

# Università di Trieste

## Laurea in ingegneria elettronica e informatica

Enrico Piccin - Corso di Probabilità e Statistica - Prof. Marco Barchiesi

Anno Accademico 2021/2022 - 3 Marzo 2022

### Indice

<b>1</b>	<b>Introduzione</b>	<b>2</b>
1.1	Matematica della probabilità . . . . .	4
1.2	Regole della famiglia di eventi complessi . . . . .	4
1.3	Regole dell'applicazione di probabilità . . . . .	5
1.4	Spazio di probabilità . . . . .	5
1.4.1	Proprietà degli spazi di probabilità finiti . . . . .	6
1.4.2	Costruzione di spazi di probabilità finiti . . . . .	8
<b>2</b>	<b>Serie</b>	<b>11</b>
2.1	Condizione necessaria per la convergenza di una serie . . . . .	12
2.1.1	Serie armonica . . . . .	12
2.2	Serie armonica generalizzata . . . . .	13
2.3	Serie geometrica . . . . .	14
2.4	Criterio del confronto per serie a termini positivi . . . . .	15
<b>3</b>	<b>Spazi di probabilità generali</b>	<b>17</b>
3.1	$\sigma$ -algebra . . . . .	17
3.2	Probabilità generale . . . . .	18
3.3	Spazio di probabilità generale . . . . .	18

3 Marzo 2022

## 1 Introduzione

Si supponga di stare in un **universo deterministico**, ovvero tale per cui tutto ciò che accadrà in futuro è determinato dalla situazione nel preciso istante in cui si sta vivendo.

Dal punto di vista fisico, si supponga di voler analizzare un certo fenomeno, a patto di conoscere

1. la legge che regola tale fenomeno
2. i dati iniziali (riferiti allo stato iniziale del fenomeno)
3. le condizioni esterne al fenomeno oggetto di interesse

è sempre possibile predire quello che accadrà nel futuro relativamente al fenomeno oggetto di studio.

**Esempio:** Si consideri il **lancio di un dado**. Taluno è un fenomeno oggetto di studio e, come tale, deve essere analizzato conoscendo

1. la legge che regola il fenomeno: la legge di caduta dei gravi
2. i dati iniziali del problema: peso del dado, altezza iniziale, forza di attrazione gravitazionale, etc.
3. le condizioni esterne: vento, umidità, stabilità dell'aria, etc.

Naturalmente, attraverso queste informazioni, è possibile predire il comportamento del dado: quando esso viene lasciato, cade e si schianta al suolo.

Tuttavia, se ora si volesse anche sapere su che faccia il dado atterrerà, non si ha a disposizione una legge fisica che ne regola tale fenomeno, in quanto la legge di caduta dei gravi presuppone il corpo come puntiforme; inoltre il movimento dell'aria influenza significativamente la rotazione del corpo. Pertanto, per la determinazione dell'esito di tale fenomeno, non si hanno a disposizione informazioni sufficienti: non si conosce la legge che regola il fenomeno, i dati iniziali sono scarsamente influenti e le condizioni esterne sono troppo variabili. Ciò fa sì che l'output finale di tale fenomeno sia completamente sconosciuto, in quanto l'evento oggetto d'analisi è totalmente **casuale**, o più propriamente **aleatorio**.

Si noti, ovviamente, che anche per fenomeni apparentemente facili da predire, le condizioni iniziali che vengono poste per lo studio degli stessi comportano sempre un margine di incertezza e, quindi, di aleatorietà: non si può sempre sapere con precisione assoluta lo stato iniziale del sistema oggetto di studio.

Per cercare di far fronte a tale incertezza si può

- impiegare una legge molto più particolareggiata (più vicina alla perfezione) che regola il fenomeno interessato; misurare con maggiore precisione i dati iniziali e definire con più raffinatezza le condizioni esterne; tuttavia, tale procedimento comporterebbe un lavoro molto oneroso e scarsamente proficuo;
- cercare di capire quali sono i possibili output del fenomeno (ossia le 6 facce del dado) e associare a ciascuno di tali output un valore che fornisca un'informazione di carattere quantitativo in riferimento alla possibilità che esso sia l'effettivo output del fenomeno interessato.

