero naturale n, si pone

$$\binom{n}{0} = 1,$$

come è ovvio, sia tenendo presente la convenzione $0! = 1$, sia tenendo presente il fatto che l'unico sottoinsieme di cardinalità zero è l'insieme vuoto.

Esercizio proposto 3.1. *Dimostrare che $\sum_{k=0}^{n} C_n^k = 2^n$, senza usare la formula della potenza del binomio di Newton*[7]*.*

3.3 Alcuni classici esempi

Consideriamo ora qualche semplice e classico esempio di probabilità combinatorie.

Esempio 3.3 (Problema del compleanno)**.** *Qual è la probabilità che, fra M persone scelte a caso, ve ne siano almeno due che festeggiano il compleanno nello stesso giorno? (Si supponga l'anno costituito da 365 giorni e che vi sia una situazione di simmetria rispetto alle nascite).*

Soluzione. Calcoliamo la probabilità dell'evento complementare

$$\overline{E} = \{\textit{Le M persone festeggiano il compleanno in tutti giorni diversi}\}$$

[7]Si veda anche la nota di approfondimento a pagina 28

Lo spazio Ω è costituito dalle disposizioni con ripetizione di classe M di 365 elementi (i giorni dell'anno solare). $\overline{E}$ è un evento composto costituito da tutte le disposizioni senza ripetizione di classe M di 365 elementi. Quindi

$$\mathbb{P}(\overline{E}) = \frac{365 \cdot 364 \cdot \ldots \cdot (365 - M + 1)}{(365)^M}$$

e la probabilità cercata è fornita da $\mathbb{P}(E) = 1 - \mathbb{P}(\overline{E})$. Indichiamo ora tale probabilità con $\mathbb{P}_M(E)$ per mettere in evidenza la sua dipendenza dal valore di M. Ovviamente $\mathbb{P}_M(E)$ è una funzione crescente di M ed è interessante notare che si ha $\mathbb{P}_M(E) > \frac{1}{2}$ per $M > 22$, in particolare si ha $\mathbb{P}_{22}(E) \simeq 0.4756$ mentre $\mathbb{P}_{23}(E) \simeq 0.5072$.

Esempio 3.4 ("Paradosso del Cavalier De Méré")**.** *È più probabile ottenere almeno un asso in 4 lanci consecutivi di un dado o un doppio asso in 24 lanci consecutivi di una coppia di dadi?*

Soluzione. Anche qui conviene calcolare le probabilità dei due eventi complementari:

$$\mathbb{P}(\{\textit{almeno un asso in 4 lanci}\}) = 1 - \mathbb{P}(\{\textit{nessun asso in 4 lanci}\})$$

$$\begin{aligned}&\mathbb{P}(\{\textit{almeno un doppio asso in 24 lanci}\}) = \\ &= 1 - \mathbb{P}(\{\textit{nessun doppio asso in 24 lanci}\}).\end{aligned}$$

I risultati possibili nei 4 lanci del dado sono rappresentati dalle disposizioni con ripetizione di classe 4 di 6 elementi; in altre parole, possiamo rappresentare Ω come lo spazio delle quaterne ordinate (x_1, x_2, x_3, x_4) con $x_i \in \{1, ..., 6\}$. Dunque $|\Omega| = 6^4$. Gli eventi elementari che costituiscono l'evento composto $\{$*nessun asso in 4 lanci*$\}$ corrispondono, invece, alle disposizioni con ripetizione di classe 4 dei 5 elementi $\{2, ..., 6\}$. Si ha quindi

$$\mathbb{P}(\{\textit{almeno un asso in 4 lanci}\}) = 1 - \frac{5^4}{6^4} \simeq 0.52.$$

Analogamente si ottiene

$$\mathbb{P}(\{\textit{almeno un doppio asso in 24 lanci}\}) = 1 - \frac{35^{24}}{36^{24}} \simeq 0.49.$$

Esempio 3.5. *Un gruppo di $4n$ persone comprende $2n$ ragazzi e $2n$ ragazze. Vengono formate a caso due squadre di $2n$ persone ciascuna.*

a) Qual è la probabilità che tutte le ragazze si trovino nella stessa squadra e tutti i ragazzi nella squadra avversaria?

b) Qual è la probabilità che ciascuna squadra sia, all'opposto, composta esattamente da n ragazzi ed n ragazze?

Soluzione. Qui il generico evento elementare è specificato da un modo di scegliere $2n$ oggetti (i componenti della prima squadra) da un insieme di $4n$ oggetti; dunque la cardinalità dello spazio degli eventi elementari Ω è data da $\binom{4n}{2n}$.

> **Suggerimento:** in effetti si può pensare che le ragazze siano numerate come $f_1,\ f_2,\ \ldots,\ f_{2n}$ ed i ragazzi come $m_1,\ m_2,\ \ldots,\ m_{2n}$. Per specificare la prima squadra basta prendere un sottoinsieme di
>
> $$R = \{f_1,\ f_2,\ \ldots,\ f_{2n},\ m_1,\ m_2,\ \ldots,\ m_{2n}\}$$
>
> di cardinalità $2n$, ovvero
>
> $$\Omega = \{\text{combinazioni dei } 4n \text{ elementi di } R \text{ di classe } 2n\}, \qquad \text{con } |\Omega| = \binom{4n}{2n}$$

Nel caso a) due soli eventi elementari sono favorevoli e dunque la probabilità cercata è $\frac{2}{\binom{4n}{2n}}$.

> **Suggerimento:** i due casi favorevoli sono le due combinazioni $\{f_1,\ f_2,\ \dots,\ f_{2n}\}$ e $\{m_1,\ m_2,\ \dots,\ m_{2n}\}$.

Nel caso b) gli eventi elementari favorevoli[8] sono in numero di $\binom{2n}{n}\binom{2n}{n}$, corrispondente al numero dei modi in cui si possono scegliere n ragazze dal gruppo di tutte le $2n$ e n ragazzi dal gruppo di tutti i $2n$. La probabilità cercata è dunque data da

$$\frac{\binom{2n}{n}\binom{2n}{n}}{\binom{4n}{2n}}.$$

Esempio 3.6. *Supponiamo che una moneta perfetta venga lanciata n volte. Per $h \leq n$, qual è la probabilità di nessuna testa sui primi h lanci?*

Soluzione. Possiamo schematizzare gli eventi elementari in questo esperimento come gli elementi dell'insieme $\{0,1\}^n$ cioè come n-uple con elementi uguali a 0 (croce) o uguali a 1 (testa), (ad esempio l'evento elementare $\omega \equiv (0,0,1,1,0,...,0)$ coincide con il fatto che i primi due lanci danno croce, poi si hanno consecutivamente due risultati testa, e poi in tutti i successivi lanci si ottiene ancora croce); dunque si ha $|\Omega| = 2^n$.

L'evento $\{$*nessuna testa sui primi h lanci*$\}$ è allora l'evento composto

$$E \equiv \{\omega \in \Omega :\ \omega = (0,0,...,0,\omega_{h+1},...,\omega_n)\,,\ \text{con}\ (\omega_{h+1},...,\omega_n) \in \{0,1\}^{n-h}\}.$$

Traduciamo la condizione che la moneta sia perfetta con la posizione

$$p(\omega) = \frac{1}{2^n}, \qquad \forall \omega \in \Omega.$$

Si ha $|E| = 2^{n-h}$ e dunque $\mathbb{P}(E) = \frac{2^{n-h}}{2^n} = \left(\frac{1}{2}\right)^h$.

Esempio 3.7. *Qual è la probabilità di k risultati testa negli n lanci di una moneta?*

Soluzione. Si ha lo stesso spazio di probabilità dell'esercizio precedente; questa volta $|E| = \binom{n}{k}$ e dunque $\mathbb{P}(E) = \frac{\binom{n}{k}}{2^n}$; come vedremo in seguito si tratta di un caso particolare di probabilità ***binomiali***.

Esempio 3.8. *Trovare la probabilità di k voti per lo schieramento A in un sondaggio elettorale di ampiezza n in un gruppo di M elettori di cui è noto che m_1 votano per A e $m_2 = M - m_1$ votano per B.*

Soluzione. Si è sottinteso che gli n elettori siano stati selezionati senza reinserimento. L'esperimento consiste dunque nel selezionare un sottoinsieme di cardinalità n (*il campione*) dall'insieme degli M elettori (*la popolazione*) e quindi $|\Omega| = \binom{M}{n}$. Si sottointende che il sondaggio sia condotto in modo casuale, cioè che ogni "campione" abbia uguale probabilità $\frac{1}{\binom{M}{n}}$ di essere estratto. Fra tali "campioni", ve ne sono $\binom{m_1}{k} \cdot \binom{m_2}{n-k}$ che contengono k elettori per A e $(n-k)$ per B. Infatti, ci sono $\binom{m_1}{k}$ modi di selezionare k elettori fra i votanti per A, ci sono $\binom{m_2}{n-k}$ modi di selezionare $(n-k)$ elettori fra i votanti per B e, inoltre, una qualunque scelta di k elettori fra i votanti per A e di $(n-k)$ elettori fra i votanti per B dà luogo ad una n-upla di elettori (un campione) che contiene k elettori per A.

[8]Per convincersene si consiglia il lettore di considerare il caso $n = 1$ ed $n = 2$, elencando esplicitamente sia tutti i casi possibili che tutti i casi favorevoli.

Dunque la probabilità cercata è data da

$$\frac{\binom{m_1}{k}\cdot\binom{m_2}{n-k}}{\binom{M}{n}}=\frac{\binom{m_1}{k}\cdot\binom{m_2}{n-k}}{\binom{m_1+m_2}{n}}.$$

Come vedremo in seguito, si tratta di un caso particolare di probabilità ***ipergeometriche***. Osserviamo che i valori possibili per k devono rispettare la condizione

$$0\leq k\leq m_1,\quad 0\leq n-k\leq m_2,\quad \textit{con}\quad n\leq M=m_1+m_2,$$

che, dopo semplici passaggi, diviene

$$0\vee(n-m_2)=\max(0,n-m_2)\leq k\leq\min(n,m_1)=n\wedge m_1.$$

Esercizio proposto 3.2. *Siano M, m_1, n e k numeri assegnati e tali che*

$$m_1<M,\quad n<M,\quad \max(0,m_1+n-M)\leq k\leq\min(n,m_1).$$

Verificate l'identità

$$\frac{\binom{m_1}{k}\binom{M-m_1}{n-k}}{\binom{M}{n}}=\frac{\binom{n}{k}\binom{M-n}{m_1-k}}{\binom{M}{m_1}}.$$

Riuscite a darne un'interpretazione probabilistica?

Interpretazione dell'identità

$$\frac{\binom{m_1}{k}\binom{M-m_1}{n-k}}{\binom{M}{n}} = \frac{\binom{n}{k}\binom{M-n}{m_1-k}}{\binom{M}{m_1}},$$

del precedente esercizio-proposto 3.2

Pensate ad esempio di avere un'urna e di estrarre tutte le palline in ordine: si suppone come al solito che l'urna contenga m_1 palline di tipo A ed $m_2 = M - m_1$ di tipo B. Siamo interessati alla probabilità che tra le prime n (con $1 \leq n \leq M$) estratte ce ne siano (esattamente) k di tipo A (e quindi $n-k$ di tipo B). Un esperimento si può descrivere annotando solo l'ordine con cui vengono estratti i tipi. Ad esempio se $m_1 = 3$ ed $m_2 = 2$, i risultati possibili sono

(A, A, A, B, B)	(A, A, B, A, B)	(A, A, B, B, A)	(A, B, A, A, B)	(A, B, A, B, A)
(A, B, B, A, A)	(B, A, A, A, B)	(B, A, A, B, A)	(B, A, B, A, A)	(B, B, A, A, A)

Non è difficile convincersi che tutti i casi possibili sono equiprobabili e che sono in tutto quante le combinazioni di 5 elementi di classe 3, infatti basta specificare le posizioni in cui sono uscite le palline di tipo A. Con questa corrispondenza i casi precedenti sono in corrispondenza con

{1,2,3}	{1,2,4}	{1,2,5}	{1,3,4}	{1,3,5}
{1,4,5}	{2,3,4}	{2,3,5}	{2,4,5}	{3,4,5}

Più in generale si ottiene che i casi possibili sono appunto

$$C_{m_1}^{M} = \binom{M}{m_1},$$

in quanto per elencare tutti i casi possibili basta specificare solo le m_1 posizioni occupate dalle A su tutte le M posizioni.
Per quanto riguarda i casi favorevoli, iniziamo di nuovo con il caso particolare dell'urna precedente e con $n = 3$ e $k = 2$. I casi favorevoli sono

(A, A, B, A, B)	(A, A, B, B, A)
(A, B, A, A, B)	(A, B, A, B, A)
(B, A, A, A, B)	(B, A, A, B, A)

e si ottengono nel seguente modo: basta specificare quali sono le due A tra le prime 3 posizioni, e dove si trova la terza A tra le ultime $2 = 5 - 3$ posizioni. Il numero totale risulta quindi $\binom{3}{2} \cdot \binom{2}{1} = 3 \cdot 2 = 6$.

Più in generale per elencare tutti casi favorevoli all'evento che tra le prime n estratte ce ne siano k di tipo A, basta specificare le k posizioni tra le prime n in cui si trovano le A (e questo si può fare in C_k^n=$\binom{n}{k}$ modi, e inoltre specificare le $m_1 - k$ posizioni occupate dalle A, nelle ultime $M - n$ posizioni (e questo si può fare in $C_{m_1-k}^{M-n}$=$\binom{M-n}{m_1-k}$ modi). In totale quindi si hanno

$$\binom{n}{k} \cdot \binom{M-n}{m_1-k}$$

casi favorevoli.

3.4 Alcune proprietà dei coefficienti binomiali.

Nello studio del Calcolo delle Probabilità è opportuno tenere presente alcune identità fondamentali riguardanti i coefficienti binomiali. Ne presentiamo intanto alcune qui di seguito.

Una prima semplice identità è la seguente (nota come formula di Stiefel)

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \tag{17}$$

purché $k-1 \geq 0$ e $k \leq n-1$, ovvero per $1 \leq k \leq n-1$.

Si noti inoltre che per $k=0$ e per $k=n$ ovviamente si ha $\binom{n}{0} = \binom{n}{n} = 1$. La dimostrazione di tale formula è immediata; basterà infatti sviluppare i coefficienti binomiali (provare come esercizio).

Qui vogliamo comunque anche darne una semplice dimostrazione probabilistica[9], ricordando quanto visto nel precedente Esempio 3.7, relativo ad n lanci di una moneta.

Poniamo $E \equiv \{$*si ottengono k risultati testa in n lanci di una moneta perfetta*$\}$. Sappiamo che la probabilità di ottenere tale risultato è uguale a $\frac{\binom{n}{k}}{2^n}$. D'altra parte, ponendo

$$E_1 \equiv \{\textit{(k-1) teste sui primi (n-1) lanci}\} \cap \{\textit{testa all'n-esimo lancio}\}\,,$$

$$E_2 \equiv \{\textit{k teste sui primi (n-1) lanci}\} \cap \{\textit{croce all'n-esimo lancio}\}\,,$$

possiamo anche scrivere

$$E \equiv E_1 \cup E_2$$

e, essendo chiaramente $E_1 \cap E_2 = \emptyset$,

$$\mathbb{P}(E) = \mathbb{P}(E_1) + \mathbb{P}(E_2). \tag{18}$$

Ora possiamo notare che gli eventi composti E_1 ed E_2 hanno rispettivamente cardinalità uguale a $\binom{n-1}{k-1}$ e $\binom{n-1}{k}$; e dunque la (18) diventa

$$\frac{\binom{n}{k}}{2^n} = \frac{\binom{n-1}{k-1}}{2^n} + \frac{\binom{n-1}{k}}{2^n}.$$

L'identità (17) è in particolare alla base della costruzione del ben noto ***Triangolo di Tartaglia***.

$$\begin{array}{ccccccccccccccccc}
&&&&&&&&1&&&&&&&&\\
&&&&&&&1&&1&&&&&&&\\
&&&&&&1&&2&&1&&&&&&\\
&&&&&1&&3&&3&&1&&&&&\\
&&&&1&&4&&6&&4&&1&&&&\\
&&&1&&5&&10&&10&&5&&1&&&\\
&&1&&6&&15&&20&&15&&6&&1&&\\
&1&&7&&21&&35&&35&&21&&7&&1&\\
1&&8&&28&&56&&70&&56&&28&&8&&1
\end{array}$$

Triangolo di Tartaglia

[9] Si veda anche la nota di approfondimento a pagina 29

Utilizzando (17) è anche facile, per $0 \le k \le n$, ottenere la seguente uguaglianza

$$\sum_{r=k}^{n} \binom{r}{k} = \binom{n+1}{k+1}$$

È ben noto (e comunque si verifica immediatamente per induzione, utilizzando la (17)) che i coefficienti binomiali intervengono come segue nello sviluppo della potenza di un binomio: siano a, b due arbitrari numeri reali non nulli e sia n un numero naturale; allora risulta

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \cdot b^{n-k}. \tag{19}$$

Ponendo nella (19) $a = x, b = 1$, otteniamo l'identità

$$(x+1)^n = \sum_{k=0}^{n} \binom{n}{k} x^k. \tag{20}$$

In particolare ponendo $a = b = 1$, otteniamo

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Tenendo presente che $\binom{n}{k}$ coincide con il numero di sottoinsiemi di cardinalità k contenuti in un insieme composto da n elementi, otteniamo che 2^n è uguale alla cardinalità della famiglia delle parti di un insieme di n elementi. Dunque se in un esperimento vi sono n eventi elementari, vi sono allora in tutto 2^n eventi fra elementari, composti e contando anche l'evento certo e quello impossibile (si veda anche la nota di approfondimento a pagina 28).

Un'altra utile identità si può ottenere facilmente per una qualunque terna di numeri naturali r, s, n con $n \le r + s$:

$$\sum_{k=0 \bigvee (n-s)}^{n \bigwedge r} \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}, \tag{21}$$

dove la somma è estesa a tutti gli indici k per i quali $0 \le k \le r$ e $0 \le n-k \le s$.

Quest'ultima uguaglianza può essere verificata ad esempio osservando intanto quanto segue: per ogni x reale

$$(1+x)^{r+s} = (1+x)^r \cdot (1+x)^s.$$

Da una parte, utilizzando la (20) con $n = r$, e $n = s$, e la proprietà che il prodotto di sommatorie è la sommatoria del prodotto (si veda il richiamo nel riquadro a pagina 25)

$$\begin{aligned}
(1+x)^r \cdot (1+x)^s &= \left(\sum_{k=0}^{r} \binom{r}{k} x^k \right) \cdot \left(\sum_{h=0}^{s} \binom{s}{h} x^h \right) \\
&= \sum_{k=0}^{r} \sum_{h=0}^{s} \binom{r}{k} \binom{s}{h} x^{k+h} = \sum_{n=0}^{r+s} \sum_{\substack{0 \le k \le r,\, 0 \le h \le s \\ k+h=n}} \binom{r}{k} \binom{s}{h} x^{k+h} \\
&= \sum_{n=0}^{r+s} \sum_{0 \le k \le r,\, 0 \le n-k \le s} \binom{r}{k} \binom{s}{n-k} x^n = \sum_{n=0}^{r+s} \left(\sum_{k=0 \bigvee (n-s)}^{r \bigwedge n} \binom{r}{k} \binom{s}{n-k} \right) x^n.
\end{aligned}$$

D'altra parte, per la (20) con $n = r + s$,

$$(1+x)^{r+s} = \sum_{n=0}^{r+s} \binom{r+s}{n} x^n,$$

e la (21) si ottiene confrontando termine a termine tali due sviluppi. Ponendo in particolare, ad esempio, $r \leq s$ e $n = s$, e utilizzando il fatto che $\binom{s}{s-k} = \binom{s}{k}$ si ottiene

$$\sum_{k=0}^{r} \binom{r}{k}\binom{s}{k} = \binom{r+s}{s}$$

e potremo dunque anche scrivere

$$\sum_{k=0}^{r \wedge s} \binom{r}{k}\binom{s}{k} = \binom{r+s}{s} = \binom{r+s}{r}, \tag{22}$$

che per $r = s = n$ diviene

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}. \tag{23}$$

Esercizio proposto 3.3. Una moneta perfetta viene lanciata r volte da Renato e s volte da Stefano ($r > 3, s > 3$). Si ponga
$X =$ numero dei lanci in cui Renato ottiene il risultato testa
$Y =$ numero dei lanci in cui Stefano ottiene il risultato testa.
Calcolare la probabilità dei seguenti eventi.
(a) $\{X = 3\} \cap \{Y = 3\}$ (b) $\{X = 3\} \cup \{Y = 3\}$ (c) $\{\max(X, Y) = 3\}$ (d) $\{X = Y\}$.

Suggerimento per la soluzione dell'Esercizio proposto 3.3

Per calcolare la probabilità dell'evento $\{X = Y\}$ si consiglia di considerare il fatto che la moneta è perfetta e quindi, ai fini del calcolo delle probabilità è equivalente scambiare successo (esce testa) con insuccesso (esce croce) nel caso di uno dei due giocatori, ad esempio Stefano.
In altre parole si sta considerando che la probabilità di $\{X = Y\}$ è la stessa di $\{X = Y'\}$, dove $Y' = s - Y$ è il numero di croci ottenute da Stefano.
Ma allora $\{X = Y'\} = \{X = s - Y\} = \{X + Y = s\}$ e ciò permette di ottenere immediatamente, e con un ragionamento probabilistico, la relazione (22) (nota anche come formula di Vandermonde).