Da quest'ultima alternativa segue la definizione di **probabilità**:

### PROBABILITÀ

La **probabilità** è un modo per **quantificare** quanto un possibile risultato sia **facilmente ottenibile**.

**Esempio:** Si consideri il lancio di un dado a 6 facce e si prendano in considerazione dei possibili risultati

1. esce il numero 6
2. esce un numero pari
3. esce un numero  $\leq 6$
4. esce un numero  $\leq 4$
5. esce un numero  $\geq 7$

Si capisce facilmente come il primo risultato (o evento) sia molto elementare, in quanto prende in considerazione una sola faccia del dado, mentre i restanti sono dei risultati (o eventi) più complessi, che si ottengono tramite aggregazione dei risultati elementari.

Dal punto di vista matematica, per l'analisi di questo fenomeno, si definisce un insieme  $\Omega$  dei possibili risultati elementari del lancio di un dado, ovvero

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

Naturalmente, ora, i risultati che sono stati esposti in principio non sono altro che dei **sottoinsiemi** dell'insieme  $\Omega$  appena definito, come mostrato di seguito:

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

Per attribuire a ciascuno di tali risultati un valore quantitativo che ne descriva la possibilità di verificarsi, si definisce  $p = p(A)$  come la **probabilità associata al risultato**  $A$ , ovvero sia un numero all'interno di una scala che, per convenzione, viene indicata nell'intervallo  $[0, 1]$ , che quantifica la facilità con cui il risultato si presenta.

Per esempio, la probabilità che esca un numero maggiore di 7 nel lancio di un dado a 6 facce è ovviamente nulla, in quanto a tale evento viene associato l'insieme vuoto. Questo significa che tale risultato (o evento) è **impossibile**, pertanto si assegna ad esso un valore di probabilità di fondo scala, ovvero

$$p(\emptyset) = 0$$

Analogamente, la probabilità che esca un numero minore o uguale a 6 nel lancio di un dado è ovviamente massima, in quanto a tale evento viene associato l'insieme  $\Omega$  stesso. Questo significa che tale risultato (o evento) è **certo**, pertanto si assegna ad esso un valore di probabilità di fine scala, ovvero

$$p(\Omega) = 1$$

È facile capire che se si considerano due eventi a cui sono associati due sottoinsiemi  $A$  e  $B$  disgiunti, tali per cui  $A \cap B = \emptyset$ , allora si avrà che

$$p(A \cup B) = p(A) + p(B)$$

Inoltre, se il dado è regolare, ha senso ed è plausibile associare lo stesso valore di probabilità a ciascuno degli eventi elementari, ovvero

$$p(\{1\}) = p(\{2\}) = p(\{3\}) = p(\{4\}) = p(\{5\}) = p(\{6\})$$

e se ora si sommano le probabilità di tutti gli eventi elementari, non si può non ottenere la probabilità dell'evento certo, ovvero

$$p(\{1\}) + p(\{2\}) + p(\{3\}) + p(\{4\}) + p(\{5\}) + p(\{6\}) = p(\Omega) = 1$$

Per implicazione logica, siccome la somma delle probabilità degli eventi elementari è pari a 1 ed essi sono **equiprobabili**, deve essere necessariamente che

$$p(\{1\}) = p(\{2\}) = p(\{3\}) = p(\{4\}) = p(\{5\}) = p(\{6\}) = \frac{1}{6}$$

A partire da tale evidenza, è possibile ora andare ad associare agli eventi complessi, aggregati di eventi elementari, una probabilità, come di seguito esposto

$$p(\{2, 4, 6\}) = p(\{2\}) + p(\{4\}) + p(\{6\}) = \frac{3}{6} = \frac{1}{2}$$

e similmente

$$p(\{1, 2, 3, 4\}) = p(\{1\}) + p(\{2\}) + p(\{3\}) + p(\{4\}) = \frac{4}{6} = \frac{2}{3}$$