Richiamo: PROPRIETÀ DISTRIBUTIVA DELLA SOMMA RISPETTO AL PRODOTTO

Dati R ed S interi ($R, S \geq 1$), e $a_1, \dots, a_R$ e $b_1, \dots, b_S$, numeri reali, la sommatoria di tutti i prodotti del tipo $a_r\, b_s$ coincide con il prodotto delle somme di a_r per le somme di b_r:

$$\sum_{r=1}^{R}\sum_{s=1}^{S} a_r b_s = \left(\sum_{r=1}^{R} a_r\right)\left(\sum_{s=1}^{S} b_s\right)$$

Infatti basta osservare che

$$\begin{aligned}
\sum_{r=1}^{R}\sum_{s=1}^{S} a_r b_s &= a_1 b_1 + a_1 b_2 + \cdots a_1 b_S \\
&\quad + a_2 b_1 + a_2 b_2 + \cdots a_2 b_S \\
&\quad + \cdots \\
&\quad + a_R b_1 + a_R b_2 + \cdots a_R b_S \\
&= a_1 \big(b_1 + b_2 + \cdots b_S\big) \\
&\quad + a_2 \big(b_1 + b_2 + \cdots b_S\big) \\
&\quad + \cdots \\
&\quad + a_R \big(b_1 + b_2 + \cdots b_S\big) \\
&= \sum_{r=1}^{R} a_r \big(\textstyle\sum_{s=1}^{S} b_s\big) \\
&= \left(\sum_{s=1}^{S} b_s\right)\left(\sum_{r=1}^{R} a_r\right)
\end{aligned}$$

3.5 Il problema delle concordanze

Supponiamo di avere la seguente situazione: una segretaria ha n lettere indirizzate a n persone distinte ed n buste con già scritti gli n indirizzi: le cadono tutte le n lettere e quindi mette le n lettere a caso nelle n buste.

(i) **Quanto vale la probabilità $p(n)$ che nessuna lettera sia nella busta corrispondente?**

(ii) **Quanto vale la probabilità $q(n)$ che ci sia almeno una lettera nella busta corrispondente?**

In altre parole, se diciamo che c'è una *concordanza* quando una lettera viene messa nella busta corrispondente $p(n)$ è la probabilità che non ci siano concordanze, e $q(n)$ è la probabilità che ci sia almeno una concordanza. La probabilità in ***(i)*** è quindi il complemento a 1 della probabilità in ***(ii)***, ossia

$$p(n) = 1 - q(n)$$

Si tratta di un problema classico che ha anche altre diverse interpretazioni. Tutte però, in modo astratto, possono essere ricondotte alla seguente situazione: data un'urna che contiene n palline numerate da 1 ad n, si estraggono e palline una alla volta contando il numero d'ordine di estrazione.

Se, per un $h \in \{1, 2, ..., n\}$, si estrae la pallina con il numero h proprio all'h-sima estrazione si dice che c'è una concordanza. In altre parole stiamo assumendo che lo spazio Ω sia l'insieme di tutte le permutazioni $(j_1, j_2, ..., j_n)$ e si ha almeno una concordanza se $j_h = h$ per almeno un $h \in \{1, 2, ..., n\}$. Infine stiamo assumendo che le $n!$ permutazioni abbiano tutte la stessa probabilità.

Per calcolare $q(n)$ poniamo

$$A_h = \{\text{la lettera } h\text{-sima è nella busta } h\text{-sima}\},$$

di modo che per la formula di inclusione ed esclusione (vedere la nota a pagina 14 nella Lezione 7)

$$\begin{aligned} q(n) &= \mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) \\ &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{\{i_1, i_2, \cdots, i_k\}} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \end{aligned}$$

ora, qualunque sia la combinazione $\{i_1, i_2, \cdots, i_k\}$ di $\{1, 2, \cdots, N\}$, si ha[10]

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}) = \frac{(n-k)!}{n!}$$

e quindi la somma

$$\sum_{\{i_1, i_2, \cdots, i_k\}} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})$$

fatta su tutte le $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ combinazioni di n elementi di classe k vale

$$\frac{n!}{k!(n-k)!} \, \frac{(n-k)!}{n!} = \frac{1}{k!}$$

di conseguenza

$$q(n) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{k-1} \frac{1}{k!} + \cdots + (-1)^{n-1} \frac{1}{n!}$$

e

$$p(n) = 1 - q(n) = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} + \cdots + (-1)^k \frac{1}{k!} + \cdots + (-1)^n \frac{1}{n!} = \sum_{k=0}^{n} (-1)^k \frac{1}{k!}$$

[10]È importante osservare che stiamo chiedendo che a ciascuno dei posti i_1, i_2, $\cdots$, i_k compaia il numero corrispondente, ma non stiamo escludendo che nei posti rimanenti ci possa essere qualche altro *punto fisso* .
Quindi, una volta fissati i k valori i_1, i_2, ..., i_k, nei rimanenti $n-k$ posti possiamo mettere i rimanenti valori in tutti gli $(n-k)!$ ordinamenti possibili.
Si consiglia di pensare ad un esempio particolare, ad esempio $n = 7$, $k = 3$ e $i_1 = 2$, $i_2 = 3, i_3 = 5$: stiamo contando le permutazioni di $\{1, 2, 3, 4, 5, 6, 7\}$ del tipo $(j_1, 2, 3, j_4, 5, j_6, j_7)$ dove j_1, j_4, j_6, j_7 possono variare solo tra i numeri $\{1, 4, 6, 7\}$ e di fatto (j_1, j_4, j_6, j_7) è una permutazione degli elementi $\{1, 4, 6, 7\}$, che formano appunto un insieme di $7 - 3 = 4$ elementi e sono quindi in totale $(7-3)! = 4!$ permutazioni.

È interessante notare che $p(n)$ converge (per n che tende ad infinito) a e^{-1} (ossia poco più di un terzo, anzi circa $\frac{1}{2{,}71}$):
basta ricordare lo sviluppo in serie di Taylor di e^x in $x_0 = 0$, ovvero lo sviluppo di Mac Laurin:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \qquad \text{da cui} \qquad e^{-x} = \sum_{k=0}^{\infty} (-1)^k \frac{x^k}{k!}.$$

Per controllo possiamo calcolare $p(n)$ per $n = 3,\ 4,\ 5$:

$$p(3) = 1 - 1/1! + 1/2! - 1/3! = 1/2 - 1/6 = 2/3$$
$$p(4) = 1 - 1/1! + 1/2! - 1/3! + 1/4! = 1/2 - 1/6 + 1/24 = 17/24$$
$$p(5) = 1 - 1/1! + 1/2! - 1/3! + 1/4! - 1/5! = 1/2 - 1/6 + 1/24 - 1/120 = 84/120$$

Inoltre è interessante osservare che il numero delle permutazioni $(i_1, i_2, ...i_n)$ di n elementi che non presentano punti fissi, ossia per le quali, qualunque sia h, il valore i_h è diverso da h, ovvero senza concordanze vale $n!p(n)$, infatti basta moltiplicare $n!$ per la probabilità $p(n)$ che non ci siano concordanze.

3.6 Approfondimenti sul calcolo combinatorio

Una dimostrazione elementare di
$\sum_{k=0}^{n} C_k^n = \sum_{k=0}^{n} \binom{n}{k} = 2^n$

La relazione

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

si può dedurre anche nel seguente modo (senza conoscere la formula del binomio di Newton): l'insieme delle parti $\mathcal{P}(\{a_1, \cdots, a_n\}$ di $\{a_1, \cdots, a_n\}$ è l'unione delle famiglie dei sottoinsiemi di cardinalità k al variare di k da 0 ad n, per cui la cardinalità dell'insieme delle parti, $|\mathcal{P}(\{a_1, \cdots, a_n\}|$, soddisfa la relazione seguente

$$\sum_{k=0}^{n} C_k^n = |\mathcal{P}(\{a_1, \cdots, a_n\})|.$$

Già sappiamo che $C_k^n = \binom{n}{k}$, inoltre si può ricavare facilmente che $\mathcal{P}(\{a_1, \cdots, a_n\})$ ha cardinalità 2^n: infatti $\mathcal{P}(\{a_1, \cdots, a_n\}$ è in corrispondenza biunivoca con l'insieme delle funzioni $g : \{a_1, \cdots, a_n\} \mapsto \{0, 1\}$. La corrispondenza è data da $G \longleftrightarrow g = \mathbf{1}_G$, dove

$$\mathbf{1}_G(a_i) = 1, \qquad se \quad a_i \in G$$

$$\mathbf{1}_G(a_i) = 0, \qquad se \quad a_i \notin G.$$

Queste ultime sono tante quanti sono gli elementi del prodotto cartesiano di $\{0, 1\} \times \{0, 1\} \times \cdots \{0, 1\}$ (n volte) e sono quindi 2^n. Per capire meglio la corrispondenza poniamo $n = 4$ e consideriamo il sottoinsieme $G = \{a_1, a_3\}$:

$$\{a_1, a_3\} \qquad \Longleftrightarrow \qquad \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad \Longleftrightarrow \qquad (1, 0, 1, 0) \in \{0, 1\}^4.$$

Ancora, ad esempio $(0, 0, 0, 0)$ corrisponde all'insieme vuoto, e $(1, 1, 1, 1)$ a tutto $A = \{a_1, a_2, a_3, a_4\}$. Nel caso generale di $A = \{a_1, a_2, \ldots, a_n\}$, la corrispondenza tra i sottoinsiemi di A e gli elementi di $\{0, 1\}^n$ dovrebbe essere quindi chiara, e di conseguenza l'uguaglianza $|\mathcal{P}(\{a_1, \cdots, a_n\})| = 2^n$.

Dimostrazione della formula di Stiefel, senza l'uso dei fattoriali

Prendendo come interpretazione di $\binom{n}{k}$ il numero C_k^n dei sottoinsiemi di cardinalità k di un insieme $\{a_1, a_2, \cdots a_{n-1}, a_n\}$ di cardinalità n, la formula di Stiefel diviene

$$C_k^n = C_{k-1}^{n-1} + C_k^{n-1}$$

Per dimostrare la precedente uguaglianza si può ragionare anche nel seguente modo (attenzione, si noti che ***non è necessario sapere il valore esplicito di C_k^n***, cioè nella dimostrazione ***non si usa*** il fatto che $C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}$).

I sottoinsiemi di cardinalità k si possono dividere in due classi:
1) i sottoinsiemi C che contengono a_n
2) i sottoinsiemi D che non contengono a_n.
Quindi il numero C_k^n dei sottoinsiemi di cardinalità k si può esprimere come la somma del numero dei sottoinsiemi del primo tipo e del numero dei sottoinsiemi del secondo tipo.
D'altra parte
I) gli insiemi C del primo tipo si possono esprimere come $C = C' \cup \{a_n\}$, con C' sottoinsieme di $\{a_1, a_2, \cdots a_{n-1}\}$ che ha cardinalità $k-1$, e quindi sono tanti quanti i sottoinsiemi di cardinalità $k-1$ di un insieme di cardinalità $n-1$, ovvero C_{k-1}^{n-1},
II) gli insiemi D del secondo tipo, sono sottoinsiemi di $\{a_1, a_2, \cdots a_{n-1}\}$ di cardinalità k e quindi sono esattamente C_k^{n-1}.

Come si contano le disposizioni-I

In questa nota vogliamo vedere come si può arrivare a contare sia le disposizioni che le combinazioni. Iniziamo con un esempio.
Consideriamo $n = 5$ e $k = 1, 2, 3$ e proviamo a scrivere tutte le disposizioni di 5 elementi di classe k. Indichiamo il nostro insieme di 5 elementi con $\{1, 2, 3, 4, 5\}$. Ovviamente tutte le disposizioni di classe 1 sono $D_1^5 = 5$:

(1)	(2)	(3)	(4)	(5)

Per ottenere le disposizioni di classe 2 si può procedere mettendo insieme le disposizioni che iniziano per 1, quelle che iniziano per 2, etc. come segue

(1,2)	(2,1)	(3,1)	(4,1)	(5,1)
(1,3)	(2,3)	(3,2)	(4,2)	(5,2)
(1,4)	(2,4)	(3,4)	(4,3)	(5,3)
(1,5)	(2,5)	(3,5)	(4,5)	(5,4)

Vengono quindi 20 disposizioni, in quanto si tratta di 5 colonne ciascuna di lunghezza 4: infatti una volta scelto il primo elemento, rimangono $5 - 1 = 4$ scelte per il secondo elemento.
Per ottenere le disposizioni di classe 3 possiamo ancora dividere le disposizioni mettendo insieme le disposizioni che iniziano per una disposizione (i, j) di classe 2. Le disposizioni di classe 3 che iniziano per tali elementi sono del tipo (i, j, k) con k che varia nell'insieme $\{1,2,3,4,5\}\setminus\{i, j\}$ di cardinalità $5 - 2 = 3$. Otteremo quindi $60 = 20 \cdot 3$ disposizioni.

(1,2,3) (1,2,4) (1,2,5)	(2,1,3) (2,1,4) (2,1,5)	(3,1,2) (3,1,4) (3,1,5)	(4,1,2) (4,1,3) (4,1,5)	(5,1,2) (5,1,3) (5,1,4)
(1,3,2) (1,3,4) (1,3,5)	(2,3,1) (2,3,4) (2,3,5)	(3,2,1) (3,2,4) (3,2,5)	(4,2,1) (4,2,3) (4,2,5)	(5,2,1) (5,2,3) (5,2,4)
(1,4,2) (1,4,3) (1,4,5)	(2,4,1) (2,4,3) (2,4,5)	(3,4,1) (3,4,2) (3,4,5)	(4,3,1) (4,3,2) (4,3,5)	(5,3,1) (5,3,2) (5,3,4)
(1,5,2) (1,5,3) (1,5,4)	(2,5,1) (2,5,3) (2,5,4)	(3,5,1) (3,5,2) (3,5,4)	(4,5,1) (4,5,2) (4,5,3)	(5,4,1) (5,4,2) (5,4,3)

Come si contano le disposizioni-II

Quanto fatto nella nota di approfondimento precedente si generalizza al caso n e k (con $1 \leq k \leq n$) ottenendo

$$D_1^n = n, \qquad \text{e la formula ricorsiva} \qquad D_k^n = D_{k-1}^n \big(n-(k-1)\big)$$

da cui si ricava immediatamente che

$$D_1^n = n, \quad D_2^n = n(n-1), \quad D_3^n = n(n-1)(n-2), \quad D_k^n = n(n-1)\cdots\big(n-(k-1)\big),$$

ossia $\quad D_k^n = \frac{n!}{(n-k)!} \qquad$ e per le permutazioni $\qquad P_n = D_n^n = n!$

Relazioni tra disposizioni e combinazioni-I

Abbiamo visto che per trovare il numero C_k^n delle combinazioni di n elementi di classe k, basta osservare che

$$D_k^n = C_k^n \cdot P_k,$$

da cui immediatamente

$$C_k^n = \frac{D_k^n}{P_k} = \frac{n!}{k!\,(n-k)!}.$$

Iniziamo con il caso particolare $n = 5$ e $k = 3$. Riscriviamo tutte le 60 disposizioni mettendo in ogni riga tutte quelle che contengono gli stessi tre elementi, e poi permutiamo i tre elementi in tutti i modo possibili.

{1,2,3}	(1,2,3)	(1,3,2)	(2,1,3)	(2,3,1)	(3,1,2)	(3,2,1)
{1,2,4}	(1,2,4)	(1,4,2)	(2,1,4)	(2,4,1)	(4,1,2)	(4,2,1)
{1,2,5}	(1,2,5)	(1,5,2)	(2,1,5)	(2,5,1)	(5,1,2)	(5,2,1)
{1,3,4}	(1,3,4)	(1,4,3)	(3,1,4)	(3,4,1)	(4,1,3)	(4,3,1)
{1,3,5}	(1,3,5)	(1,5,3)	(3,1,5)	(3,5,1)	(5,1,3)	(5,3,1)
{1,4,5}	(1,4,5)	(1,5,4)	(4,1,5)	(4,5,1)	(5,1,4)	(5,4,1)
{2,3,4}	(2,3,4)	(2,4,3)	(3,2,4)	(3,4,2)	(4,2,3)	(4,3,2)
{2,3,5}	(2,3,5)	(2,5,3)	(3,2,5)	(3,5,2)	(5,2,3)	(5,3,2)
{2,4,5}	(2,4,5)	(2,5,4)	(4,2,5)	(4,5,2)	(5,2,4)	(5,4,2)
{3,4,5}	(3,4,5)	(3,5,4)	(4,3,5)	(4,5,3)	(5,3,4)	(5,4,3)

Abbiamo ottenuto quindi una tabella con 10 righe, ossia il numero dei sottoinsiemi di cardinalità 3 dell'insieme $\{1,2,3,4,5\}$ e 6 colonne, in quanto ogni riga ha 6 elementi, in quanto 6 sono le permutazioni di tre elementi. Questa tabella contiene effettivamente tutte le 60 disposizioni.

Relazioni tra disposizioni e combinazioni-II

In generale, per dimostrare che il numero delle disposizioni D_k^n è il prodotto del numero delle combinazioni C_k^n per il numero delle permutazioni P_k si può procedere osservando che le disposizioni di n elementi di classe k si possono raggruppare in modo che ciascun gruppo è formato da disposizioni che contengono gli stessi elementi, ma differiscono solo per l'ordine: ognuno di tali gruppi individua quindi il sottoinsieme degli elementi comuni, ovvero una combinazione di classe k di n elementi.
Chiaramente ciascuno dei C_k^n gruppi è composto dallo stesso numero (P_k) di disposizioni: da ciascuna combinazione di classe k di n elementi si ottengono P_k disposizioni diverse, permutando tra loro i k elementi distinti che compongono la combinazione. Quest'ultima osservazione significa appunto che vale la relazione $D_k^n = C_k^n \cdot P_k$.

3.7 Esercizi di verifica

Esercizio 3.1. Le lettere AAMMM vengono ordinate a caso. Qual è la probabilità di ottenere la parola MAMMA?

Esercizio 3.2. Si fanno n lanci di una moneta perfetta. Per $1 \leq h \leq n$, qual è la probabilità di ottenere il risultato testa per la prima volta all'h-esimo lancio?

Esercizio 3.3. Da un'urna, che contiene 6 oggetti numerati da 1 a 6, si estraggono a caso tre oggetti contemporaneamente. Qual è la probabilità che il minimo numero estratto sia superiore a 2?

Esercizio 3.4. In una mano del gioco della roulette si punta su {*pair*}, {*passe*}, {*16*}. Qual è la probabilità di vincere almeno una di queste puntate?

Esercizio 3.5. Qual è la probabilità che il numero 16 esca almeno una volta su cinque mani del gioco della roulette?

Esercizio 3.6. Qual è la probabilità che esca il numero 16 in una delle cinque estrazioni su una ruota del lotto? (Si estrae senza reinserimento da un'urna contenente i numeri {1,2, ..., 90}).

Esercizio 3.7. Qual è la probabilità che esca la coppia di numeri 16 e 48 nelle cinque estrazioni su una ruota del lotto?

Esercizio 3.8. Qual è la probabilità che esca la terna di numeri 16, 48, 90 nelle cinque estrazioni su una ruota del lotto?

Esercizio 3.9. Vengono lanciati contemporaneamente 5 dadi perfetti.
Calcolate la probabilità degli eventi elencati qui di seguito:
a) {*tutti i dadi danno punteggi diversi fra loro*}
b) {*due dadi danno punteggi uguali fra loro e gli altri tre danno punteggi tutti diversi*}
("coppia")
c) {*tre dadi danno punteggi uguali fra loro e gli altri due danno due punteggi diversi*}
("tris")
d) {*quattro dadi danno punteggi uguali fra loro e uno da un punteggio diverso*}
("poker")

e) $\{$*tutti i dadi danno lo stesso punteggio*$\}$
("jazzi")
f) $\{$*due diverse coppie di punteggi fra loro uguali e un punteggio diverso dagli altri due*$\}$
("doppia coppia")
g) $\{$*tre punteggi uguali fra loro e gli altri due uguali fra loro e diversi dal precedente*$\}$
("full").

Esercizio 3.9 bis. Riformulate da soli ove possibile, con gli opportuni cambiamenti, e poi risolvete l'analogo dell'esercizio precedente per il caso del lancio di soli tre dadi. (Questo esercizio si può saltare se non si sono trovate eccessive difficoltà a risolvere completamente l'esercizio precedente).

Esercizio 3.10. Un servizio da tè consiste di quattro tazzine e quattro piattini con due tazzine e due piattini di un colore e i rimanenti di un altro colore. Le tazzine sono poste a caso sopra i piattini. Calcolare le probabilità degli eventi:

$$\{\textit{Nessuna tazzina è su un piattino dello stesso colore}\}$$

$$\{\textit{Una sola tazzina è su un piattino dello stesso colore}\}$$

$$\{\textit{Due sole tazzine sono su un piattino dello stesso colore}\}$$

Calcolare la probabilità dell'evento

$$\{\textit{Nessuna tazzina su un piattino dello stesso colore}\}$$

se il servizio è composto di quattro tazzine e quattro piattini di quattro colori diversi.

4 Probabilità condizionate

In questa lezione verrà introdotta la definizione di probabilità condizionata e ne verranno illustrate alcune conseguenze immediate: la ***Formula delle probabilità composte***, la ***Formula delle probabilità totali*** e la ***Formula di Bayes***.

Abbiamo già visto che, in uno spazio di probabilità finito, la teoria della probabilità è in effetti già tutta contenuta nella formula (4), che mostra come si ottenga la probabilità di un evento composto, una volta assegnate le probabilità a ciascun evento elementare. Si vedrà comunque che le formule che verranno ottenute nel seguito (e la nozione di indipendenza stocastica che verrà illustrata a partire dalla prossima lezione) costituiscono spesso una guida al ragionamento probabilistico, che può rivelarsi complementare all'uso della formula (4). Tali nozioni infatti permettono, alcune volte, di assegnare probabilità ad eventi composti (oppure di calcolarle sulla base di probabilità assegnate ad altri eventi) in modo più diretto, senza necessariamente far intervenire tutta la collezione degli eventi semplici. Vedremo nelle successive lezioni, in particolare, come si possano risolvere, in modo alternativo, alcuni degli esercizi affrontati nella lezione precedente.

Prima di iniziare tale studio è opportuno ricordare il significato "logico" della nozione di ***partizione di un insieme***.

Prima di tutto ricordiamo che una collezione di sottoinsiemi $H_1, ..., H_m$ dello spazio Ω ($H_l \in \mathcal{P}(\Omega), l = 1, ..., m$), costituisce una partizione di Ω, se e solo se è tale che

$$\bigcup_{l=1}^{m} H_l = \Omega; \;\; H_{l_1} \cap H_{l_2} = \emptyset, \text{ per } l_1 \neq l_2.$$

Interpretando $H_1, ..., H_m$ come eventi, abbiamo che essi sono a ***due a due incompatibili*** (cioè è *certo che non se ne possono verificare due contemporaneamente*) e, d'altra parte, essi sono ***esaustivi*** (*è certo che se ne verifichi almeno uno*); dunque: è certo che si verifichi uno ed uno soltanto degli eventi $H_1, ..., H_m$ (la nostra situazione di incertezza risiede nel fatto che non sappiamo **quale** di essi sia verificato).

Come si era visto quale immediata conseguenza degli assiomi della probabilità si ha che, se $H_1, ..., H_m$ costituisce una partizione di Ω, allora deve risultare

$$\sum_{l=1}^{m} \mathbb{P}(H_l) = 1.$$

Inoltre osserviamo che, per qualunque evento $E \in \mathcal{P}(\Omega)$, possiamo scrivere

$$E = (E \cap H_1) \cup ... \cup (E \cap H_m)$$

e dunque

$$\mathbb{P}(E) = \mathbb{P}\left(E \cap H_1\right) + ... + P\left(E \cap H_m\right). \tag{24}$$

4.1 Definizione di probabilità condizionata

Cominciamo con un esempio

Esempio 4.1. *In un lancio di un dado a sei facce, quale probabilità dobbiamo assegnare all'evento* $A \equiv \{X \; dispari\}$*, sapendo che si è verificato l'evento* $B \equiv \{X \geq 2\}$*?*

Soluzione. Tutti gli eventi elementari

$$\{X = 1\}, \quad \{X = 2\}, ..., \quad \{X = 6\}$$

sono inizialmente ritenuti equiprobabili.

Sapere che si è verificato l'evento B equivale a sapere che si è verificato uno dei seguenti eventi elementari:

$$\{X = 2\}, ..., \{X = 6\} \quad \text{(non sappiamo però “quale”).}$$

È naturale, a questo punto, **assumere** quanto segue:
l'informazione che si è verificato l'evento B non modifica la situazione di equiprobabilità fra gli eventi $\{X = 2\}, ..., \{X = 6\}$.

A seguito di tale informazione, quindi, la probabilità di osservare l'evento A deve essere dunque valutata come la probabilità del verificarsi di uno fra 2 eventi elementari favorevoli su un totale di 5 eventi elementari possibili, equiprobabili fra loro; valuteremo quindi tale probabilità “condizionata” uguale a $\frac{2}{5}$.

La soluzione del precedente esempio mostra che, nel caso di un numero finito di eventi elementari equiprobabili, è naturale imporre che la probabilità da attribuire ad un evento A, quando si sappia per certo che si è verificato un evento B, sia data da

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

Ciò suggerisce la seguente

Definizione 4.1 (Probabilità condizionata)**.** *Siano E ed H due eventi, con $\mathbb{P}(H) > 0$. Viene detta* **probabilità condizionata di E dato** *H, ed indicata con il simbolo $\mathbb{P}(E|H)$, la quantità*

$$\mathbb{P}(E|H) = \frac{\mathbb{P}(E \cap H)}{\mathbb{P}(H)}. \tag{25}$$

Osservazione 1 **(di carattere euristico)**. All'interno di ciascuna delle interpretazioni della probabilità (classica, frequentista, soggettivista, ...) cui si è accennato in precedenza, il numero $\mathbb{P}(E|H)$ definito nella (25) coincide effettivamente con la probabilità che, coerentemente con tale interpretazione, dovremmo attribuire al verificarsi di E, se sapessimo che si è verificato H.

Ciò costituisce la motivazione per definire “assiomaticamente” la nozione di probabilità condizionata attraverso la (25).

Esercizio proposto 4.1 (continuazione dell'Esempio 2.2, dado non equilibrato)**.** *Un dado ha sei facce numerate da 1 a 6; esso è pesato in modo tale che ciascuna faccia abbia una probabilità di presentarsi (in un singolo lancio) proporzionale al suo valore. Siano*

$$A \equiv \{\text{Si presenta un numero pari}\}, \qquad B \equiv \{\text{Si presenta un numero primo}\}.$$

Calcolare $\mathbb{P}(B|A)$ e $\mathbb{P}(A|B)$.

Vediamo ora le semplici, ma importanti, conseguenze della definizione di probabilità condizionata, già menzionate in precedenza.

4.2 Conseguenze immediate della definizione di probabilità condizionata

4.2.1 Formula delle probabilità composte

Dalla definizione di probabilità condizionata si ottiene immediatamente che, se $\mathbb{P}(E_1) > 0$, allora

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_2|E_1)\,\mathbb{P}(E_1). \tag{26}$$

Questa formula si può generalizzare.

Proposizione 1. **(Formula delle probabilità composte)** Consideriamo n eventi E_1, E_2,..., E_n, tali che

$$\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_{n-1}) > 0. \tag{27}$$

Si ha

$$\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_n) = \mathbb{P}(E_1) \cdot \mathbb{P}(E_2|E_1) \cdot \mathbb{P}(E_3|E_1 \cap E_2) \cdots \mathbb{P}(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1}). \tag{28}$$

Dimostrazione. Iniziamo con l'osservare che, essendo $E_1 \supseteq E_1 \cap E_2 \supseteq \cdots \supseteq E_1 \cap E_2 \cap \cdots \cap E_{n-1}$, per la proprietà di monotonia della probabilità si ha $\mathbb{P}(E_1) \geq \mathbb{P}(E_1 \cap E_2) \geq \cdots \geq \mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_{n-1}) > 0$. Quindi la condizione (27) implica che $\mathbb{P}(E_1) > 0$, $\mathbb{P}(E_1 \cap E_2) > 0$, ..., $\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_{n-1}) > 0$, per cui il prodotto a destra della (28) ha senso.

L'uguaglianza (28) segue immediatamente dalla definizione di probabilità condizionata: possiamo scrivere

$$\mathbb{P}(E_1 \cap E_2 \cap ... \cap E_n) = \mathbb{P}(E_n|E_1 \cap E_2 \cap ... \cap E_{n-1})\,\mathbb{P}(E_1 \cap E_2 \cap ... \cap E_{n-1})$$

A sua volta $\mathbb{P}(E_1 \cap E_2 \cap ... \cap E_{n-1})$ può essere scritto come

$$\mathbb{P}(E_{n-1}|E_1 \cap E_2 \cap ... \cap E_{n-2})\,\mathbb{P}(E_1 \cap E_2 \cap ... \cap E_{n-2}).$$

La dimostrazione quindi si ottiene facilmente proseguendo così di seguito, fino a scrivere

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_2|E_1)\,\mathbb{P}(E_1).$$

La precedente dimostrazione si può formalizzare utilizzando il principio di induzione.
Il caso $n = 2$ corrisponde alla formula (26). Supposta vera l'affermazione della Proposizione per $m - 1$ eventi, ovvero

$$\begin{aligned}
&\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_{m-1}) = \\
&= \mathbb{P}(E_1) \cdot \mathbb{P}(E_2|E_1) \cdot \mathbb{P}(E_3|E_1 \cap E_2) \cdot \cdots \cdot \mathbb{P}(E_{m-1}|E_1 \cap E_2 \cap \cdots \cap E_{m-2}).
\end{aligned}$$

mostriamo ora che vale per m eventi, e infatti

$$\begin{aligned}
&\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_{m-1} \cap E_m) \\
&= \mathbb{P}\big((E_1 \cap E_2 \cap \cdots \cap E_{m-1}) \cap E_m\big) \\
&= \mathbb{P}\big(E_1 \cap E_2 \cap \cdots \cap E_{m-1}\big) \cdot \mathbb{P}(E_m|E_1 \cap E_2 \cap \cdots \cap E_{m-1}) \\
&= \big[\,\mathbb{P}(E_1) \cdot \mathbb{P}(E_2|E_1) \cdot \mathbb{P}(E_3|E_1 \cap E_2) \cdot \cdots \cdot \mathbb{P}(E_{m-1}|E_1 \cap E_2 \cap \cdots \cap E_{m-2})\big] \\
&\quad \cdot \mathbb{P}(E_m|E_1 \cap E_2 \cap \cdots \cap E_{m-1}).
\end{aligned}$$

□

La formula delle probabilità composte di solito viene usata per trovare la probabilità dell'intersezione di un numero finito di eventi, specialmente quando è più facile valutare le probabilità condizionate rispetto alle probabilità dell'intersezione. Questa idea viene illustrata nel seguente esempio (confrontare anche le osservazioni al successivo Esempio 5.2 della Lez. 5, e ancora l'Osservazione 1 della Lez. 6).

Esempio 4.2. *Un uomo ha un mazzo di n chiavi, una sola delle quali apre un porta. Egli prova le chiavi a caso ad una ad una, escludendo dal mazzo quelle già provate, finché non trova la chiave giusta. Vogliamo trovare la probabilità dell'evento*

$$E \equiv \{\textit{chiave giusta al } k\textit{-esimo tentativo}\}$$

Soluzione. Scriviamo E come intersezione di diversi eventi come segue:
$E \equiv \{$*chiave errata al primo tentativo*$\} \cap \{$ *chiave errata al secondo tentativo*$\} \cap ... \cap \{$*chiave giusta al* k*-esimo tentativo*$\}$.

Utilizzando la formula delle probabilità composte possiamo scrivere dunque

$$\mathbb{P}(E) = \mathbb{P}(\textit{chiave errata al primo tentativo})\cdot$$

$$\mathbb{P}(\{\textit{chiave errata al secondo tentativo }\} \,|\, \{\textit{ chiave errata al primo tentativo}\}) \cdot ...\cdot$$

$$\cdot ... \cdot \mathbb{P}(\{\textit{chiave giusta al } k\textit{-esimo tentativo}\} \,|\, \{\textit{chiave errata al primo tentativo}\} \cap ...\cap$$

$$... \cap \{\textit{chiave errata al } (k-1)\textit{-esimo tentativo}\}) =$$

$$= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot ... \cdot \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} = \frac{1}{n}.$$

In simboli, ponendo $C_h = \{\textit{chiave giusta al } k\textit{-esimo tentativo}\}$

$$E = \overline{C}_1 \cap \overline{C}_2 \cap \cdots \cap \overline{C}_{k-1} \cap C_k$$

e quindi

$$\begin{aligned}
\mathbb{P}(E) &= \mathbb{P}(\overline{C}_1)\,\mathbb{P}(\overline{C}_2|\overline{C}_1)\cdots\mathbb{P}(\overline{C}_{k-1}|\overline{C}_1 \cap \overline{C}_2 \cap \cdots \cap \overline{C}_{k-2})P(C_k|\overline{C}_1 \cap \overline{C}_2 \cap \cdots \cap \overline{C}_{k-1}) \\
&= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot ... \cdot \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} = \frac{1}{n}.
\end{aligned}$$

4.2.2 Formula delle probabilità totali

***Proposizione* 4.1.** *Sia $\{H_1, ..., H_m\}$ una partizione di Ω, con $P(H_l) > 0$, $l = 1, ..., m$. Per un qualunque evento $E \in \mathcal{P}(\Omega)$, risulta*

$$\mathbb{P}(E) = \mathbb{P}\left(E|H_1\right)\mathbb{P}\left(H_1\right) + ... + \mathbb{P}\left(E|H_m\right)\mathbb{P}\left(H_m\right),$$

o più brevemente

$$\mathbb{P}(E) = \sum_{k=1}^{m} \mathbb{P}\left(E|H_k\right)\mathbb{P}\left(H_k\right).$$

Dimostrazione. Basta ricordare la precedente formula (24), ovvero $\mathbb{P}(E) = \sum_{k=1}^{m} \mathbb{P}(E \cap H_k)$, e tener conto che, essendo $\mathbb{P}(H_l) > 0$, si ha, per $l = 1, ..., m$, $\mathbb{P}\left(E \cap H_l\right) = \mathbb{P}\left(E|H_l\right)\mathbb{P}\left(H_l\right)$.

□

Come nel caso della formula delle probabilità composte, anche la formula delle probabilità totali si mostra particolarmente utile quando è più semplice valutare le probabilità condizionate $\mathbb{P}(E|H_l)$ che quelle dell'intersezione $\mathbb{P}(E \cap H_l)$.

Esempio 4.3 (Estrazione di un numero al lotto)**.** *Qual è la probabilità che sia uguale a 16 il secondo estratto su una ruota del lotto?*

Soluzione. Indichiamo con X_1, X_2 i valori rispettivamente ottenuti nella prima e nella seconda estrazione e poniamo

$$E \equiv \{X_2 = 16\}.$$

Consideriamo la partizione $\{H, \overline{H}\}$, dove

$$H \equiv \{X_1 = 16\},$$

e applichiamo la formula delle probabilità totali; si ha così

$$\mathbb{P}(E) = \mathbb{P}(E|H)\,\mathbb{P}(H) + \mathbb{P}(E|\overline{H})\,\mathbb{P}(\overline{H}),$$

da cui otteniamo facilmente

$$\mathbb{P}(E) = 0 \times \frac{1}{90} + \frac{1}{89} \times \frac{89}{90} = \frac{1}{90}.$$

Esempio 4.4. *Una moneta perfetta viene lanciata n volte. Qual è la probabilità di ottenere un numero pari di risultati testa?*

Soluzione. Poniamo, per $k = 1, 2, ..., n$

$$E_k \equiv \{\text{numero pari di risultati testa sui primi } k \text{ lanci}\}.$$

Si capisce subito che, se la moneta è perfetta, si avrà per motivi di simmetria,

$$\mathbb{P}(E_n) = \mathbb{P}(\overline{E}_n) = \frac{1}{2}.$$

È però utile anche ragionare come segue, adoperando la formula delle probabilità totali:

$$\mathbb{P}(E_n) = \mathbb{P}(E_{n-1}) \cdot \mathbb{P}(E_n|E_{n-1}) + \mathbb{P}(\overline{E}_{n-1}) \cdot \mathbb{P}(E_n|\overline{E}_{n-1}), \tag{29}$$

e risulta

$$\mathbb{P}(E_n|E_{n-1}) = \mathbb{P}(\{\textit{risultato croce all'}n\textit{-esimo lancio}\}|E_{n-1})$$

$$\mathbb{P}(E_n|\overline{E}_{n-1}) = \mathbb{P}(\{\textit{risultato testa all'}n\textit{-esimo lancio}\}|\overline{E}_{n-1})$$

Il fatto che la moneta sia perfetta ci porterà ora a valutare:

$$\mathbb{P}(E_n|E_{n-1}) = \mathbb{P}(E_n|\overline{E}_{n-1}) = \frac{1}{2}.$$

Otteniamo dunque dalla (29)

$$\mathbb{P}(E_n) = \frac{1}{2}\left(\mathbb{P}(E_{n-1}) + \mathbb{P}(\overline{E}_{n-1})\right) = \frac{1}{2}.$$

4.2.3 Formula di Bayes

Applicando la definizione di probabilità condizionata e poi la formula delle probabilità composte si ottiene, se $P(E) > 0$ e $P(H) > 0$,

$$\mathbb{P}(H|E) = \frac{\mathbb{P}(H \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|H)\,\mathbb{P}(H)}{\mathbb{P}(E)},$$

che rappresenta la ***forma elementare della formula di Bayes***. Quando $H = H_l$, dove $\{H_1, ..., H_m\}$ è una partizione dell'evento certo questa formula si generalizza nel seguente modo.

***Proposizione 3*. (Formula di Bayes)** Sia ancora $\{H_1, ..., H_m\}$ una partizione di Ω, con $P(H_l) > 0$, $l = 1, ..., m$. Per un qualunque evento $E \in \mathcal{P}(\Omega)$, risulta

$$\mathbb{P}(H_l|E) = \frac{\mathbb{P}(E|H_l)\,\mathbb{P}(H_l)}{\sum_{r=1}^{m} \mathbb{P}(E|H_r)\,\mathbb{P}(H_r)}, \quad l = 1, ..., m. \tag{30}$$

Dimostrazione.
Per la definizione di probabilità condizionata, si ha

$$\mathbb{P}(H_l|E) = \frac{\mathbb{P}(H_l \cap E)}{\mathbb{P}(E)}.$$

Applicando la formula delle probabilità composte al numeratore del membro a destra otteniamo

$$\mathbb{P}(H_l \cap E) = \mathbb{P}(E|H_l)\,\mathbb{P}(H_l)\,;$$

applicando la formula delle probabilità totali al denominatore otteniamo

$$\mathbb{P}(E) = \sum_{r=1}^{m} \mathbb{P}(E|H_r)\,\mathbb{P}(H_r)\,.$$

***Osservazione 2* (di carattere euristico)**. La formula di Bayes trova naturale applicazione nei problemi in cui si debba analizzare come un'"osservazione" porti a modificare lo stato di informazione sugli eventi di una partizione; spesso problemi di tale tipo sono originati da questioni di "inferenza statistica".

Fissiamo l'attenzione su una partizione di $\Omega, \mathcal{H} \equiv \{H_1, ..., H_m\}$: sappiamo che è verificato uno ed uno soltanto degli eventi $H_1, ..., H_m$, ma non sappiamo quale.

Attribuiamo, rispettivamente, probabilità $\mathbb{P}(H_1), ..., \mathbb{P}(H_m)$ a ciascuno di tali eventi (possiamo pensare che tali probabilità esprimono il nostro stato di informazione "iniziale" su tale partizione).

Supponiamo poi di avere l'informazione che è verificato l'evento E e ci chiediamo come, in conseguenza di ciò, si debbano modificare le probabilità da attribuire agli eventi $H_1, ..., H_m$ (cioè come ciò modifichi il nostro stato di informazione "iniziale" su $\mathcal{H}$).

Tali "nuove" probabilità coincideranno con le probabilità condizionate $\mathbb{P}(H_l|E)$ ($l = 1, ..., m$), che vanno calcolate attraverso la formula (30). Dunque la formula di Bayes può essere vista come la regola secondo cui lo stato di informazione su $\mathcal{H}$ si modifica sulla base dell'osservazione dell'evento E.

Più in generale, si può considerare la nuova probabilità

$$\mathbb{P}_E : \mathcal{P}(\Omega) \mapsto [0, 1];\ A \mapsto \mathbb{P}_E(A) := \mathbb{P}(A|E).$$

È facile verificare che $\mathbb{P}_E$ così definita è una misura di probabilità su $(\Omega, \mathcal{P}(\Omega))$. Si lascia al lettore la verifica, di queste proprietà.

Esercizio proposto 4.2 ($\mathbb{P}_E$ è una probabilità). *Controllare che $\mathbb{P}_E$ verifica gli assiomi i), ii) e iii) della Definizione 2.1, ovvero che*
i) $\mathbb{P}_E(A) \in [0,1]$
ii) $\mathbb{P}_E(\Omega) = 1$,
iii) se A_1 e A_2 sono eventi incompatibili, cioè $A_1 \cap A_2 = \emptyset$, allora $\mathbb{P}_E(A_1 \cup A_2) = \mathbb{P}_E(A_1) + \mathbb{P}_E(A_2)$.

Il seguente esempio costituisce un paradigma di tale uso della formula di Bayes; seguendo infatti la logica illustrata in tale esempio, la formula di Bayes può essere applicata in molti altri problemi, sostanzialmente analoghi, suggeriti in diversi campi di applicazione della teoria delle probabilità.

Esempio 4.5. *In un lotto di pezzi (che risultano, all'apparenza, simili) vi sono elementi di tipo A, B e C, rispettivamente nelle proporzioni del 50%, 30%, 20%.*

Quelli di tipo A hanno una probabilità del 10% di guastarsi durante il loro utilizzo. Le analoghe probabilità per quelli di tipo B e C sono rispettivamente del 15% e del 18%, rispettivamente. Viene scelto un pezzo a caso dal lotto. Quale probabilità si deve attribuire al fatto che esso sia di tipo C, se si osserva un suo guasto durante l'utilizzo?

Soluzione. Poniamo

$$E \equiv \{\textit{si osserva un guasto del pezzo scelto durante l'utilizzo}\}$$

$$H_1 \equiv \{\textit{il pezzo scelto è di tipo A}\}$$
$$H_2 \equiv \{\textit{il pezzo scelto è di tipo B}\}$$
$$H_3 \equiv \{\textit{il pezzo scelto è di tipo C}\}.$$

La condizione che il pezzo sia scelto "a caso" si traduce nella assegnazione di probabilità

$$\mathbb{P}(H_1) = 0.5, \quad \mathbb{P}(H_2) = 0.3, \quad \mathbb{P}(H_3) = 0.2.$$

Gli altri dati del problema forniscono:

$$\mathbb{P}(E|H_1) = 0.10, \quad \mathbb{P}(E|H_2) = 0.15, \quad \mathbb{P}(E|H_3) = 0.18.$$

Dalla formula delle probabilità totali otteniamo

$$\mathbb{P}(E) = 0.5 \times 0.10 + 0.3 \times 0.15 + 0.2 \times 0.18 = 0.131$$

e la formula di Bayes fornisce:

$$\mathbb{P}(H_1|E) = \frac{50}{131}, \quad \mathbb{P}(H_2|E) = \frac{45}{131}, \quad \mathbb{P}(H_3|E) = \frac{36}{131}.$$

Osservazione 3. La formula di Bayes (30) può essere più brevemente scritta nella forma

$$\mathbb{P}(H_l|E) \propto \mathbb{P}(H_l) \cdot \mathbb{P}\left(E|H_l\right), \qquad l = 1, ..., m.$$

Notiamo che, essendo $\mathcal{H}$ una partizione di Ω, sappiamo già a priori che deve risultare

$$\sum_{l=1}^{m} \mathbb{P}\left(H_l|E\right) = 1.$$

Infatti considerando quanto detto nell'***Osservazione 2***,

$$\sum_{l=1}^{m} \mathbb{P}\left(H_l|E\right) = \sum_{l=1}^{m} \mathbb{P}_E\left(H_l\right) = 1.$$

La quantità $K = \dfrac{1}{\mathbb{P}(E)} = \dfrac{1}{\sum_{r=1}^{m} \mathbb{P}\left(E|H_r\right) \mathbb{P}\left(H_r\right)}$ ha dunque il ruolo di ***costante di normalizzazione***.

Esempio 4.6. ***(Un esempio medico)*** *Siamo nel 1990, agli inizi del problema dell'HIV, e siamo nella seguente situazione: c'è un test (poco costoso) che permette di verificare la sieropositività per l'HIV, ma che non è infallibile, nel senso che può accadere di risultare sieropositivi, anche senza essere stati colpiti dal virus. È noto che un individuo colpito dal virus HIV risulta sieropositivo in 999 casi su 1000, mentre un individuo "sano" risulta sieropositivo in un caso su 100. Infine è noto che, nella popolazione di New York, la percentuale dei colpiti dal virus è dello* $0,6\%$. *John, che è stato scelto a caso tra gli abitanti di New York, si sottopone al test. Calcolare la probabilità che John risulti sieropositivo e, nel caso in cui sia risultato sieropositivo, la probabilità che sia stato colpito veramente dal virus HIV. Ripetere i calcoli considerando la sieronegatività.*

Soluzione. Iniziamo esplicitando le probabilità sulla base dei dati del problema. Indicando con S^+ l'evento $\{John\, risulta\; sieropositivo\}$, con S^- l'evento complementare $\{John\; risulta\; sieronegativo\}$ e con H l'evento $\{John\, ha\, il\, virus\, HIV\}$, abbiamo

$$\mathbb{P}(H) = \frac{6}{1000}, \qquad \mathbb{P}(\overline{H}) = \frac{994}{1000},$$

$$\mathbb{P}(S^+|H) = \frac{999}{1000}, \quad \big(\text{quindi } \mathbb{P}(S^-|H) = \frac{1}{1000}\big), \quad \mathbb{P}(S^+|\overline{H}) = \frac{1}{100}, \quad \big(\text{quindi } \mathbb{P}(S^-|\overline{H}) = \frac{99}{100}\big).$$

Di conseguenza

$$\mathbb{P}(S^+) = \mathbb{P}(H)\,\mathbb{P}(S^+|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^+|\overline{H}) = \frac{6}{1000}\,\frac{999}{1000} + \frac{994}{1000}\,\frac{1}{100} = 0,015934,$$

e analogamente

$$\mathbb{P}(S^-) = \mathbb{P}(H)\,\mathbb{P}(S^-|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^-|\overline{H}) = \frac{6}{1000}\,\frac{1}{1000} + \frac{994}{1000}\,\frac{99}{100} = 0,984066\big(= 1 - \mathbb{P}(S^+)\big).$$

Dalla formula di Bayes si ha

$$\begin{aligned}\mathbb{P}(H|S^+) &= \frac{\mathbb{P}(H)\,\mathbb{P}(S^+|H)}{\mathbb{P}(H)\,\mathbb{P}(S^+|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^+|\overline{H})} = \frac{\frac{6}{1000}\,\frac{999}{1000}}{\frac{6}{1000}\,\frac{999}{1000} + \frac{994}{1000}\,\frac{1}{100}} = \frac{\frac{6}{1000}\,\frac{999}{1000}}{\frac{6}{1000}\,\frac{999}{1000} + \frac{994}{1000}\,\frac{10}{1000}}\\ &= \frac{6 \cdot 999}{6 \cdot 999 + 994 \cdot 10} \simeq 0,37617672\end{aligned}$$

e analogamente

$$\begin{aligned}\mathbb{P}(H|S^-) &= \frac{\mathbb{P}(H)\,\mathbb{P}(S^-|H)}{\mathbb{P}(H)\,\mathbb{P}(S^-|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^-|\overline{H})} = \frac{\frac{6}{1000}\,\frac{1}{1000}}{\frac{6}{1000}\,\frac{1}{1000} + \frac{994}{1000}\,\frac{99}{100}} = \frac{\frac{6}{1000}\,\frac{1}{1000}}{\frac{6}{1000}\,\frac{1}{1000} + \frac{994}{1000}\,\frac{990}{1000}}\\ &= \frac{6 \cdot 1}{6 \cdot 1 + 994 \cdot 990} \simeq 0,00000609.\end{aligned}$$

La probabilità $\mathbb{P}(H|S^+)$ ottenuta può sembrare sorprendentemente bassa: alcune osservazioni a questo proposito sono contenute nella seguente nota.

Alcune osservazioni sull'esempio medico 4.6

Come abbiamo detto, la probabilità $\mathbb{P}(H|S^+)$ ottenuta può sembrare sorprendentemente bassa, ma (come è chiarito nel seguito di questa nota) il risultato è meno strano se si pone l'attenzione su due fatti:
1) John non è stato sottoposto al test perché si era esposto al rischio di infezione, ma è stato scelto a caso.
2) Il test non è simmetrico, nel senso che dà errore con frequenza diversa a seconda del caso in cui uno sia affetto dal virus o no: infatti la probabilità di errore nel caso di individuo sano, $\mathbb{P}(S^-|\overline{H}) = \frac{1}{100}$, è dieci volte $\mathbb{P}(S^-|H) = \frac{1}{1000}$, la probabilità di errore nel caso di individuo infetto.

Per quanto riguarda il punto 1), supponiamo invece che John sia stato a contatto con una o più persone malate di HIV. La probabilità (a priori) di essere stato colpito dal virus non sarebbe stata uguale a $6/1000$, ma molto più alta. A titolo di esempio supponiamo che sia $\mathbb{P}(H) = \mathbb{P}(\overline{H}) = 1/2$, allora l'essere risultato positivo al test avrebbe comportato il seguente calcolo:

$$\begin{aligned}\mathbb{P}(H|S^+) &= \frac{\mathbb{P}(H)\,\mathbb{P}(S^+|H)}{\mathbb{P}(H)\,\mathbb{P}(S^+|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^+|\overline{H})}\\ &= \frac{\frac{1}{2}\,\frac{999}{1000}}{\frac{1}{2}\,\frac{999}{1000} + \frac{1}{2}\,\frac{10}{1000}} = \frac{999}{999+10} \simeq 0,99.\end{aligned}$$

Sembra quindi che questo test sia abbastanza buono nel caso in cui si abbiano seri motivi per sospettare un contagio (ovviamente sarà bene in ogni caso fare altre indagini).
Tuttavia, il risultato ottenuto nel caso di un individuo scelto a caso suggerisce il dubbio se un tale test sia valido per fare uno screening per una popolazione, nel senso che sembra dare troppi falsi allarmi. Questo accade perché il test non è simmetrico, nel senso del punto 2). Se la asimmetria non fosse stata così grande il test sarebbe stato migliore per uno screening. A titolo di esempio, sempre nel caso in cui John fosse scelto a caso, cioè con $\mathbb{P}(H) = \frac{6}{1000}$, $\mathbb{P}(\overline{H}) = \frac{994}{1000}$, la precisione del test nel caso di infezione fosse sempre la stessa, cioè $\mathbb{P}(S^-|H) = \frac{1}{1000}$, ma la precisione del test fosse migliore nel caso di non infezione, ad esempio se fosse $\mathbb{P}(S^+|\overline{H}) = \frac{1{,}5}{1000}$, allora si avrebbe

$$\begin{aligned}\mathbb{P}(H|S^+) &= \frac{\mathbb{P}(H)\,\mathbb{P}(S^+|H)}{\mathbb{P}(H)\,\mathbb{P}(S^+|H) + \mathbb{P}(\overline{H})\,\mathbb{P}(S^+|\overline{H})}\\ &= \frac{\frac{6}{1000}\,\frac{999}{1000}}{\frac{6}{1000}\,\frac{999}{1000} + \frac{994}{1000}\,\frac{1{,}5}{1000}} = \frac{6\cdot 999}{6\cdot 999 + 994\cdot 1,5} \simeq 0,8008,\end{aligned}$$

che è una probabilità abbastanza alta (ovviamente anche in questo caso sarebbero opportune ulteriori indagini).

4.3 Esercizi di verifica

Esercizio 4.1. A e B sono due eventi tali che

$$\mathbb{P}(A \cap B) = 0.3, \quad \mathbb{P}(A \cap \overline{B}) = 0.2, \quad \mathbb{P}(\overline{A} \cap B) = 0.1.$$

Calcolare $\mathbb{P}(A \cup B)$, $\mathbb{P}(A|B)$.

Esercizio 4.2. Vengono estratti a caso, senza reinserimento, due elementi dall'insieme $\{1, 2, ..., 9\}$. Poniamo

$$A_i \equiv \{X_i \textit{ pari}\}, \ i = 1, 2.$$

Utilizzando la formula delle probabilità composte, calcolare le probabilità degli eventi $A_1 \cap A_2$, $A_1 \cap \overline{A}_2$, $\overline{A}_1 \cap A_2$, $\overline{A}_1 \cap \overline{A}_2$.

Esercizio 4.3. Nel lancio di due dadi, qual è la probabilità condizionata che nessuno dei due punteggi sia superiore a 4 sapendo che la somma dei due punteggi è uguale a 7?

Esercizio 4.4. Abbiamo due urne: l'urna U contiene una sola pallina gialla ed r palline rosse; l'urna V contiene una sola pallina rossa ed r palline gialle. Viene scelta a caso una fra queste due urne e ne estraiamo (ancora a caso) una pallina.
a) Calcolare la probabilità dell'evento

$$E \equiv \{\textit{la pallina estratta è gialla}\}$$

b) Condizionatamente all'osservazione dell'evento E, qual è la probabilità di aver eseguito le estrazioni dall'urna U?

Esercizio 4.5. Nel gioco della roulette, qual è la probabilità condizionata del risultato *pair*, dato che si è ottenuto il risultato *passe*? (Ricordiamo che il risultato $\{0\}$, non è *passe*, ne' *manque*, ne' *pair*, ne' *unpair*).

Esercizio 4.6. Un'urna contiene 3 palline rosse e 7 bianche; si esegue un'estrazione casuale e se ne reinserisce una pallina di colore opposto a quella estratta; si procede quindi ad una successiva estrazione casuale.
a) Qual è la probabilità di una pallina rossa alla seconda estrazione?
b) Sapendo che le palline estratte nelle due successive estrazioni sono dello stesso colore, qual è la probabilità che siano entrambe bianche?

Esercizio 4.7. Sto organizzando un appuntamento per una cena fra amici per questa sera. Non riesco a raggiungere Emilio per telefono e chiedo a Aldo e a Bruno di provare ad avvertirlo. Aldo e a Bruno proveranno separatamente ad avvertirlo, Aldo inviandogli un messaggio di posta elettronica e Bruno inviando un messaggio sul telefono cellulare. Dò le seguenti valutazioni di probabilità

$$\mathbb{P}(\{\textit{Emilio leggerà la sua posta elettronica}\}) = 0.7$$

$$\mathbb{P}(\{\textit{Emilio riceverà il messaggio sul suo cellulare}\}) = 0.8$$

$$\mathbb{P}(\{\textit{Emilio leggerà la posta elettronica e riceverà il messaggio sul cellulare}\}) = 0.56.$$

a) Come devo valutare la probabilità che Emilio venga all'appuntamento?
b) Dato che Emilio effettivamente si presenta all'appuntamento, come devo valutare la probabilità che egli abbia letto la sua posta elettronica?

Esercizio 4.8. Relativamente a due eventi A, B, suppongo di aver assegnato le probabilità

$$\mathbb{P}(A), \quad \mathbb{P}(B), \quad \mathbb{P}(A|B).$$

Come devo valutare la probabilità che si sia verificato A, condizionatamente all'informazione che si è verificato almeno uno fra i due eventi A e B?

5 Correlazione e indipendenza fra eventi

5.1 Il caso di due eventi: correlazione positiva, negativa e indipendenza

Riprendiamo l'Esempio 4.5 della precedente Lezione 4, sull'estrazione da un lotto di pezzi apparentemente identici. Come ci si poteva già aspettare intuitivamente prima di svolgere i calcoli, risulta

$$\mathbb{P}(H_1|E) < \mathbb{P}(H_1), \quad \mathbb{P}(H_3|E) > \mathbb{P}(H_3).$$

Da tale osservazione prendiamo spunto per formulare le seguenti definizioni.

Definizione 5.1. *Siano A e B due eventi, con $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$. A e B si dicono* **correlati positivamente** *se risulta*

$$\mathbb{P}(A|B) > \mathbb{P}(A).$$

Notiamo che tale condizione è equivalente a

$$\mathbb{P}(A \cap B) > \mathbb{P}(A) \cdot \mathbb{P}(B)$$

e che, dunque, tale relazione è simmetrica.

Definizione 5.2. *Due eventi A e B, con $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$, si dicono* **correlati negativamente** *se risulta*

$$\mathbb{P}(A|B) < \mathbb{P}(A)$$

oppure

$$\mathbb{P}(A \cap B) < \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Definizione 5.3. *Due eventi A e B si dicono* **stocasticamente indipendenti** *se risulta*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Notiamo che non abbiamo richiesto necessariamente la condizione $\mathbb{P}(A) > 0$, $\mathbb{P}(B) > 0$. Se tale condizione è verificata e A e B sono indipendenti allora risulta

$$\mathbb{P}(A|B) = \mathbb{P}(A), \quad \mathbb{P}(B|A) = \mathbb{P}(B).$$

Esempio 5.1. *Consideriamo l'esperimento relativo al lancio di due dadi, prendendo $\Omega = \big\{(i, j),\ i,\ j \in \{1, 2, 3, 4, 5, 6\}\big\}$. Imponiamo la condizione che ciascuno dei trentasei eventi elementari possibili abbia probabilità $\mathbb{P}(\{(i, j)\}) = \frac{1}{36}$. Denotiamo con X_1 il risultato del primo dado, e con X_2 quello del secondo dado. Consideriamo gli eventi composti $E_1 \equiv \{X_1 \text{ pari}\}$, $E_2 \equiv \{X_1 + X_2 \text{ pari}\}$, $E_3 \equiv \{X_1 + X_2 \leq 4\}$, $E_4 \equiv \{X_1 \leq 2\}$, $E_5 \equiv \{\max(X_1, X_2) > 3\}$.*

È facile verificare che risulta:

E_1 ed E_2 sono stocasticamente indipendenti,

E_3 ed E_4 sono correlati positivamente,

E_3 ed E_5 sono correlati negativamente.

È importante a questo punto tener presente quanto segue.

Osservazione 1. Consideriamo ancora, a mò di esempio, l'esperimento del lancio di due dadi. Si verifica immediatamente che assegnare uguali probabilità $\frac{1}{36}$ a tutti gli eventi elementari implica che gli eventi (composti) del tipo $\{X_1 = i\}$, $\{X_2 = j\}$ sono tutti equiprobabili con probabilità uguale ad $\frac{1}{6}$ e $\{X_1 = i\}$, $\{X_2 = j\}$, per $1 \leq i \leq 6$, $1 \leq j \leq 6$, costituiscono coppie di eventi indipendenti.

È d'altra parte immediato verificare anche il viceversa, cioè che la condizione

$$\mathbb{P}(X_1 = i) = \mathbb{P}(X_2 = j) = \frac{1}{6}$$

insieme all'indipendenza stocastica per tutte le coppie $\{X_1 = i\}$, $\{X_2 = j\}$ implica che tutti gli eventi elementari $\{X_1 = i\} \cap \{X_2 = j\}$ hanno probabilità uguale ad $\frac{1}{36}$.

Questa equivalenza prefigura un fatto piuttosto generale: *spesso l'assegnazione di probabilità non avviene imponendo le probabilità degli eventi elementari ma, piuttosto, imponendo che certi eventi composti abbiano probabilità assegnate ed imponendo l'indipendenza stocastica fra opportune coppie di eventi*.

Ciò può permettere di individuare quali debbano essere le corrispondenti probabilità per tutti gli eventi semplici (e quindi, attraverso la (4), per tutti gli eventi composti) oppure può permettere di individuare quali siano le probabilità almeno per certi eventi cui siamo effettivamente interessati.

Esempio 5.2. *Nel lancio di due dadi assumiamo che*

$$\mathbb{P}(X_1 = i) = \mathbb{P}(X_2 = j) = \frac{1}{6}, \qquad 1 \leq i \leq 6, \quad 1 \leq j \leq 6$$

e che

$$\mathbb{P}(\{X_1 \in I\} \cap \{X_2 \in J\}) = \mathbb{P}(X_1 \in I) \cdot \mathbb{P}(X_2 \in J),$$

essendo $I,\ J$ *una arbitraria coppia di sottoinsiemi di* $\{1, 2, ..., 6\}$.

Calcolare la probabilità dell'evento $\{$almeno un punteggio ≥ 5 nei due lanci$\}$.

Soluzione. Basta osservare che

$$\begin{aligned}
&\mathbb{P}(\{(X_1 \geq 5\} \cup \{X_2 \geq 5)\}) \\
&= \mathbb{P}(X_1 \geq 5) + \mathbb{P}(X_2 \geq 5) - \mathbb{P}(\{X_1 \geq 5\} \cap \{X_2 \geq 5)\}) \\
&= 2\,\mathbb{P}(X_1 \geq 5) - [\mathbb{P}(X_1 \geq 5)]^2 = \frac{2 \cdot 2}{6} - \frac{1}{9} = \frac{20}{36} = \frac{5}{9}.
\end{aligned}$$

Osserviamo che abbiamo calcolato tale probabilità senza usare la formula (4) che richiede di specificare quali sono gli eventi elementari che costituiscono l'evento di interesse (e quindi senza ricorrere ad un discorso di tipo combinatorio). Abbiamo invece calcolato tale probabilità imponendo i valori delle probabilità per alcuni eventi e di indipendenza stocastica fra certe coppie di eventi; tale procedimento è più sintetico, e ciò può costituire una caratteristica importante nei casi in cui $|\Omega|$ è un numero molto grande.

Esempio 5.3. *Tizio ha comprato due biglietti di ciascuna di due diverse lotterie. Sono stati emessi 700 biglietti per ciascuna lotteria e, in ogni lotteria, vengono estratti tre biglietti vincenti. Quale probabilità ha di vincere almeno un premio?*

Soluzione. Si sottointende che, per ciascuna lotteria, le estrazioni sono casuali e senza reinserimento; si sottointende inoltre che vi sia indipendenza stocastica fra quello che succede nelle due lotterie diverse.

Dunque, ponendo $E \equiv \{$*Tizio vince almeno un premio*$\}$ ed $E_i \equiv \{$ *Tizio vince almeno un premio nella lotteria i*$\}$, con $i = 1, 2$, si ha

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E_1 \cup E_2)\\ &= \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2)\\ &= 2\,\mathbb{P}(E_1) - [\mathbb{P}(E_1)]^2 .\end{aligned}$$

Per quanto riguarda il calcolo di $\mathbb{P}(E_1)$, osserviamo che, applicando la formula delle probabilità composte, si ottiene

$$\mathbb{P}(E_1) = 1 - \mathbb{P}\left(\overline{E}_1\right) = 1 - \frac{697}{700} \cdot \frac{696}{699}.$$

Suggerimento per il calcolo della probabilità $\mathbb{P}(\overline{E}_1)$

Per calcolare la probabilità di $\overline{E}_1$ possiamo ragionare in due modi:

(I MODO) Posto A_1^i l'evento all'estrazione i-sima non viene estratto nessuno dei due biglietti si ha

$$\mathbb{P}\left(\overline{E}_1\right) = \mathbb{P}(A_1^1 \cap A_1^2 \cap A_1^3) = \mathbb{P}(A_1^1)\,\mathbb{P}(A_1^2|A_1^1)\,\mathbb{P}(A_1^3|A_1^1 \cap A_1^2) = \frac{698}{700}\frac{697}{699}\frac{696}{698} = \frac{697}{700}\cdot\frac{696}{699},$$

in quanto ci si riconduce a tre estrazioni senza reinserimento da un'urna che contiene 2 palline bianche (i due numeri dei due biglietti posseduti da Tizio, e 698 rosse, tutti gli altri: l'evento E_1 corrisponde allora all'evento nelle tre estrazioni escono solo palline rosse.

(II MODO) È interessante notare anche che a questo risultato si può arrivare anche pensando che invece si tratti di estrazioni in blocco, ossia di scegliere 2 biglietti tra i 700 di cui si sa che esattamente 3 sono vincenti, come sarebbe logico in un gratta e vinci, o una lotteria in cui si comprano biglietti su cui può essere scritta la frase *Non hai vinto ritenta* oppure *Hai vinto!*. Allora la situazione si riconduce a due estrazioni senza reinserimento da un'urna che contiene 3 palline verdi e 697 arancioni, e l'evento $\overline{E}_1$ diviene l'evento $\{$*nelle due estrazioni si estraggono solo palline arancioni*$\}$: la probabilità di non vincere diviene allora immediatamente

$$\frac{\binom{3}{0}\binom{697}{2}}{\binom{700}{2}} = \frac{697}{700} \cdot \frac{696}{699}.$$

Per finire notiamo che l'indipendenza di due eventi non è una proprietà intrisenca degli eventi ma è determinata dalla probabilità assegnata, come mostra il seguente esempio:

Esempio 5.4. *Come nell'Esempio 5.2, consideriamo l'esperimento relativo al lancio di due dadi, prendendo* $\Omega = \big\{(i,j),\ i,j \in \{1,2,3,4,5,6\}\big\}$. *Imponiamo invece la condizione che i trentasei eventi elementari possibili abbiano probabilità* $\mathbb{Q}(\{(i,j)\}) = C\,i \cdot j$. *Osserviamo che, per mettere in evidenza la differenza con la situazione dell'Esempio 5.2, abbiamo usato il simbolo* $\mathbb{Q}$ *per denotare la misura di probabilità, invece di* $\mathbb{P}$. *Per calcolare il valore di* C *basta imporre che* $\sum_{i=1}^{6}\sum_{j=1}^{6} C i \cdot j = 1$, *e* *dei semplici calcoli mostrano che* $C = \frac{1}{21^2}$. *Denotiamo con* X_1 *il risultato del primo dado, e con* X_2 *quello del secondo dado e, come nell'Esempio 5.2, consideriamo gli eventi composti* $E_1 \equiv \{X_1 \text{ pari}\}$ *ed* $E_2 \equiv \{X_1 + X_2 \text{ pari}\}$. *È facile calcolare che* $\mathbb{Q}(X_1 = i) = \frac{i}{21}$, *per* $i \in \{1, 2, 3, 4, 5, 6\}$, *e che*

$$\mathbb{Q}(E_1) = \frac{2+4+6}{21} = \frac{4}{7},$$

$$\mathbb{Q}(E_2) = \frac{1(1+3+5)+2(2+4+6)+3(1+3+5)+4(2+4+6)+5(1+3+5)+6(2+4+6)}{21^2} = \frac{25}{49},$$

$$\mathbb{Q}(E_1 \cap E_2) = \frac{2(2+4+6)+4(2+4+6)+6(2+4+6)}{21^2} = \frac{16}{49} \neq \mathbb{Q}(E_1) \cdot \mathbb{Q}(E_2),$$

e quindi, gli eventi E_1 *ed* E_2 *non sono stocasticamente indipendenti rispetto alla misura di probabilità* $\mathbb{Q}$, *mentre lo sono rispetto alla misura di probabilità* $\mathbb{P}$ *dell'Esempio 5.2.*

In quanto segue approfondiremo alcuni aspetti critici della nozione di indipendenza stocastica fra eventi. D'ora in poi verrà utilizzato il simbolo $A \perp\!\!\!\perp B$ per indicare l'indipendenza stocastica fra due eventi A e B.

5.2 Indipendenza fra due partizioni e fra due algebre di eventi

In uno spazio di probabilità, consideriamo due eventi A e B e le loro rispettive negazioni $\overline{A}$ e $\overline{B}$.

È facile verificare che le seguenti relazioni sono fra di loro equivalenti:

$$A \perp\!\!\!\perp B, \quad \overline{A} \perp\!\!\!\perp B, \quad A \perp\!\!\!\perp \overline{B}, \quad \overline{A} \perp\!\!\!\perp \overline{B},$$

ad esempio, assumiamo $A \perp\!\!\!\perp B$ e mostriamo che $\overline{A} \perp\!\!\!\perp B$; infatti possiamo scrivere

$$\begin{aligned}\mathbb{P}\left(\overline{A} \cap B\right) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B) \\ &= \mathbb{P}(B) \cdot [1 - \mathbb{P}(A)] = \mathbb{P}(B) \cdot \mathbb{P}\left(\overline{A}\right),\end{aligned}$$

e dunque $\overline{A} \perp\!\!\!\perp B$.

Possiamo riassumere quanto sopra affermando che prendendo un qualunque evento della partizione $\{A, \overline{A}\}$ ed un qualunque evento della partizione $\{B, \overline{B}\}$ otteniamo una coppia di eventi indipendenti.

Ciò suggerisce la seguente definizione.

Definizione 5.4. *Siano* $\mathcal{A} \equiv \{A_1, A_2, ..., A_n\}$ *e* $\mathcal{B} \equiv \{B_1, B_2, ..., B_m\}$ *due diverse partizioni (finite) di uno stesso spazio campione* Ω. $\mathcal{A}$ *e* $\mathcal{B}$ *sono due* **partizioni indipendenti** *se risulta*

$$A_i \perp\!\!\!\perp B_j, \quad \forall\, i = 1, ..., n;\ j = 1, 2, ..., m.$$

Esempio 5.5. *Consideriamo di nuovo l'esperimento del lancio di due dadi perfetti (o ben equilibrati) e gli eventi*

$$A_i \equiv \{X_1 = i\}, \quad i = 1, 2, ..., 6; \qquad B_i \equiv \{X_2 = j\}, \quad j = 1, 2, ..., 6.$$

La condizione che tutti gli eventi elementari siano equiprobabili implica, come è immediato verificare (confrontare anche l' ***Osservazione 1***), *che le partizioni* $\mathcal{A} \equiv \{A_1, A_2, ..., A_6\}$ *e* $\mathcal{B} \equiv \{B_1, B_2, ..., B_6\}$ *sono indipendenti.*

Esercizio proposto 5.1. *Nelle condizioni del precedente Esempio 5.5, verificare che gli eventi* $E = \{X_1 \leq 4\}$ *ed* $F = \{X_2 \text{ pari}\}$ *sono indipedenti.*

Esercizio proposto 5.2. *Lanciano due dadi e denotiamo con* X_1 *il risultato del primo dado, e con* X_2 *quello del secondo dado. Poniamo* $A_i = \{X_1 = i\}$ *e* $B_j = \{X_2 = j\}$*, per* $i,\ j \in \{1,2,3,4,5,6\}$.
Supponiamo che i dadi siano entrambi truccati in modo che $P(X_1 = i) = P(X_2 = i) = Ci$*,* $i \in \{1,2,3,4,5,6\}$ *e che le le partizioni* $\mathcal{A} = \{A_1, A_2, ..., A_n\}$ *e* $\mathcal{B} \equiv \{B_1, B_2, ..., B_m\}$ *siano partizioni indipendenti.*
(a) *Verificare che gli eventi* $E = \{X_1 \leq 4\}$ *ed* $F = \{X_2 \text{ pari}\}$ *sono indipendenti.*
(b) *Verificare che comunque presi* $I,\ J \subseteq \{1,2,3,4,5,6\}$ *gli eventi* $E = \{X_1 \in I\}$ *ed* $F = \{X_2 \in J\}$ *sono indipedenti.*

I precedenti esercizi sono un caso particolare del seguente risultato più generale:

***Proposizione* 5.1.** *Siano* $\mathcal{A} = \{A_1, A_2, ..., A_n\}$ *e* $\mathcal{B} \equiv \{B_1, B_2, ..., B_m\}$ *due partizioni (finite) indipendenti di* Ω*. Allora, comunque scelti due eventi*

$$E = \bigcup_{i \in I} A_i \quad e \quad F = \bigcup_{j \in J} B_j,$$

dove $I = \{i_1, i_2, \cdots, i_k\} \subseteq \{1, 2, \cdots, n\}$ *e* $J = \{j_1, j_2, \cdots, j_h\} \subseteq \{1, 2, \cdots, m\}$*, risulta*

$$E \perp\!\!\!\perp F.$$

Dimostrazione. Siano

$$I = \{i_1, i_2, \cdots, i_k\} \subseteq \{1, 2, \cdots, n\} \quad \text{e} \quad J = \{j_1, j_2, \cdots, j_h\} \subseteq \{1, 2, \cdots, m\},$$

posto

$$E = \bigcup_{i \in I} A_i = \bigcup_{r=1}^{k} A_{i_r}, \quad \text{ed} \quad F = \bigcup_{j \in J} B_j = \bigcup_{s=1}^{h} B_{j_s},$$

allora

$$E \cap F = \bigcup_{i \in I} \bigcup_{j \in J} A_i \cap B_j = \bigcup_{r=1}^{k} \bigcup_{s=1}^{h} A_{i_r} \cap B_{j_s}.$$

Si osservi che finora non abbiamo utilizzato ne' il fatto che $\mathcal{A}$ e $\mathcal{B}$ sono partizioni, ne' il fatto che sono partizioni indipendenti.
Utilizziamo ora il fatto che $\mathcal{A} \equiv \{A_1, A_2, ..., A_n\}$ e $\mathcal{B} \equiv \{B_1, B_2, ..., B_m\}$ sono due partizioni (finite) dell'evento certo. Allora

$$\mathbb{P}(E \cap F) = \sum_{r=1}^{k} \sum_{s=1}^{h} \mathbb{P}(A_{i_r} \cap B_{j_s}),$$

come si deduce immediatamente dalla proprietà di additività delle probabilità, e dal fatto, che essendo $\mathcal{A}$ e $\mathcal{B}$ partizioni, se $(i', j') \neq (i'', j'')$ allora gli eventi del tipo $A_{i'} \cap B_{j'}$ e $A_{i''} \cap B_{j''}$ sono disgiunti.
Si osservi che per il momento abbiamo utilizzato solo il fatto che $\mathcal{A}$ e $\mathcal{B}$ sono partizioni, ma non che sono partizioni indipendenti.

A questo punto si arriva immediatamente al risultato osservando che, se le due partizioni sono indipendenti, allora

$$\mathbb{P}(E\cap F)=\sum_{r=1}^{k}\sum_{s=1}^{h}\mathbb{P}(A_{i_r}\cap B_{j_s})=\sum_{r=1}^{k}\sum_{s=1}^{h}\mathbb{P}(A_{i_r})\,\mathbb{P}(B_{j_s}),$$

e che, dall'altra parte,

$$\mathbb{P}(E)\ \mathbb{P}(F)=\left(\sum_{r=1}^{k}\mathbb{P}(A_{i_r})\right)\left(\sum_{s=1}^{h}\mathbb{P}(B_{j_s})\right).$$

Per ottenere l'indipendenza, basta poi osservare che, dalla proprietà distributiva della somma rispetto al prodotto (vedere la nota a pagina 25) si ottiene quindi

$$\mathbb{P}(E\cap F)=\mathbb{P}(E)\ \mathbb{P}(F),$$

□

Osservazione 2. La nozione di indipendenza fra partizioni ha, nella teoria della probabilità, un importante significato concettuale, su cui ritorneremo in seguito. Per il momento ci limitiamo ad accennare che, in un certo senso, la nozione di indipendenza stocastica esprime una relazione che si addice ad una coppia di partizioni piuttosto che ad una coppia di eventi. In ogni caso vedremo presto che la nozione di indipendenza fra due partizioni ci servirà per definire in modo semplice e concettualmente efficiente la nozione di indipendenza fra due *variabili aleatorie* (Lez. 7 e 8). In tale prospettiva è utile presentare qui le seguenti nozioni.

Definizione 5.5 (Algebra)**.** *Una famiglia $\mathcal{G}$ di eventi di Ω (dunque $\mathcal{G}\subseteq\mathcal{P}(\Omega)$) è un'***algebra***, se sono verificate le seguenti proprietà*

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

ii) $E\in\mathcal{G}\Rightarrow\overline{E}\in\mathcal{G}$

iii) $E_1,E_2\in\mathcal{G}\Rightarrow E_1\cup E_2\in\mathcal{G}$

È ovvio che se $\mathcal{G}\subseteq\mathcal{P}(\Omega)$ è un'algebra allora si ha anche che $\emptyset\in\mathcal{G}$ $\big($usare le proprietà *i)* e *ii)* con $E=\Omega\big)$ e inoltre che $E_1,E_2\in\mathcal{G}\Rightarrow E_1\cap E_2\in\mathcal{G}$ $\big($usarela Legge di De Morgan (1) e le proprietà *ii)*, *iii)*$\big)$.

Esercizio proposto 5.3. *Si dimostri che $\mathcal{P}(\Omega)$ è un'algebra.*

Esercizio proposto 5.4. *Data una partizione dell'evento certo $\mathcal{A}=\{A_1,\,A_2,\,...,\,A_n\}$, si dimostri che la famiglia degli insiemi $E\in\mathcal{P}(\Omega)$ che sono le unioni di eventi della partizione è un'algebra: in formule si tratta della famiglia degli eventi del tipo*

$$E=\bigcup_{i\in I}A_i=A_{i_1}\cup A_{i_2}\cup\cdots\cup A_{i_k},\quad \textit{con } I=\{i_1,\cdots,i_k\}\subseteq\{1,2,\cdots,n\},$$

con la convenzione che se $I=\emptyset$ allora $E=\bigcup_{i\in I}A_i=\emptyset$.

Definizione 5.6 (Algebra generata da una famiglia di eventi)**.** *Sia $\mathcal{A}\equiv\{A_1,\,A_2,\cdots,\,A_n\}$ una famiglia di eventi in uno spazio campione Ω. Si definisce* **algebra generata** *da $\mathcal{A}$, la famiglia $\mathcal{G}(\mathcal{A})$ di eventi di Ω caratterizzata dalle seguenti proprietà:*

* *$\mathcal{G}(\mathcal{A})$ è un'algebra*

** *$\mathcal{A}\subseteq\mathcal{G}(\mathcal{A})$*

*** *se $\mathcal{C}\subseteq\mathcal{P}(\Omega)$ è un'algebra e $\mathcal{A}\subseteq\mathcal{C}$ allora $\mathcal{G}(\mathcal{A})\subseteq\mathcal{C}$.*

Possiamo dire cioè che $\mathcal{G}(\mathcal{A})$ è la più piccola famiglia di sottoinsiemi di Ω che abbia contemporaneamente le due proprietà di essere un'algebra e di contenere al suo interno tutti i sottoinsiemi della famiglia $\mathcal{A}$.

Esercizio proposto 5.5 (Algebra generata da una partizione)**.** *Si dimostri che se $\mathcal{A} \equiv \{A_1,\ A_2, ...,\ A_n\}$ è una partizione (finita) dell'evento certo, l'algebra degli insiemi $E \in \mathcal{P}(\Omega)$ che sono le unioni di eventi della partizione è l'algebra generata da $\mathcal{A}$, ovvero che*

$$\mathcal{G}(\mathcal{A}) = \big\{E = \bigcup_{i \in I} A_i,\ I \subseteq \{1, 2, ..., n\}\big\}$$

Vale in proposito la seguente riformulazione della precedente Proposizione 5.1

***Proposizione* 5.2.** *Siano $\mathcal{A}$ e $\mathcal{B}$ due partizioni (finite) indipendenti di Ω. Allora, comunque scelti due eventi $E \in \mathcal{G}(\mathcal{A}), F \in \mathcal{G}(\mathcal{B})$ risulta $E \perp\!\!\!\perp F$.*

Come applicazione della precedente ***Proposizione* 5.2** si suggerisce di risolvere il seguente esercizio.

Esercizio proposto 5.6. *Siano $\mathcal{A} \equiv \{A_1, A_2, ..., A_6\}$ e $\mathcal{B} \equiv \{B_1, B_2, ..., B_6\}$ le partizioni dell'Esempio 5.5, relativo al lancio di due dadi. Si verifichi che*
(a) $\mathcal{G}(\mathcal{A}) = \big\{\{X_1 \in I\},$ per $I \subseteq \{1, 2, 3, 4, 5, 6\}\big\}$;
(b) $\mathcal{G}(\mathcal{B}) = \big\{\{X_2 \in J\},$ per $J \subseteq \{1, 2, 3, 4, 5, 6\}\big\}$;
*Come nell'Esempio 5.5 si assuma la sola condizione che tutti gli eventi elementari sono equiprobabili, in modo che le due partizioni $\mathcal{A}$ e $\mathcal{B}$ sono indipendenti. Utilizzando la **Proposizione 5.2** e i precedenti punti (a) e (b) si verifichi che*
(c) qualunque siano $I,\ J \subseteq \{1, 2, 3, 4, 5, 6\}$, gli eventi $\{X_1 \in I\}$ e $\{X_2 \in J\}$ sono indipendenti.

5.3 Indipendenza completa e prove bernoulliane

Dovrebbe essere abbastanza chiaro il seguente significato intuitivo della condizione di indipendenza stocastica fra due eventi A e B: A e B sono indipendenti se il sapere con certezza che si è verificato B, o anche il sapere con certezza che non si è verificato B, non modifica le aspettative circa il verificarsi, o meno, dell'evento A.

Ovviamente si tratta di un concetto limite, di una condizione ideale, che viene assunta quale ipotesi di lavoro per ottenere delle rilevanti semplificazioni nell'analisi di un problema reale (è utile, per fare un'analogia, pensare ad esempio al concetto di *punto* in Geometria o di *punto materiale* in Meccanica: si tratta di una condizione limite, mai realizzata, ma che viene assunta ogni qualvolta sia accettabile entro una discreta approssimazione).

Come si è già detto, tale nozione è fondamentale nella costruzione di modelli probabilistici. Infatti, in pratica, nell'assegnare una misura di probabilità su uno spazio campione, si parte molto spesso dall'individuazione di famiglie di eventi a ciascuno dei quali si impone uguale probabilità e a coppie di eventi tra i quali si impone l'indipendenza stocastica, ad esempio, si assume in genere che le estrazioni del lotto su ruote diverse in una stessa settimana siano fenomeni indipendenti fra di loro, i successivi lanci di una moneta perfetta siano indipendenti fra di loro, etc.... In base a tali posizioni si può dedurre quanto vale la probabilità di alcuni eventi composti, interessanti nel problema stesso.

Vedremo comunque presto che vi sono delle situazioni naturali in cui la condizione di indipendenza è palesemente contraddetta; per ora accenniamo soltanto che ciò accade nei casi in cui una situazione di mancanza di informazione fa sì che ciascun evento osservato contiene un forte valore informativo, che si riflette sulle aspettative relative ad altri eventi connessi. Tale punto verrà sviluppato nella successiva Lezione 11 (in particolare per eventi condizionatamente indipendenti).

Veniamo ora ad aspetti tecnici della nozione di indipendenza. Dobbiamo rilevare a questo proposito che la definizione precedentemente formulata, si rivela non adeguata ad esprimere compiutamente una condizione di indipendenza reciproca fra molti eventi diversi.

Ciò è efficacemente illustrato dal seguente semplice esempio.

Esempio 5.6. *Riprendiamo ancora una volta il caso del lancio di due dadi e consideriamo gli eventi:* $A \equiv \{X_1 \text{ pari}\}$*,* $B \equiv \{X_2 \text{ pari}\}$*,* $C \equiv \{X_1 + X_2 \text{ dispari}\}$*. Imponendo la condizione di equiprobabilità fra gli eventi elementari abbiamo le relazioni:* $A \perp\!\!\!\perp B$*,* $A \perp\!\!\!\perp C$*,* $B \perp\!\!\!\perp C$*. Notiamo però che ovviamente risulta*

$$\mathbb{P}(C|A \cap B) = 0.$$

Tale conclusione contrasta, naturalmente, con il significato di indipendenza e mostra l'esigenza di una definizione appropriata per il caso di più di due eventi.

Si dà allora le seguente definizione. Sia $\{E_1, E_2, ..., E_n\}$ una famiglia di eventi in uno stesso spazio di probabilità e consideriamo le partizioni $\mathcal{P}_1 \equiv \{E_1, \overline{E}_1\}, ..., \mathcal{P}_n \equiv \{E_n, \overline{E}_n\}$.

Definizione 5.7. *Gli eventi* $E_1, E_2, ..., E_n$ *sono una famiglia di eventi* **completamente (o globalmente) indipendenti** *se comunque presi degli indici* $\{j_1, ..., j_m\}$ *dall'insieme* $\{1, 2, ..., n\}$ *(*$2 \leq m \leq n$*) e comunque scelti degli eventi* $A_i \in \mathcal{P}_{j_i}$ *(dunque* $A_i = E_{j_i}$ *oppure* $A_i = \overline{E}_{j_i}$*) risulta*

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_m) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_m). \tag{31}$$

Chiaramente la precedente definizione implica le seguenti due condizioni:
considerando solo il caso $A_i = E_{j_i}$ (e non il caso $A_i = \overline{E}_{j_i}$)

$$\mathbb{P}(E_{j_1} \cap E_{j_2} \cap ... \cap E_{j_m}) = \mathbb{P}(E_{j_1}) \cdot \mathbb{P}(E_{j_2}) \cdot ... \cdot \mathbb{P}(E_{j_m}) \tag{32}$$
$$\text{per ogni } \{j_1, ..., j_m\} \subseteq \{1, 2, ..., n\}, \text{ con } 2 \leq m \leq n.$$

oppure, considerando solo il caso $m = n$,

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_n), \qquad \text{con } A_i = E_i \text{ oppure } A_i = \overline{E}_i \tag{33}$$

L'INDIPENDENZA COMPLETA (o GLOBALE) NEL CASO DI $n = 3$ EVENTI

Dati tre eventi A, B, C, la condizione (32), equivale alle seguenti 4 condizioni:

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C),$$

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B),\; \mathbb{P}(A \cap C) = \mathbb{P}(A)\,\mathbb{P}(C),\; \mathbb{P}(B \cap C) = \mathbb{P}(B)\,\mathbb{P}(C).$$

Invece la condizione (33) equivale alle seguenti 8 condizioni:

$$\begin{aligned}
&\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C), \quad &&\mathbb{P}(A \cap B \cap \overline{C}) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(\overline{C}),\\
&\mathbb{P}(A \cap \overline{B} \cap C) = \mathbb{P}(A)\,\mathbb{P}(\overline{B})\,\mathbb{P}(C), \quad &&\mathbb{P}(A \cap \overline{B} \cap \overline{C}) = \mathbb{P}(A)\,\mathbb{P}(\overline{B})\,\mathbb{P}(\overline{C}),\\
&\mathbb{P}(\overline{A} \cap B \cap C) = \mathbb{P}(\overline{A})\,\mathbb{P}(B)\,\mathbb{P}(C), \quad &&\mathbb{P}(\overline{A} \cap B \cap \overline{C}) = \mathbb{P}(\overline{A})\,\mathbb{P}(B)\,\mathbb{P}(\overline{C}),\\
&\mathbb{P}(\overline{A} \cap \overline{B} \cap C) = \mathbb{P}(\overline{A})\,\mathbb{P}(\overline{B})\,\mathbb{P}(C), \quad &&\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) = \mathbb{P}(\overline{A})\,\mathbb{P}(\overline{B})\,\mathbb{P}(\overline{C});
\end{aligned}$$

infine la condizione (31) equivale a chiedere, oltre alle precedenti 8 condizioni, anche le seguenti 12 condizioni

$$\begin{aligned}
&\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B), \quad &&\mathbb{P}(A \cap C) = \mathbb{P}(A)\,\mathbb{P}(C), \quad &&\mathbb{P}(B \cap C) = \mathbb{P}(B)\,\mathbb{P}(C),\\
&\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A)\,\mathbb{P}(\overline{B}), \quad &&\mathbb{P}(A \cap \overline{C}) = \mathbb{P}(A)\,\mathbb{P}(\overline{C}), \quad &&\mathbb{P}(B \cap \overline{C}) = \mathbb{P}(B)\,\mathbb{P}(\overline{C}),\\
&\mathbb{P}(\overline{A} \cap B) = \mathbb{P}(\overline{A})\,\mathbb{P}(B), \quad &&\mathbb{P}(\overline{A} \cap C) = \mathbb{P}(\overline{A})\,\mathbb{P}(C), \quad &&\mathbb{P}(\overline{B} \cap C) = \mathbb{P}(\overline{B})\,\mathbb{P}(C),\\
&\mathbb{P}(\overline{A} \cap \overline{B}) = \mathbb{P}(\overline{A})\,\mathbb{P}(\overline{B}), \quad &&\mathbb{P}(\overline{A} \cap \overline{C}) = \mathbb{P}(\overline{A})\,\mathbb{P}(\overline{C}), \quad &&\mathbb{P}(\overline{B} \cap \overline{C}) = \mathbb{P}(\overline{B})\,\mathbb{P}(\overline{C}),
\end{aligned}$$

Osservazione 3 È importante notare che in alcuni testi la definizione di famiglia di eventi completamente indipendenti è data attraverso la relazione (32), mentre in altri è data attraverso la (33). Ciò è dovuto al fatto che le relazioni (32) e (33) sono equivalenti, ed entrambe implicano la (31) (risultando quindi equivalenti alla (31)). Ciò è ovvio per $n = 2$, come detto all'inizio del paragrafo 5.2. Non diamo qui la dimostrazione di questa proprietà, cioè l'equivalenza tra le tre proprietà (32), (33) e (31), ma proponiamo al lettore il seguente esercizio.

Esercizio proposto 5.7. *Nel caso di $n = 3$ eventi A, B e C dimostrare l'equivalenza tra (32), (33) e (31).*
Suggerimento: *la soluzione è basata sull'osservazione che per tre eventi A, B, C si ha*

$$A \cap B = (A \cap B \cap C) \cup (A \cap B \cap \overline{C}),$$

e che quindi

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap \overline{C}),$$

da cui anche

$$\mathbb{P}(A \cap B \cap \overline{C}) = \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B \cap C).$$

Un caso assai particolare ma di notevole interesse è quello individuato dalla seguente definizione di ***schema di Bernoulli*** o delle ***prove di Bernoulli*** (detto anche delle ***prove ripetute***).

Definizione 5.8 (Schema di Bernoulli, o prove bernoulliane)**.** *Gli eventi $E_1, E_2, ..., E_n$ costituiscono delle* **prove bernoulliane** *se sono completamente indipendenti ed hanno tutti una stessa probabilità θ, con $0 < \theta < 1$.*

Se $E_1, E_2, ..., E_n$ costituiscono delle ***prove bernoulliane*** si ha dunque, in particolare, per $m \leq n$ e per qualunque $\{j_1, ..., j_m\} \subseteq \{1, 2, ..., n\}$,

$$\mathbb{P}(E_{j_1} \cap E_{j_2} \cap ... \cap E_{j_m}) = \theta^m.$$

Inoltre, se $\{i_1, i_2, ..., i_k\} \subseteq \overline{J} = \{1, 2, ..., n\} \backslash \{j_1, ..., j_m\}$ allora

$$\mathbb{P}(E_{j_1} \cap E_{j_2} \cap ... \cap E_{j_m} \cap \overline{E}_{i_1} \cap \overline{E}_{i_2} ... \cap \overline{E}_{i_k}) = \theta^m\,(1-\theta)^k.$$

5.4 Indipendenza completa di partizioni

Abbiamo già osservato che la nozione di indipendenza completa tra n eventi E_i, $i = 1, \dots, n$, si può esprimere attraverso una proprietà che coinvolge gli elementi delle partizioni $\mathcal{P}_i$, $i = 1, \dots, n$, generate da tali eventi, ossia $\mathcal{P}_i = \{E_i, \overline{E}_i\}$.

Date n partizioni $\mathcal{P}_i = \{H_1^i, H_2^i, \dots, H_n^i\}$, $i = 1, \dots, n$, si può generalizzare immediatamente la definizione di indipendenza completa:

Definizione 5.9. *Le partizioni $\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_n$ sono una famiglia di partizioni* **completamente (o globalmente)** indipendenti *se comunque estratti degli indici $\{j_1, ..., j_m\}$ dall'insieme $\{1, 2, ..., n\}$ (con la condizione che $2 \leq m \leq n$) e comunque scelti degli eventi $A_i \in \mathcal{P}_{j_i}$, per $i = 1, \dots m$, risulta*

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_m) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_m). \tag{34}$$

Chiaramente la precedente definizione implica la seguente condizione:
considerando solo il caso $m = n$, comunque scelti $A_i \in \mathcal{P}_i$, per $i = 1, \dots n$,

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_n). \tag{35}$$

Si può dimostrare che la condizione (35), per $A_i \in \mathcal{P}_i$, implica la condizione (34), ossia che le due precedenti condizioni sono equivalenti. Si può anche dimostrare che tali condizioni implicano la condizione (35), ma con $A_i \in \mathcal{G}(\mathcal{P}_i)$, l'algebra generata dalla partizione $\mathcal{P}_i$. Più precisamente vale la seguente generalizzazione della Proposizione 5.2.

***Proposizione* 5.3.** *Siano $\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_n$ una famiglia di partizioni completamente (o globalmente) indipendenti Allora, comunque scelti n eventi $E_i \in \mathcal{G}(\mathcal{P}_i)$, per $i = 1, \dots, n$, risulta*

$$\mathbb{P}(E_1 \cap E_2 \cap ... \cap E_n) = \mathbb{P}(E_1) \cdot \mathbb{P}(E_2) \cdot ... \cdot \mathbb{P}(E_n). \tag{36}$$

Pur non dandone una dimostrazione facciamo notare che dalla (36) si può ottenere la seguente relazione: comunque scelto m (con $2 \leq m \leq n$), comunque estratti degli indici $\{j_1, ..., j_m\}$ dall'insieme $\{1, 2, ..., n\}$ e comunque scelti degli eventi $E_{j_i} \in \mathcal{G}\big(\mathcal{P}_{j_i}\big)$, per $i = 1, \dots m$, risulta

$$\mathbb{P}(E_{j_1} \cap E_{j_2} \cap ... \cap E_{j_m}) = \mathbb{P}(E_{j_1}) \cdot \mathbb{P}(E_{j_2}) \cdot ... \cdot \mathbb{P}(E_{j_m}).$$

Basterà infatti considerare $E_\ell = \Omega$ per $\ell \notin \{j_1, \dots, j_m\}$ e applicare la (36).

Esempio 5.7. *Si lanciano tre dadi e sia X_i il risultato dell'i-esimo lancio. Si considerino le partizioni*

$$\mathcal{P}_i = \big\{\{X_i = 1\}, \{X_i = 2\}, \{X_i = 3\}, \{X_i = 4\}, \{X_i = 5\}, \{X_i = 6\}\big\}, \qquad i = 1, 2, 3.$$

Chiaramente, comunque scelti j_1, j_2*, e* j_3 *si ha*

$$\mathbb{P}(\{X_1 = j_1, X_2 = j_2, X_3 = j_3\}) = \frac{1}{216} = \mathbb{P}(\{X_1 = j_1\})\, \mathbb{P}(\{X_2 = j_2\})\, \mathbb{P}(\{X_3 = j_3\})$$

e quindi, osservando che ciascuno degli eventi $\{X_1 = j_1, X_2 = j_2, X_3 = j_3\}$ *equivale all'evento* $\{X_1 = j_i\} \cap \{X_2 = j_2\} \cap \{X_3 = j_3\}$*, e tenuto conto del fatto che la relazione* (35) *è verificata, e che è equivalente alla* (32)*, le tre partizioni sono indipendenti. Inoltre, grazie alla Proposizione 5.3, risulta anche che, comunque scelti tre sottoinsiemi* I_1*,* I_2 *ed* I_3 *di* $\{1, 2, 3, 4, 5, 6\}$*, si ha*

$$\mathbb{P}\big(\{X_1 \in I_1\} \cap \{X_2 \in I_2\} \cap \{X_3 \in I_3\}\big) = \mathbb{P}\big(\{X_1 \in I_1\}\big)\, \mathbb{P}\big(\{X_2 \in I_2\}\big)\, \mathbb{P}\big(\{X_3 \in I_3\}\big).$$

5.5 Esercizi di verifica

Esercizio 5.1. A e B sono due eventi tali che

$$\mathbb{P}(A \cap B) = 0.3, \quad \mathbb{P}(A \cap \overline{B}) = 0.2, \quad \mathbb{P}(\overline{A} \cap B) = 0.1.$$

Verificare se A e B sono stocasticamente indipendenti.

Esercizio 5.2. Mostrare che se A e B sono due eventi indipendenti e $A \subseteq B$, allora si ha $\mathbb{P}(A) = 0$, oppure $\mathbb{P}(B) = 1$.

Esercizio 5.3. Siano A e B due eventi fra loro incompatibili. Mostrare che A e B risultano stocasticamente indipendenti se e solo se almeno uno di essi ha probabilità nulla.

Esercizio 5.4. X ed Y indicano i punteggi ottenuti nel lancio di due dadi a sei facce. Poniamo

$$A \equiv \{\max(X, Y) < 5\}, \quad B \equiv \{\min(X, Y) > 3\}$$

a) Calcolare $\mathbb{P}(A \cap B)$, $\mathbb{P}(A \cup B)$, $\mathbb{P}(A|B)$, $\mathbb{P}(B|A)$.
b) Gli eventi A e B sono indipendenti? A e B sono incompatibili?

Esercizio 5.5. Consideriamo i due risultati *pair* e *passe* nel gioco della roulette. Sono stocasticamente indipendenti, correlati positivamente o correlati negativamente?

Esercizio 5.6. Indichiamo con X un numero selezionato a caso nell'insieme dei primi 120 numeri naturale e consideriamo gli eventi

$$E \equiv \{X \textit{ pari}\}, \quad F \equiv \{X \textit{ divisibile per 3}\}.$$

Fra E ed F sussiste correlazione positiva, negativa o indipendenza stocastica?

Esercizio 5.7. a) Qual è la probabilità che il numero 16 venga estratto su una data ruota del lotto in una fissata giornata?
b) Qual è la probabilità che il numero 16 non venga mai estratto su una data ruota del lotto per n giornate consecutive?
c) Qual è la probabilità condizionata che il numero 16 venga estratto l'$(n+1)$-esima giornata, dato che non è mai stato estratto nelle n giornate precedenti?

Esercizio 5.8. Non riesco ad avvertire Emilio per telefono dell'appuntamento per la cena di questa sera. Come al solito Aldo gli invierà allora un messaggio di posta elettronica, Bruno gli invierà un messaggio sul telefono cellulare, e interverrà anche Carla, cercando di avvertire di persona la sorella di Emilio. Essi avranno successo rispettivamente con probabilità $\mathbb{P}(A) = 0.8$, $\mathbb{P}(B) = 0.7$ e $\mathbb{P}(C) = 0.6$; i tre eventi, inoltre, sono completamente indipendenti.
a) Trovare la probabilità che Emilio venga informato dell'appuntamento.
b) Dato che Emilio si presenta effettivamente all'appuntamento, come devo valutare la probabilità che egli abbia letto la sua posta elettronica?

Esercizio 5.9. Una moneta viene lanciata n volte. Supponiamo che valga

$$\mathbb{P}(\underbrace{T, T, ..., T}_{n \text{ lanci}}) = \mathbb{P}(\underbrace{C, C, ..., C}_{n \text{ lanci}}) = \frac{1}{2}.$$

Posto, per $i = 1, \cdots, n$, $E_i = \{$ *esce testa all'i-simo lancio* $\}$,
a) i due eventi E_1 ed E_2 sono indipendenti, correlati positivamente o correlati negativamente?
b) i due eventi E_1 ed $\overline{E}_2$ sono indipendenti, correlati positivamente o correlati negativamente?
c) al variare di n, qual è la probabilità che vi sia un numero pari di risultati testa sugli n lanci? (considerando 0 come un numero pari)

Esercizio 5.10. Siano E_1 ed E_2 due eventi in uno spazio campione Ω. Elencate gli eventi appartenenti all'algebra generata da $\{E_1, E_2\}$.

6 Probabilità binomiali e ipergeometriche; estrazioni casuali da urne

In questo paragrafo vogliamo discutere in modo più sistematico due particolari modelli probabilistici, che sono già comparsi in precedenti esempi e che portano alle probabilità binomiali e ipergeometriche.

6.1 Probabilità binomiali

Consideriamo, su uno spazio finito di probabilità, n prove bernoulliane $E_1, ..., E_n$; cioè, ricordando la Definizione 5.8, assumiamo che $E_1, ..., E_n$ sono completamente indipendenti ed equiprobabili: ponendo $P(E_i) = \theta$, per $i = 1, 2..., n$ e ponendo

$$X_i = \begin{cases} 1 & \text{se si verifica } E_i \\ 0 & \text{se si verifica } \overline{E}_i \end{cases} \qquad i = 1, 2, \cdots, n, \tag{37}$$

si ha, per ogni n-upla $\mathbf{x} \equiv (x_1, ..., x_n) \in \{0, 1\}^n$,

$$\mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\}) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}, \tag{38}$$

dove $\{X_1 = x_1, ..., X_n = x_n\}$ è un modo rapido di scrivere l'evento $\{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\}$.

Poniamo ora

$$S_n \equiv \sum_{i=1}^n X_i$$

e consideriamo, per $k = 0, 1, ..., n$, la probabilità dell'evento composto $\{S_n = k\}$; osserviamo che potremo scrivere

$$\{S_n = k\} = \left\{ \sum_{i=1}^n X_i = k \right\} = \bigcup_{\mathbf{x} \in \{0,1\}^n : \sum_{i=1}^n x_i = k} \{X_1 = x_1, ..., X_n = x_n\}.$$

Due eventi del tipo

$$\{X_1 = x'_1, ..., X_n = x'_n\}, \quad \{X_1 = x''_1, ..., X_n = x''_n\}$$

sono ovviamente incompatibili nel caso $\mathbf{x}' \equiv (x'_1, ..., x'_n) \neq \mathbf{x}'' \equiv (x''_1, ..., x''_n)$ e, nel caso in cui

$$\sum_{i=1}^n x'_i = \sum_{i=1}^n x''_i = k,$$

essi risultano equiprobabili, entrambi di probabilità $\theta^k(1-\theta)^{n-k}$, in virtù dell'equazione (38).

E dunque, dal momento che la cardinalità dell'insieme

$$\{\mathbf{x} \in \{0, 1\}^n : \sum_{i=1}^n x_i = k\} \tag{39}$$

è uguale a $\binom{n}{k}$, potremo scrivere, per $k = 0, 1, ..., n$,

$$\begin{aligned} \mathbb{P}(\{S_n = k\}) &= \sum_{\mathbf{x} \in \{0,1\}^n : \sum_{i=1}^n x_i = k} \mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\}) \\ &= \binom{n}{k} \theta^k (1-\theta)^{n-k}. \end{aligned} \tag{40}$$

Probabilità del tipo in (40) sull'insieme $\{0, 1, \ldots, n\}$, prendono il nome di ***probabilità binomiali***.

Esempio 6.1. *Un dado viene lanciato* 10 *volte. Sia* S *il numero di volte in cui si ottiene il risultato asso. Calcolare* $\mathbb{P}(\{S=5\})$.

Soluzione. Si tratta di 10 prove bernoulliane, di probabilità $\frac{1}{6}$; dunque

$$\mathbb{P}(\{S=5\}) = \binom{10}{5} \cdot \left(\frac{5}{6}\right)^5 \cdot \left(\frac{1}{6}\right)^5 = \binom{10}{5} \cdot \frac{5^5}{6^{10}}.$$

Esempio 6.2. *Ciascun viaggiatore che occupa un posto in uno scompartimento "per fumatori" in un treno EuroStar è effettivamente un fumatore (o fumatrice) con probabilità uguale al* 70%. *Se lo scompartimento contiene* 5 *posti (oltre a quello da me prenotato), qual è la probabilità che io vi incontri meno di tre fumatori?*

Soluzione. Si sta sottointendendo che lo scompartimento venga riempito e che i viaggiatori si comportino, rispetto al fumo, in modo ciascuno indipendente dall'altro. Se S indica il numero dei fumatori nei 5 posti rimanenti, si avrà:

$$\begin{aligned}\mathbb{P}(\{S<3\}) &= \mathbb{P}(\{S=0\}) + \mathbb{P}(\{S=1\}) + \mathbb{P}(\{S=2\}) \\ &= \binom{5}{0}\left(\frac{7}{10}\right)^0\left(\frac{3}{10}\right)^5 + \binom{5}{1}\left(\frac{7}{10}\right)^1\left(\frac{3}{10}\right)^4 + \binom{5}{2}\left(\frac{7}{10}\right)^2\left(\frac{3}{10}\right)^3.\end{aligned}$$

Esempio 6.3. *Un testo, contenente* 20 *errori di stampa, viene sottoposto a due diversi correttori di bozze. Ciascun errore contenuto nel testo viene individuato da ciascun correttore con probabilità* $p = 0.6$ *ed indipendentemente da quello che accade per gli altri errori. Inoltre i due correttori lavorano indipendentemente uno dall'altro.*

Trovare la probabilità che il numero degli errori individuati da almeno uno dei due correttori sia superiore a 15.

Soluzione. Ciascun errore viene individuato (da almeno uno dei due correttori) con probabilità[11]

$$\theta = 2p - p^2 = \frac{84}{100}$$

(non ci interessa se viene individuato da uno dei due correttori o dall'altro o, eventualmente, da entrambi; a noi interessa che almeno uno dei due individui l'errore).

Si tratta quindi di 20 prove bernuolliane, in ognuna delle quali vi è una probabilità di successo uguale a θ. Indicando dunque con S il numero complessivo degli errori individuati, si avrà

$$\mathbb{P}(\{S>15\}) = \sum_{k=16}^{20} \binom{20}{k} \left(\frac{84}{100}\right)^k \left(\frac{16}{100}\right)^{20-k}.$$

6.2 Estrazioni casuali da urne con reiserimento

Illustriamo ora una tipica situazione in cui si incontra uno schema di prove bernoulliane.

Pensiamo ad una popolazione composta di M oggetti, diciamo $c_1, ..., c_M$, di cui m_1 di tipo A e i rimanenti $m_2 = M - m_1$ di un diverso tipo B. Supponiamo, ad esempio, di aver numerato $c_1, ..., c_M$ in modo tale che

$$c_1, ..., c_{m_1} \text{ sono di tipo A} \quad \text{e} \quad c_{m_1+1}, ..., c_M \text{ di tipo B.} \tag{41}$$

[11] Il ragionamento per arrivare al calcolo di $\theta = 2p - p^2$ è simile a quello usato negli Esempi 5.2 e 5.3

Eseguiamo ora n estrazioni *casuali con reinserimento* da tale popolazione; con tale termine si vuole esprimere il fatto che si eseguono n estrazione successive, reinserendo ogni volta nella popolazione l'oggetto estratto ed estraendo in modo tale che **ciascun** oggetto abbia, ogni volta, la **stessa probabilità** $\frac{1}{M}$ di essere estratto, sia esso di tipo A o di tipo B.

Lo spazio campione in tale esperimento (consistente nell'eseguire le n estrazioni) può essere identificato come l'insieme

$$\Omega \equiv \{1, ..., M\}^n$$

costituito delle n-uple ordinate di elementi in $\{1, ..., M\}$; esso ha dunque cardinalità M^n. In tale spazio campione, consideriamo, per $i = 1, ..., n$, ora gli eventi del tipo:

$$E_i \equiv \{\textit{l'oggetto estratto nella i-esima estrazione è di tipo } A\}.$$

Proposizione 1. Gli eventi $E_1, ..., E_n$ (definiti qui sopra) costituiscono delle prove bernoulliane con $\mathbb{P}(E_i) = \frac{m_1}{M}$.

Dimostrazione. La condizione che le estrazioni sono casuali con reinserimento corrisponde all'assegnazione della stessa probabilità $\frac{1}{M^n}$ a ciascuno degli eventi elementari $(j_1, ..., j_n) \in \Omega$. Inoltre, in virtù della posizione (41) possiamo scrivere gli eventi E_i come segue

$$\begin{aligned} E_i &\equiv \{(j_1, ..., j_n) \in \Omega : \ 1 \leq j_i \leq m_1\} \\ &= \overbrace{\{1, ..., M\} \times \cdots \times \{1, ..., M\}}^{i-1\, volte} \times \{1, 2, \cdots, m_1\} \times \overbrace{\{1, ..., M\} \times \cdots \times \{1, ..., M\}}^{n-i\, volte}, \end{aligned}$$

ed in modo analogo

$$\begin{aligned} \overline{E}_i &\equiv \{(j_1, ..., j_n) \in \Omega : \ m_1 + 1 \leq j_i \leq M\} \\ &= \overbrace{\{1, ..., M\} \times \cdots \times \{1, ..., M\}}^{i-1\, volte} \times \{m_1 + 1, m_1 + 2, \cdots, m_1 + m_2\} \times \overbrace{\{1, ..., M\} \times \cdots \times \{1, ..., M\}}^{n-i\, volte}. \end{aligned}$$

Dunque, considerando per $1 \leq i \leq n$, la quantità X_i come definita nella (37), si ha

$$\begin{aligned} \mathbb{P}(\{X_i = 1\}) = P(E_i) &= \frac{|\{(j_1, ..., j_n) \in \Omega : \ 1 \leq j_i \leq m_1\}|}{M^n} \\ &= \frac{m_1 \cdot M^{n-1}}{M^n} = \frac{m_1}{M}, \end{aligned}$$

e, analogamente

$$\mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\}) = \frac{m_1^{\sum_{i=1}^n x_i} m_2^{n - \sum_{i=1}^n x_i}}{M^n} = \left(\frac{m_1}{M}\right)^{\sum_{i=1}^n x_i} \left(\frac{m_2}{M}\right)^{n - \sum_{i=1}^n x_i}.$$

Mettendo insieme le due relazioni dimostrate si ottiene che, per tutti gli eventi

$$A_i \in \mathcal{P}_i = \{E_i, \overline{E}_i\} = \Big\{\{X_i = 1\}, \ \{X_i = 0\}\Big\}$$

si ha

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_n),$$

che corrisponde alla relazione (33), che come osservato in Osservazione 3 della Lez. 5, è equivalente alla relazione (31) che definisce l'indipendenza completa. □

Una precisazione sulla dimostrazione della precedente Proposizione 6.2

Al lettore più attento la dimostrazione potrebbe sembrare non soddisfacente, in quanto abbiamo dimostrato la condizione (33), senza aver dato la dimostrazione che essa è equivalente alla (31), che abbiamo preso come definizione dell'indipendenza completa. Tuttavia va osservato che, in modo assolutamente analogo, si potrebbe dimostrare direttamente la (31), ossia verificare che per ogni $2 \leq m \leq n$,

$$\begin{aligned}\mathbb{P}(\{X_{j_1} = x_{j_1}, ..., X_{j_m} = x_{j_m}\}) &= \frac{m_1^{\sum_{i=1}^m x_{j_i}} m_2^{m-\sum_{i=1}^m x_{j_i}} M^{n-m}}{M^n} \\ &= \left(\frac{m_1}{M}\right)^{\sum_{i=1}^m x_{j_i}} \left(\frac{m_2}{M}\right)^{m-\sum_{i=1}^m x_{j_i}} \\ &= \prod_{i=1}^m \mathbb{P}(\{X_{j_i} = x_{j_i}\}).\end{aligned}$$

Consideriamo ora

$$S_n = \sum_{i=1}^n X_i.$$

S_n rappresenta dunque il numero di elementi di tipo A in un campionamento casuale *con reinserimento* di n oggetti da una popolazione complessivamente costituita da M elementi, di cui m_1 di tipo A e $m_2 = M - m_1$ di tipo B. Ricordando la formula (40) ed in virtù della ***Proposizione 1***, possiamo concludere scrivendo

$$\mathbb{P}(\{S_n = k\}) = \binom{n}{k} \left(\frac{m_1}{M}\right)^k \left(\frac{m_2}{M}\right)^{n-k}, \qquad \text{per } k = 0, 1, \cdots, n. \tag{42}$$

6.3 Estrazioni casuali da urne senza reiserimento e Probabilità ipergeometriche

Consideriamo ora la stessa situazione come descritta nel paragrafo precedente, ma con la differenza che le n estrazioni siano eseguite ***senza reinserimento***. Poniamo di nuovo

$$E_i \equiv \{\textit{oggetto di tipo A alla i-esima estrazione}\}, \qquad i = 1, 2, ..., n$$

$$X_i = \begin{cases} 1 & \text{se si verifica } E_i, \\ 0 & \text{se si verifica } \overline{E}_i, \end{cases} \qquad i = 1, 2, ..., n.$$

$$S_n = \sum_{i=1}^n X_i.$$

Nel caso di estrazioni senza reinserimento potremo considerare come spazio campione lo spazio costituito dalle n-uple di elementi di $\{1, ..., M\}$, senza ripetizione:

$$\Omega \equiv \{(j_1, ..., j_n) : \ 1 \leq j_i \leq M, \ j_1 \neq ... \neq j_n\} = \{(j_1, ..., j_n) : \ 1 \leq j_i \leq M, \ j_1, \cdots, j_n \text{ tutti distinti}\}.$$

Dunque

$$|\Omega| = M\big(M-1\big)...\big(M-(n-1)\big) = M(M-1)...(M-n+1) = \frac{M!}{(M-n)!}.$$

Lo schema di estrazioni casuali senza reinserimento si traduce nella condizione che tutti gli elementi di Ω (eventi elementari) hanno uguale probabilità

$$\frac{1}{M(M-1)...(M-n+1)} = \frac{(M-n)!}{M!}.$$

Un qualunque evento composto, della forma $\{X_1 = x_1, ..., X_n = x_n\}$, ha una cardinalità che dipende soltanto dal numero $k = \sum_{i=1}^{n} x_i$ ed, esattamente, è data da

$$\begin{aligned}
&\overbrace{m_1 \cdot (m_1-1) \cdot ... \cdot \big(m_1 - (k-1)\big)}^{k\ fattori} \cdot \overbrace{m_2 \cdot (m_2-1) \cdot ... \cdot \big(m_2 - (n-k-1)\big)}^{n-k\ fattori} \\
=& m_1 \cdot (m_1-1) \cdot ... \cdot (m_1-k+1) \cdot m_2 \cdot (m_2-1) \cdot ... \cdot (m_2-(n-k)+1) \\
=& \frac{m_1!}{(m_1-k)!} \cdot \frac{m_2!}{\big(m_2-(n-k)\big)!}
\end{aligned}$$

e dunque

$$\begin{aligned}
&\mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\}) \\
&= \frac{m_1 \cdot (m_1-1) \cdot ... \cdot (m_1-k+1) \cdot m_2 \cdot (m_2-1) \cdot ... \cdot (m_2-(n-k)+1)}{M(M-1)...(M-n+1)} \\
&= \frac{m_1!}{(m_1-k)!} \frac{m_2!}{(m_2-(n-k))!} \frac{(M-n)!}{M!}
\end{aligned}$$

da cui

$$\begin{aligned}
\mathbb{P}(\{S_n = k\}) &= \sum_{\mathbf{x} \in \{0,1\}^n; \sum_i x_i = k} \mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\}) \\
&= \binom{n}{k} \frac{m_1 \cdot (m_1-1) \cdot ... \cdot (m_1-k+1) \cdot m_2 \cdot (m_2-1) \cdot ... \cdot (m_2-(n-k)+1)}{M(M-1)...(M-n+1)} \\
&= \binom{n}{k} \frac{m_1!}{(m_1-k)!} \frac{m_2!}{(m_2-(n-k))!} \frac{(M-n)!}{M!};
\end{aligned}$$

riscrivendo in forma più compatta tale ultima frazione, attraverso la notazione dei coefficienti binomiali, possiamo concludere: per $k \in \{0, 1, \dots, n\}$

$$\begin{aligned}
&\mathbb{P}(\{S_n = k\}) = \frac{\binom{m_1}{k}\binom{m_2}{n-k}}{\binom{M}{n}} \qquad \text{per } \max(0, n+m_1-M) \le k \le \min(n, m_1), \\
&\mathbb{P}(\{S_n = k\}) = 0 \qquad \text{altrimenti}
\end{aligned} \tag{43}$$

Probabilità del tipo (43) prendono il nome di ***probabilità ipergeometriche***. In effetti già conosciamo questo risultato dall'Esempio 3.8 (si veda anche l'Esercizio proposto 3.2).

Dal momento che, per fissati M, m_1, n, la famiglia degli eventi $\{S_n = k\}$, per $\max(0, n+m_1-M) \le k \le \min(n, m_1)$, costituisce una partizione dello spazio campione, deve ovviamente risultare

$$\sum_{k=0 \vee (n+m_1-M)}^{n \wedge m_1} \mathbb{P}(\{S_n = k\}) = 1$$

cioè

$$\sum_{k=0 \vee (n+m_1-M)}^{n \wedge m_1} \frac{\binom{m_1}{k}\binom{m_2}{n-k}}{\binom{M}{n}} = 1.$$

In effetti quest'ultima identità coincide con la (21).

Osservazione 1. Si può pervenire al risultato (43) anche in un modo alternativo, applicando direttamente la formula delle probabilità composte agli eventi del tipo $\{X_1 = x_1, ..., X_n = x_n\} = \{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\}$:

$$\mathbb{P}(\{X_1 = x_1, ..., X_n = x_n\})$$
$$= \mathbb{P}(\{X_1 = x_1\})P(\{X_2 = x_2\}|\{X_1 = x_1\}) \cdots \mathbb{P}(\{X_n = x_n\}|\{X_1 = x_1...., X_{n-1} = x_{n-1}\}).$$

Possiamo infatti *imporre* direttamente, a partire dalla descrizione del problema, che le probabilità condizionate del tipo

$$\mathbb{P}(\{X_{r+1} = 1\}|\{X_1 = x_1...., X_r = x_r\}),$$

per $1 \leq r \leq n-1$ siano uguali a

$$\frac{m_1 - \sum_{i=1}^r x_i}{M - r + 1},$$

cioè uguali al rapporto fra il numero $m_1 - \sum_{i=1}^r x_i$ degli elementi di tipo A rimasti nella popolazione dopo le prime $(r-1)$ estrazioni ed il numero complessivo $M - (r-1) = M - r + 1$ degli elementi rimasti nella popolazione.

Notiamo dunque che tali probabilità condizionate non vengono ***calcolate*** *utilizzando la loro definizione (data in Definizione 4.1 della Lezione 4), ma vengono direttamente* ***assegnate*** *a partire dalle condizioni del problema.*

Ciò costituisce un esempio di quanto era stato già accennato più in generale in merito alla nozione di probabilità condizionata: *si giunge cioè a calcolare delle probabilità di eventi composti non tramite un calcolo combinatorio, bensì assegnando delle probabilità condizionate e delle condizioni di simmetria fra eventi; tali considerazioni sono analoghe a quelle già svolte nell'Esempio 6.2.*

Osservazione 2. Il problema qui affrontato riguarda il calcolo della probabilità di avere k elementi di tipo A in un campionamento casuale di n oggetti da una popolazione complessivamente costituita da M elementi, di cui m_1 di tipo A e $m_2 = (M - m_1)$ di tipo B. Tale calcolo si applica a problemi di diverso tipo (quali *estrazioni da urne*, *sondaggio elettorale*, *analisi statistica di una popolazione*, etc...) tutti, fra di loro, sostanzialmente isomorfi. È interessante confrontare fra loro le due formule (42) e (43). Entrambe risolvono il problema detto; la prima riguarda però estrazioni con reinserimento, mentre la seconda riguarda estrazioni senza reinserimento. Intuitivamente ci possiamo aspettare che le due formule tendano a coincidere nel caso in cui M, m_1 ed m_2 siano numeri molto grandi rispetto a n; infatti in tal caso ci si può aspettare che non vi sia grande differenza fra estrazioni con o senza reinserimento. Ciò può essere formalizzato come segue: è possibile dimostrare che, per fissati valori di n, k (con $0 \leq k \leq n$) e θ (con $0 < \theta < 1$), mandando M ed m_1 all'infinito in modo che $\frac{m_1}{M}$ tende a θ, allora risulta

$$\lim_{\substack{M\to\infty, m_1\to\infty \\ \frac{m_1}{M}\to\theta}} \frac{\binom{m_1}{k}\binom{M-m_1\rfloor}{n-k}}{\binom{M}{n}} = \binom{n}{k}\theta^k(1-\theta)^{n-k}.$$

Ad esempio, indicando con $\lfloor x \rfloor$ la *parte intera*[12] di un numero reale x, si può prendere $m_1 = \lfloor \theta M \rfloor$, e in effetti risulta

$$\lim_{M\to\infty} \frac{\binom{\lfloor \theta M \rfloor}{k}\binom{M-\lfloor \theta M \rfloor}{n-k}}{\binom{M}{n}} = \binom{n}{k}\theta^k(1-\theta)^{n-k}.$$

[12]Ricordiamo che la parte intera $\lfloor x \rfloor$ di x è quel numero intero k tale che $k \leq x < k + 1$

Convergenza delle probabilità ipergeometriche alle probabilità binomiali

Dimostriamo che

$$\lim_{\substack{M\to\infty, m_1\to\infty\\ \frac{m_1}{M}\to\theta}} \frac{\binom{m_1}{k}\binom{M-m_1}{n-k}}{\binom{M}{n}} = \binom{n}{k}\theta^k(1-\theta)^{n-k}.$$

Ricordiamo che, per arrivare alla (43), abbiamo mostrato che

$$\begin{aligned}
&\binom{n}{k}\frac{m_1!}{(m_1-k)!}\frac{m_2!}{(m_2-(n-k))!}\frac{(M-n)!}{M!}\\
&=\frac{n!}{k!(n-k)!}\frac{m_1!}{(m_1-k)!}\frac{m_2!}{(m_2-(n-k))!}\frac{(M-n)!}{M!}\\
&=\frac{m_1!}{k!\,(m_1-k)!}\frac{m_2!}{(n-k)!\,(m_2-(n-k))!}\frac{(M-n)!\,n!}{M!}\\
&=\frac{\binom{m_1}{k}\binom{m_2}{n-k}}{\binom{M}{n}}.
\end{aligned}$$

Utilizzando la precedente relazione con $m_2 = M - m_1$, basta quindi verificare che

$$\lim_{\substack{M\to\infty, m_1\to\infty\\ \frac{m_1}{M}\to\theta}} \frac{m_1!}{(m_1-k)!}\frac{(M-m_1)!}{(M-m_1-n-k)!}\frac{(M-n)!}{M!} = \theta^k(1-\theta)^{n-k}.$$

Infatti si ha

$$\begin{aligned}
&\frac{m_1!}{(m_1-k)!}\frac{(M-m_1)!}{(M-m_1-n-k)!}\frac{(M-n)!}{M!}\\
&=\frac{m_1}{M}\left(\frac{m_1}{M-1}-\frac{1}{M-1}\right)\cdots\left(\frac{m_1}{M-(k-1)}-\frac{k-1}{M-(k-1)}\right)\cdot\\
&\cdot\left(\frac{M}{M-k}-\frac{m_1}{M-k}\right)\left(\frac{M-1}{M-(k+1)}-\frac{m_1}{M-(k+1)}\right)\cdots\left(\frac{M-(n-k-1)}{M-(n-1)}-\frac{m_1}{M-(n-1)}\right)
\end{aligned}$$

e la tesi si ottiene considerando che, per ogni i, j,

$$\lim_{M\to\infty}\frac{j}{M-j}=0, \qquad \lim_{M\to\infty}\frac{M-i}{M-j}=1,$$

$$\lim_{\substack{M\to\infty, m_1\to\infty\\ \frac{m_1}{M}\to\theta}}\frac{m_1}{M-j} = \lim_{\substack{M\to\infty, m_1\to\infty\\ \frac{m_1}{M}\to\theta}}\frac{m_1}{M}\,\frac{M}{M-j}=\theta.$$

Nel caso in cui $m_1 = \lfloor\theta M\rfloor$ basta mandare M all'infinito in quanto chiaramente $m_1 = \lfloor\theta M\rfloor$ tende ad infinito se M tende all'infinito, e inoltre

$$\lim_{M\to\infty}\frac{\lfloor\theta M\rfloor}{M}=\theta.$$

L'ultimo limite si ottiene tenendo conto che

$$\lfloor\theta M\rfloor \le \theta M < \lfloor\theta M\rfloor + 1$$

e che quindi

$$0 \le \theta M - \lfloor\theta M\rfloor < 1, \quad \text{ovvero} \quad 0 \le \frac{\theta M}{M} - \frac{\lfloor\theta M\rfloor}{M} < \frac{1}{M}.$$

6.4 Esercizi di verifica

Esercizio 6.1. Un candidato ad un'elezione ha bisogno di almeno 50 voti per essere eletto. Prepara allora una lettera per informare i potenziali elettori circa la sua candidatura, il suo programma elettorale, etc....
Egli valuta che ogni persona che riceve la lettera si recherà effettivamente a votare per lui con una probabilità del 40%, indipendentemente dal comportamento degli altri (e si sottointende che egli certamente non ottiene voti da coloro cui non ha inviato la lettera).
a) Qual è la probabilità che egli riceva esattamente 51 voti se invia la lettera a 200 persone? (basta trovare l'espressione)
b) Qual è la probabilità di essere eletto se invia la lettera a 100 persone? (basta trovare l'espressione)
c) Caratterizzare il numero minimo di persone cui deve inviare copia della lettera affinché la probabilità di essere eletto sia superiore all'80%?

Esercizio 6.2. Si prendono a caso $n = 5$ viti da una scatola contenente complessivamente $M = 26$ viti, di cui alcune nuove ed altre usurate.
a) Supponendo che la scatola contiene $m_1 = 20$ viti nuove e $m_2 = 6$ viti usurate, calcolare la probabilità che almeno quattro delle cinque viti scelte siano nuove.
b) Si supponga ora di non conoscere inizialmente il numero M_1 delle viti nuove nella scatola e si ponga

$$\mathbb{P}(\{M_1 = h\}) = \binom{26}{h} \left(\frac{4}{5}\right)^h \left(\frac{1}{5}\right)^{26-h}, \qquad h = 0, 1, 2, ..., 26$$

Dopo aver verificato che tutte le cinque viti scelte sono nuove, come va calcolata la probabilità dell'ipotesi $\{M_1 = 26\}$?

Esercizio 6.3. In un lotto di 15 lampadine, ve ne sono 5 guaste. Se ne estraggono a caso 3. Calcolare la probabilità dei seguenti eventi:
a) nessuna lampadina difettosa fra le tre estratte
b) esattamente una lampadina difettosa fra le tre estratte
c) almeno una lampadina difettosa fra le tre estratte.
Si considerino separatamente i due diversi casi in cui
i) si estraggano le tre lampadine contemporaneamente
ii) le estrazioni sono con reimbussolamento

Esercizio 6.4. Si hanno m esemplari di un certo tipo di telecomando (TC) per televisore; ciascun TC ha bisogno di due batterie per il suo funzionamento. Si hanno a disposizione $2m$ batterie, di cui però h cariche e $(2m - h)$ scariche. Da tale gruppo di batterie vengono costituite in modo casuale m coppie, che vengono inserite negli m TC.
Calcolare la probabilità che un fissato TC abbia entrambe le batterie cariche.

Esercizio 6.5. Riottenere la formula (43) seguendo le indicazioni contenute nella precedente ***Osservazione 1*** di pagina 61.

7 Variabili aleatorie e distribuzioni di probabilità

Varie questioni incontrate nelle precedenti lezioni trovano una adeguata formalizzazione tramite l'introduzione della nozione di ***variabile aleatoria***. Nei precedenti esempi, infatti, ci siamo ripetutamente imbattuti in oggetti quali: somma dei punteggi nel lancio di due dadi, numero di votanti per uno schieramento in un sondaggio elettorale, numero di successi su n prove bernoulliane, massimo fra i cinque numeri risultanti da un'estrazione del lotto, etc....

A parte la diversa natura dei problemi considerati, notiamo che si è trattato in ogni caso di situazioni in cui possiamo elencare i valori che possono essere assunti da una certa grandezza X, ma sussiste una situazione di incertezza relativamente a quale sarà lo specifico valore che X effettivamente assume: il valore che essa assume sarà connesso (in qualche preciso modo) al risultato elementare di un qualche esperimento aleatorio; in base alla misura di probabilità assegnata sullo spazio campione in tale esperimento, potremo valutare la probabilità che si presentino i vari possibili valori per la grandezza.

Tali considerazioni motivano le definizioni seguenti.

Definizione 7.1 (provvisoria)**.** *Sia* Ω *un insieme finito. Una applicazione* $X:\ \Omega \to X(\Omega) \subseteq \mathbb{R}$ *viene detta* **variabile aleatoria** *(definita su* $(\Omega, \mathcal{P}(\Omega))$*).*

Osservazione 1. Essendo Ω un insieme finito, l'immagine di X, ovvero $X(\Omega)$, sarà un insieme del tipo $\{x_1, ..., x_n\}$, con $x_\ell \in \mathbb{R}$, per $\ell = 1, ..., n$.

Consideriamo ora gli eventi

$$X^{-1}(\{x_\ell\}) \equiv \{\omega \in \Omega:\ X(\omega) = x_\ell\}, \quad \ell = 1, 2, ..., n.$$

Tali eventi vengono anche indicati brevemente con i simboli

$$\{X = x_1\}, ..., \{X = x_n\}.$$

È immediato verificare che la famiglia degli eventi

$$\big\{\{X = x_1\}, ..., \{X = x_n\}\big\}$$

costituisce una partizione di Ω; per tale motivo se $\mathbb{P}:\ \mathcal{P}(\Omega) \to [0,1]$ è una probabilità, ponendo

$$p_X(x_\ell) \equiv \mathbb{P}(X = x_\ell), \quad \ell = 1, ..., n,$$

risulta

$$\sum_{\ell=1}^{n} p_X(x_\ell) = \sum_{\ell=1}^{n} \mathbb{P}(X = x_\ell) = 1, \qquad p_X(x_\ell) \geq 0, \qquad \ell = 1, ..., n.$$

La funzione $p_X:\ X(\Omega) \to \mathbb{R};\ x \mapsto p_X(x) \equiv P(X = x)$, così definita viene detta ***densità discreta*** di X.

Per semplicità di notazione, a volte useremo il simbolo p_ℓ al posto della densità discreta, ossia porremo

$$p_\ell = p_X(x_\ell) \equiv \mathbb{P}(X = x_\ell), \quad \ell = 1, ..., n,$$

e ovviamente risulta

$$\sum_{\ell=1}^{n} p_\ell = 1, \qquad p_\ell \geq 0, \qquad \ell = 1, ..., n.$$

La precedente proprietà ci permette di definire un nuovo spazio di probabilità

$$(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$$

dove, per $A \in \mathcal{P}(X(\Omega))$, la probabilità $P_X(A)$ è definita da

$$P_X(A) \equiv \sum_{\ell : x_\ell \in A} p_\ell = \sum_{\ell : x_\ell \in A} p_X(x_\ell). \tag{44}$$

Come vedremo tra poco vale la seguente relazione:

$$P_X(A) \equiv \mathbb{P}\left(X^{-1}(A)\right), \tag{45}$$

dove $X^{-1}(A)$ indica la controimmagine dell'insieme A tramite la funzione X, ovvero

$$X^{-1}(A) = \{\omega \in \Omega : \ X(\omega) \in A\}.$$

In genere, per comodità di notazione, si scrive più brevemente $\{X \in A\}$ invece di $\{\omega \in \Omega : \ X(\omega) \in A\}$.

Riassumendo, possiamo interpretare $P_X(A)$ come $\mathbb{P}(\{X \in A\})$.

Tale interpretazione e l'equivalenza tra (44) e (45) sono basate sul fatto che l'evento $X^{-1}(A) = \{X \in A\}$ si può scrivere come

$$\{X \in A\} = \bigcup_{\ell : x_\ell \in A} \{X = x_\ell\},$$

e che, di conseguenza,

$$P_X(X^{-1}(A)) \equiv \mathbb{P}(\{X \in A\}) = \sum_{\ell : x_\ell \in A} \mathbb{P}(\{X = x_\ell\}) = \sum_{\ell : x_\ell \in A} p_X(x_\ell) = \sum_{\ell : x_\ell \in A} p_\ell$$

Definizione 7.2 (provvisoria)**.** *La misura di probabilità $P_X(\cdot)$ su $\mathcal{P}(X(\Omega))$ definita da* (45) *o equivalentemente da* (44) *prende il nome di* **distribuzione di probabilità** *della variabile aleatoria X.*

Osservazione 2. Ovviamente a qualunque variabile aleatoria definita su uno spazio di probabilità possiamo associare la sua distribuzione di probabilità. Per individuare la distribuzione di probabilità di una variabile aleatoria X basta specificare la sua immagine $X(\Omega)$, cioè l'insieme dei valori che può assumere X, e, grazie alla relazione 44, dalla densità discreta p_X, ossia dai valori

$$p_\ell = p_X(x_\ell) = \mathbb{P}(X = x_\ell) \quad \text{per ogni } x_\ell \in X(\Omega).$$

Due diverse variabili aleatorie, definite o meno su uno stesso spazio di probabilità, possono dar luogo ad una stessa distribuzione di probabilità (vedere i due successivi Esempi 7.1 e 7.2).

È opportuno innanzitutto richiamare l'attenzione sui due particolari tipi di variabili aleatorie: le variabili aleatorie ***degeneri*** e le variabili aleatorie ***binarie***.

Definizione 7.3 (variabili aleatorie degeneri)**.** *Diciamo che X, variabile aleatoria definita su Ω, è una variabile aleatoria* **degenere** *se esiste $\overline{x} \in \mathbb{R}$, tale che $X(\Omega) = \{\overline{x}\}$, cioè se X è costante su Ω.*

La distribuzione di una variabile aleatoria degenere è banale: $X(\Omega) = \{\overline{x}\}$, $\mathcal{P}(X(\Omega)) = \big\{\emptyset, \{\overline{x}\}\big\}$, con $P_X(\emptyset) = 0$ e $P_X(\{\overline{x}\}) = 1$.

Definizione 7.4 (variabili aleatorie binarie)**.** *Diciamo che X, variabile aleatoria definita su Ω, è una variabile aleatoria* **binaria** *se $X(\Omega) = \{0, 1\}$.*

Definizione 7.5 (indicatore di un evento)**.** *Sia $E \in \mathcal{P}(\Omega)$ un evento e sia $\mathcal{X}_E$ la funzione definita da*[13]*:*

$$\mathcal{X}_E(\omega) = \begin{cases} 1 & \textit{per } \omega \in E \\ 0 & \textit{per } \omega \notin E \end{cases}$$

La funzione $\mathcal{X}_E$ è dunque una variabile aleatoria binaria, che viene chiamata con il termine **indicatore** *di E (o anche* **funzione indicatrice***) di E.*

Ricordiamo infine che invece di $\mathcal{X}_E$ si usa talvolta anche il simbolo $\mathbf{1}_E$, per mettere in evidenza il fatto che vale 1 su E.

Osservazione 3. Per qualunque v.a. binaria X esiste $E \subseteq \Omega = \{\omega_1, \cdots, \omega_N\}$ tale che

$$X(\omega_i) = \mathcal{X}_E(\omega_i), \qquad i = 1, ..., N.$$

Poniamo infatti

$$E = \{\omega_i \in \Omega : \ X(\omega_i) = 1\}$$

Si avrà allora

$$\overline{E} = \{\omega_i \in \Omega : \ X(\omega_i) = 0\}$$

e dunque possiamo scrivere $X(\omega_i) = \mathcal{X}_E(\omega_i), \qquad i = 1, ..., N.$

La distribuzione di una variabile binaria $\mathcal{X}_E$, con $p = P(E)$, è individuata ovviamente dal fatto che $X(\Omega) = \{0, 1\}$ e da $p_0 = 1 - p$ e $p_1 = p$.

Proposizione 1. Una qualunque variabile aleatoria X si scrive come combinazione lineare di variabili aleatorie binarie.

Dimostrazione.

Sia $X(\Omega) = \{x_1, ..., x_n\}$ l'insieme dei valori assumibili da X; consideriamo gli eventi

$$H_\ell \equiv \{X = x_\ell\}, \quad \ell = 1, ..., n$$

e le variabili aleatorie binarie $X_1, ..., X_n$ definite come indicatori di tali eventi, ovvero $X_\ell = \mathcal{X}_{H_\ell}$. È ovvio allora che possiamo scrivere

$$X(\omega_i) = \sum_{\ell=1}^{n} x_\ell \cdot X_\ell(\omega_i), \qquad i = 1, ..., N.$$

Infatti basta mostrare che la funzione $Y(\omega)$, definita su Ω da

$$Y(\omega) \equiv \sum_{\ell=1}^{n} x_\ell \cdot X_\ell(\omega)$$

coincide con $X(\omega)$ per ogni ω. A questo scopo basta osservare che la famiglia di eventi H_ℓ forma una partizione dell'evento certo Ω e di conseguenza basta mostrare che

$$Y(\omega) \equiv \sum_{\ell=1}^{n} x_\ell \cdot X_\ell(\omega) = X(\omega) \qquad \text{per ogni } \ \omega \in H_\kappa, \qquad \text{e per ogni } \ \kappa = 1, ..., n.$$

[13] In altre discipline della Matematica tale funzione viene usualmente chiamata ***funzione caratteristica*** di E, ma in Probabilità si usa un termine diverso, in quanto il termine ***funzione caratteristica*** è utilizzato per un altro concetto, che verrà studiato in corsi successivi.

E infatti per ogni $\omega \in H_\kappa$ ovviamente $X(\omega) = x_\kappa$, e anche $Y(\omega) = x_\kappa$, come si vede subito, tenendo conto che $\mathcal{X}_{H_\ell}(\omega) = 0$ se $\ell \neq \kappa$ e che ovviamente $\mathcal{X}_{H_\kappa}(\omega) = 1$:

$$Y(\omega) = x_\kappa \cdot \mathcal{X}_{H_\kappa}(\omega) + \sum_{\ell \neq \kappa} x_k \cdot \mathcal{X}_{H_\ell}(\omega) = x_\kappa \cdot 1 + 0 = x_\kappa.$$

Osservazione 4. Spesso si mira a determinare direttamente la distribuzione di probabilità di una variabile aleatoria, sulla base di considerazioni circa l'esperimento consistente nell'osservare il valore della variabile stessa. In tali casi non teniamo conto dello spazio di probabilità Ω su cui la variabile può essere definita, ne' come tale variabile vi possa essere definita, ne' quale sia la misura di probabilità su $(\Omega, \mathcal{P}(\Omega))$.

Vediamo ora qualche esempio di distribuzione di probabilità.

Esempio 7.1. *Consideriamo ancora una volta l'esperimento legato al lancio di due dadi, in cui*

$$\Omega \equiv \{(h,k) : \ 1 \leq h \leq 6, \ 1 \leq k \leq 6\}.$$

$$\mathbb{P}(\{(h,k)\}) = \frac{1}{36}, \qquad 1 \leq h \leq 6, \quad 1 \leq k \leq 6$$

Su questo spazio possiamo definire diverse variabili aleatorie, ad esempio:

$X_1 : (h,k) \to h$ *("punteggio del primo dado"),* $X_2 : (h,k) \to k$ *("punteggio del secondo dado"),* $X : (h,k) \to h+k$ *("somma dei due punteggi"),* $W : (h,k) \to \frac{h}{k}$*, etc...*

X_1 *e* X_2 *hanno la stessa distribuzione di probabilità, data da :*

$$\mathbb{P}(\{X_1 = x\}) = P(\{X_2 = x\}) = \frac{1}{6}, \qquad x = 1, 2, ...6;$$

questa è la **distribuzione uniforme** *su {1,2,...,6}.*

La distribuzione di probabilità di $X = X_1 + X_2$ *è invece data da*

$$\mathbb{P}(\{X = x\}) = \frac{x-1}{36}, \quad x = 2, 3, ..., 7, \qquad \mathbb{P}(\{X = x\}) = \frac{13-x}{36}, \quad x = 8, 9, ..., 12.$$

Esempio 7.2. *Riprendiamo l'Esempio 4.2, considerando il caso* $n = 6$*. La variabile aleatoria*

$$T \equiv \textit{numero dei tentativi fino a trovare la chiave giusta}$$

può prendere i valori $1, 2, ..., 6$ *e risulta, grazie a quanto avevamo visto,*

$$\mathbb{P}(T = x) = \frac{1}{6}, \quad x = 1, 2, ...6.$$

Confrontando tale risultato con quanto visto prima, troviamo l'esempio di due variabili aleatorie (cioè X_1 e T), definite su spazi diversi, che hanno la stessa distribuzione di probabilità.

Prima di considerare il successivo esempio è utile fare mente locale sulla seguente semplice

Osservazione 4. Siano $E_1, ..., E_n$ degli eventi in uno spazio $(\Omega, \mathcal{P}(\Omega))$ e indichiamo rispettivamente con $X_1, ..., X_n$ i loro indicatori. Ovviamente $X_1, ..., X_n$ sono delle variabili aleatorie definite su $(\Omega, \mathcal{P}(\Omega))$. Su $(\Omega, \mathcal{P}(\Omega))$ possiamo definire anche la variabile aleatoria $S_n \equiv \sum_{h=1}^n X_h$; poniamo cioè

$$S_n(\omega_i) \equiv \sum_{h=1}^n X_h(\omega_i) = \sum_{h=1}^n \mathbf{1}_{E_h}(\omega_i).$$

Ovviamente S_n ha il significato di ***numero di successi fra gli eventi*** $E_1, ..., E_n$.

Prima di passare ad esaminare i due importanti Esempi 7.3 e 7.4, si noti che la variabile aleatoria S_n può assumere $n+1$ valori, ossia $S_n(\Omega) = \{0, 1, ..., n\}$, e che la famiglia degli n eventi E_h per $h = 1, ..., n$ ***non coincide*** con la partizione $\big\{S_n = 0,\ S_n = 1,\ ..., S_n = n\big\}$, che invece è costituita da $n+1$ eventi. Questa osservazione mostra anche che la rappresentazione di una variabile aleatoria come combinazione lineare di variabili aleatorie binarie non è unica: posto $H_\ell = \{S_n = \ell\}$ per $\ell = 0, 1, ..., n$ si ha che

$$S_n = \sum_{\ell=0}^{n} \ell\, \mathcal{X}_{H_\ell}$$

corrisponde alla rappresentazione usata nella dimostrazione della ***Proposizione 1***, che è diversa dalla precedente rappresentazione

$$S_n = \sum_{h=1}^{n} \mathbf{1}_{E_h}.$$

Esempio 7.3. *Consideriamo n prove bernoulliane, cioè n eventi completamente indipendenti, ciascuno di probabilità θ ($0 < \theta < 1$) e consideriamo la variabile aleatoria $S_n \equiv$ numero di successi sulle n prove. I valori possibili per tale variabile sono ovviamente $0, 1, ..., n$ e, come abbiamo visto nella lezione precedente, si ha,*

$$\mathbb{P}(\{S_n = k\}) = \binom{n}{k} \theta^k\, (1-\theta)^{n-k}\,, \qquad k = 0, 1, ..., n.$$

Si dice che S_n segue una **distribuzione binomiale di parametri** ***n*** *e* ***θ****; ciò si indica con il simbolo $S_n \sim b(n, \theta)$.*

Esempio 7.4. *Vengono eseguite n estrazioni casuali senza reinserimento da una popolazione che contiene complessivamente M elementi, di cui m_1 elementi di tipo A e m_2 elementi di tipo B. Consideriamo la variabile aleatoria $S_n \equiv$ numero di elementi di tipo A fra gli n elementi estratti. Sappiamo che vale*

$$\mathbb{P}(\{S_n = k\}) = \frac{\binom{m_1}{k}\binom{m_2}{n-k}}{\binom{M}{n}}, \qquad \max(0, n + m_1 - M) \le k \le \min\,(n, m_1)\,.$$

Si dice che S_n segue una **distribuzione ipergeometrica di parametri** *M, m_1, n e ciò si indica con il simbolo $S_n \sim Hyp(M, m_1, n)$.*

Finora abbiamo quasi esclusivamente considerato variabili aleatorie ***a valori interi*** (cioè tali che $X\,(\Omega) \subseteq \mathbb{Z}$); ma questi non sono gli unici casi di possibile interesse; nel caso considerato nel precedente Esempio 7.3, è interessante considerare anche la variabile aleatoria Y_n ("frequenza dei successi") definita dalla relazione

$$Y_n = \frac{S_n}{n}.$$

A questo proposito è interessante più in generale, date n variabili aleatorie $X_1, ..., X_n$, studiare il comportamento probabilistico della media aritmetica

$$Y_n = \frac{\sum_{h=1}^{n} X_h}{n}.$$

A tale tipo di variabile aleatoria, daremo particolare attenzione nel seguito, in particolare nella Lez. 10.

Nelle ultime lezioni ci occuperemo di variabili aleatorie che prendono valori in un intervallo continuo di numeri reali.

Tornando a variabili aleatorie a valori interi notiamo quanto segue.
Osservazione 6. Per una variabile aleatoria X, a valori interi, cioè $X(\Omega) \subseteq \mathbb{Z}$, può essere spesso conveniente calcolare la distribuzione di probabilità tenendo conto della relazione

$$\mathbb{P}(X = k) = \mathbb{P}(X \leq k) - \mathbb{P}(X \leq k-1), \qquad \text{per } k \in \mathbb{Z}.$$

Altre volte può essere conveniente tenere conto invece che

$$\begin{aligned} \mathbb{P}(X = k) &= \mathbb{P}(X \geq k) - \mathbb{P}(X \geq k+1) \\ &= \mathbb{P}(X > k-1) - \mathbb{P}(X > k), \qquad \text{per } k \in \mathbb{Z}. \end{aligned}$$

La dimostrazione delle relazioni seguenti è basato sul fatto che

$$\{X \leq k\} = \{X = k\} \cup \{X \leq k-1\} \quad \text{e} \quad \{X \geq k\} = \{X = k\} \cup \{X \geq k+1\},$$

da cui

$$\mathbb{P}(X \leq k) = \mathbb{P}(X = k) + \mathbb{P}(X \leq k-1) \quad \text{e} \quad \mathbb{P}(X \geq k) = \mathbb{P}(X = k) + \mathbb{P}(X \geq k+1)$$

e infine che $\{X \geq h\} = \{X > h-1\}$.

Esempio 7.5. *Siano X_1 e X_2 i punteggi ottenuti nel lancio di due dadi e consideriamo la variabile aleatoria Z definita come il massimo dei due punteggi. Individuare i valori che può assumere Z e con quali probabilità.*

Soluzione. I valori possibili per Z sono ovviamente $\{1, 2, ..., 6\}$; tenendo conto che le famiglie di eventi $\mathcal{A} \equiv \big\{\{X_1 = i\},\ i = 1, \cdots, 6\big\}$ e $\mathcal{B} \equiv \big\{\{X_2 = j\},\ j = 1, \cdots, 6\big\}$ sono indipendenti[14], (e quindi anche gli eventi del tipo $\{X_1 \in I\}$ e $\{X_2 \in J\}$ sono indipendenti[15]), risulta

$$\begin{aligned} \mathbb{P}(Z = x) &= \mathbb{P}(Z \leq x) - \mathbb{P}(Z \leq x-1) \\ &= \mathbb{P}(\{X_1 \leq x\} \cap \{X_2 \leq x\}) - \mathbb{P}(\{X_1 \leq x-1\} \cap \{X_2 \leq x-1\}) \\ &= \mathbb{P}(X_1 \leq x) \cdot \mathbb{P}(X_2 \leq x) - \mathbb{P}(X_1 \leq x-1) \cdot \mathbb{P}(X_2 \leq x-1) \\ &= \left(\frac{x}{6}\right)^2 - \left(\frac{x-1}{6}\right)^2 = \frac{2x-1}{36}, \qquad x = 1, 2, ..., 6. \end{aligned}$$

Esempio 7.6. *Siano X_1 e X_2 i punteggi ottenuti nel lancio di due dadi e consideriamo la variabile aleatoria W definita come il minimo dei due punteggi. Individuare i valori che può assumere W e con quali probabilità.*

Soluzione. I valori possibili per W sono ovviamente $\{1, 2, ..., 6\}$; tenendo conto che le famiglie di eventi $\mathcal{A} \equiv \big\{\{X_1 = i\},\ i = 1, \cdots, 6\big\}$ e $\mathcal{B} \equiv \big\{\{X_2 = j\},\ j = 1, \cdots, 6\big\}$ sono indipendenti (e quindi anche

[14] Questa proprietà sarà alla base della successiva definizione di indipendenza stocastica per le v.a. X_1 e X_2
[15] Si veda a questo proposito l'Esercizio proposto 5.6.

gli eventi del tipo $\{X_1 \in I\}$ e $\{X_2 \in J\}$ sono indipendenti), risulta

$$\begin{aligned}
\mathbb{P}(W = x) &= \mathbb{P}(W \geq x) - \mathbb{P}(W \geq x+1) \\
&= \mathbb{P}(\{X_1 \geq x\} \cap \{X_2 \geq x\}) - \mathbb{P}(\{X_1 \geq x+1\} \cap \{X_2 \geq x+1\}) \\
&= \mathbb{P}(X_1 \geq x) \cdot \mathbb{P}(X_2 \geq x) - \mathbb{P}(X_1 \geq x+1) \cdot \mathbb{P}(X_2 \geq x+1) \\
&= \mathbb{P}(X_1 > x-1) \cdot \mathbb{P}(X_2 > x-1) - \mathbb{P}(X_1 > x) \cdot \mathbb{P}(X_2 > x) \\
&= (1 - \mathbb{P}(X_1 \leq x-1)) \cdot (1 - \mathbb{P}(X_2 \leq x-1)) - (1 - \mathbb{P}(X_1 \leq x)) \cdot (1 - \mathbb{P}(X_2 \leq x)) \\
&= \left(1 - \frac{x-1}{6}\right)^2 - \left(1 - \frac{x}{6}\right)^2 = \frac{49 - 14x + x^2 - (36 - 12x + x^2)}{36} \\
&= \frac{13 - 2x}{36}, \qquad x = 1, 2, ..., 6.
\end{aligned}$$

Esercizio proposto 7.1. *Ripetere gli esempi precedenti nel caso in cui i due dadi sono truccati in modo che* $\mathbb{P}(X_l = i) = Ki$*, per* $i = 1, \cdots, 6$*, ed* $l = 1, 2$*, e si assuma l'indipendenza delle partizioni* $\mathcal{A}$ *e* $\mathcal{B}$*.*

Osservazione 7. Il concetto di variabile aleatoria così come introdotto in questa lezione (vedi Definizione 7.1), può talvolta apparire un po' forzato o artificiale a chi sia all'inizio dello studio della teoria assiomatica della probabilità. Di fatto, invece, esso si rivela di importanza fondamentale sia nella formalizzazione rigorosa che nella comprensione di numerose questioni specifiche del calcolo delle probabilità.

In particolare esso permette di dare un chiaro significato alle operazioni fra variabili aleatorie, estendendo in modo diretto a questi oggetti (essendo funzioni a valori reali) le operazioni definite nel campo dei numeri reali.

Ad esempio, come abbiamo precedentemente visto, la somma $X = X_1 + X_2$ di due variabili aleatorie X_1, X_2 (definite su uno stesso spazio Ω) altro non è che la funzione definita dalla relazione

$$X(\omega) = X_1(\omega) + X_2(\omega), \text{ per } \omega \in \Omega.$$

Avevamo già avvertito comunque che la Definizione 7.1 di variabile aleatoria, così come è stata formulata, è provvisoria. Come si vedrà, essa va infatti adeguatamente modificata e completata quando si passi a trattare il caso in cui Ω non è un insieme finito.

7.1 Esercizi di verifica

Esercizio 7.1. Indichiamo con $X_1, ..., X_5$ i 5 numeri estratti su una ruota del lotto (si estrae senza reinserimento da un'urna contenente i numeri {1,2,..., 90}) e sia inoltre X il valore più alto fra $X_1, ..., X_5$.
Calcolate $\mathbb{P}(X \leq k)$ e $\mathbb{P}(X = k)$ per $k = 1, 2...., 90$.

Esercizio 7.2. Supponiamo che X sia una variabile aleatoria a valori nell'insieme $\{0, 1, ..., n\}$ e che, per una coppia di costanti positive A e ρ, risulti

$$\mathbb{P}(X = k) = A \cdot \frac{\rho^k}{k!\,(n-k)!}, \qquad k = 0, 1, ..., n,$$

Dimostrate che X segue una distribuzione binomiale ed individuatene i parametri.

Esercizio 7.3. Individuate una distribuzione di probabilità (non degenere) per una variabile aleatoria X in modo tale che risulti degenere la distribuzione della variabile aleatoria $Y = X^2$.

Esercizio 7.4. Individuate una distribuzione di probabilità per una variabile aleatoria X (non binaria) in modo tale che $Y = X^2$ risulti una variabile aleatoria binaria.

Esercizio 7.5. Sia X una variabile aleatoria con distribuzione di probabilità binomiale $b(6, \frac{1}{3})$. Trovare qual è il valore più probabile per X.

Esercizio 7.6. Consideriamo una variabile aleatoria con distribuzione di probabilità ipergeometrica

$$X \propto Hyp(6, 3; 3).$$

Qual è il più probabile fra i due eventi $\{X \leq 1\}, \{X > 1\}$?

Esercizio 7.7. In una lotteria sono stati emessi 1000 biglietti e vengono distribuiti 2 primi premi del valore di 1000 Euro, 4 secondi premi del valore di 500 Euro e 20 terzi premi del valore di 100 Euro.
a) Trovate la distribuzione di probabilità della variabile aleatoria X che indica il valore della vincita associata ad un singolo biglietto.
b) Scrivete la distribuzione di probabilità della variabile aleatoria $2X$.
Un tizio ha acquistato 2 biglietti della lotteria ed indichiamo con Z la variabile aleatoria che indica il valore complessivo della sua vincita alla lotteria.
c) Trovate la distribuzione di probabilità di Z.

Esercizio 7.8. Sia S_n una variabile aleatoria binomiale di parametri n e θ.
a) Verificate che, per $k = 0, 1, ..., n-1$

$$\mathbb{P}(S_n = k+1) = \frac{n-k}{k+1} \frac{\theta}{1-\theta} \, \mathbb{P}(S_n = k).$$

b) Utilizzando la proprietà precedente, verificate che esiste un $\overline{k}$ tale che

$$\begin{aligned} \mathbb{P}(S_n = k+1) &\geq \mathbb{P}(S_n = k) \quad k \leq \overline{k} \\ \mathbb{P}(S_n = k+1) &\leq \mathbb{P}(S_n = k) \quad k > \overline{k} \end{aligned}$$