Naturalmente, ora, se si dovesse definire la probabilità associata alla somma di due eventi non disgiunti, non si può ricorrere alla formula precedentemente esposta, in quanto bisogna anche tenere conto delle sovrapposizioni. Infatti

$$\frac{7}{6} = p(\{2, 4, 6\}) + p(\{1, 2, 3, 4\}) \neq p(\{1, 2, 3, 4, 6\}) = \frac{5}{6}$$

questo perché, per quanto si è detto, i due sottoinsieme associati ai rispettivi risultati non sono disgiunti, in quanto  $\{2, 4, 6\} \cap \{1, 2, 3, 4\} = \{2, 4\}$

## 1.1 Matematica della probabilità

Il compito della probabilità è quello di fornire delle regole, a partire dalle quali riuscire ad attribuire una valutazione quantitativa della possibilità di verificarsi di eventi più complessi, **basandosi sulla probabilità associata ad eventi più elementari**.

Non è, invece, compito della probabilità quello di attribuire i valori di probabilità agli eventi elementari (si pensi, banalmente, alla differenza tra un dado regolare e un dado truccato): infatti, tale compito è affidato alla statistica, in quanto molto più legato alla praticità e alla modalità dei assegnazione.

Una volta appresa l'assegnazione della probabilità agli eventi elementari, interviene la probabilità: in particolare, la struttura matematica alla base del calcolo della probabilità prevede tre importanti elementi

1. un insieme  $\Omega$  (che per il momento si considera finito);
2. una famiglia  $\mathcal{A}$  (non vuota) di sottoinsiemi di  $\Omega$  (spesso la famiglia di tutti i sottoinsiemi); mentre gli elementi di  $\Omega$  sono chiamati risultati elementari, gli elementi sottoinsiemi di  $\Omega$  appartenenti a tale famiglia sono chiamati risultati complessi, ottenuti come aggregazione di risultati elementari;
3. un'applicazione

$$p : \mathcal{A} \longrightarrow [0, 1]$$

che ad ogni sottoinsieme appartenente alla famiglia  $\mathcal{A}$  associa un valore compreso tra 0 e 1.

Gli elementi di  $\Omega$  sono detti **eventi elementari**, per cui  $\Omega$  è detto **spazio degli eventi elementari**, mentre gli elementi della famiglia  $\mathcal{A}$  prendono il nome di **risultati** (o eventi) **casuali** (o semplicemente eventi, configurazioni, traiettorie, campioni), per cui  $\mathcal{A}$  prende il nome di **spazio degli eventi casuali**.

Se due eventi hanno un'intersezione nulla (come numeri pari e numeri dispari), si dice che essi sono **mutualmente esclusivi**.

L'applicazione  $p$  è detta **probabilità**, mentre l'immagine attraverso  $p$  di un evento  $A$ , ovvero  $p(A)$  viene chiamata **probabilità dell'evento  $A$** .

## 1.2 Regole della famiglia di eventi complessi

La famiglia di eventi complessi  $\mathcal{A}$ , così come l'applicazione  $p$ , è sempre subordinata alla presenza di alcune importanti condizioni:

1. tra gli elementi della famiglia degli eventi complessi  $\mathcal{A}$  ci deve essere sempre l'**evento certo**:

$$\Omega \in \mathcal{A}$$

2. tra gli elementi della famiglia degli eventi complessi  $\mathcal{A}$  ci deve essere sempre l'**evento impossibile**:

$$\emptyset \in \mathcal{A}$$

3. se tra gli elementi della famiglia degli eventi complessi  $\mathcal{A}$  si considerano gli eventi  $A$  e  $B$ , ad  $\mathcal{A}$  deve appartenere anche la loro **unione** e la loro **intersezione**:

$$A \cup B \in \mathcal{A} \text{ e } A \cap B \in \mathcal{A}$$

4. se tra gli elementi della famiglia degli eventi complessi  $\mathcal{A}$  si considera l'evento  $A$ , ad  $\mathcal{A}$  deve appartenere anche il suo opposto:

$$A^c \in \mathcal{A}$$

Nella pratica, se si lavora con uno spazio degli eventi che è finito, come famiglia dei risultati si considera l'insieme di tutti i risultati possibili, ovvero la famiglia costituita da tutti i sottoinsiemi di  $\Omega$  (essendo  $\Omega$  **finito**, l'insieme dei sottoinsiemi di  $\Omega$  è ancora una famiglia finita con cui si può lavorare). Quando, invece,  $\Omega$  è infinito, non si potrà lavorare con tutti i sottoinsiemi di  $\Omega$ , ma si dovrà procedere a considerare una sua restrizione (potrebbe, infatti, non essere possibile considerare l'insieme di tutti sottoinsiemi dello spazio degli eventi elementari).

Una famiglia  $\mathcal{A}$  che soddisfa i quattro punti di cui sopra prende il nome di **algebra**: l'insieme delle parti è un esempio di algebra.

### 1.3 Regole dell'applicazione di probabilità

L'applicazione  $p$  tale per cui

$$p : \mathcal{A} \longrightarrow [0, 1]$$

deve soddisfare, anch'essa, determinate regole:

1. l'**evento certo** deve avere **probabilità massima**, ovvero

$$p(\Omega) = 1$$

2. l'**evento impossibile** deve avere **probabilità nulla**, ovvero

$$p(\emptyset) = 0$$

3. se due eventi  $A$  e  $B$  sono mutualmente esclusivi, ovvero tali che i sottoinsiemi di  $\Omega$  associati siano disgiunti, ovvero  $A \cap B = \emptyset$ , allora si ha che

$$p(A) + p(B) = p(A \cup B)$$

tale proprietà prende il nome di **additività** (sempre, però, relativamente a eventi disgiunti).

Un'applicazione  $p$  definita come segue

$$p : \mathcal{A} \longrightarrow [0, 1]$$

che soddisfa le tre proprietà di cui sopra prende il nome di **misura di probabilità**.

### 1.4 Spazio di probabilità

Dopo aver introdotto il concetto di **algebra** e **probabilità** è possibile fornire la definizione di **spazio di probabilità finito**, (finito in quanto lo spazio degli eventi elementari  $\Omega$  è finito):

### SPAZIO DI PROBABILITÀ FINITO

Per definizione si chiama **spazio di probabilità finito** una **terna**  $(\Omega, \mathcal{A}, p)$  dove

- $\Omega$  è un insieme **finito** che identifica tutti e soli gli eventi elementari;
- $\mathcal{A}$  è un'algebra di sottoinsiemi di  $\Omega$  (che, in quanto algebra, soddisfa le 4 regole precedentemente esposte), generalmente data da tutti i sottoinsiemi di *Omega*;
- $p$  è una probabilità su  $\mathcal{A}$  (che, essendo una probabilità, soddisfa le 3 regole precedentemente esposte).

#### 1.4.1 Proprietà degli spazi di probabilità finiti

Le proprietà degli spazi di probabilità finiti discendono dalle richieste, ossia dalle regole, definite in merito alla famiglia degli eventi complessi  $\mathcal{A}$ , la quale deve essere un'algebra e sulla funzione  $p$ , la quale deve essere una **probabilità**:

1. **Proprietà additiva e subadditiva**: di seguito si espone la definizione di proprietà additiva e subadditiva della probabilità:

#### PROPRIETÀ ADDITIVA E SUBADDITIVA

Se  $A_1, \dots, A_n \in \mathcal{A}$  e sono a due a due **disgiunti** (ovvero mutualmente esclusivi), allora

$$\sum_{k=1}^n p(A_k) = p\left(\bigcup_{k=1}^n A_k\right)$$

ciò significa che la probabilità è **additiva**.

Se gli eventi non sono a due a due disgiunti, allora è possibile solo affermare che

$$\sum_{k=1}^n p(A_k) \geq p\left(\bigcup_{k=1}^n A_k\right)$$

In particolare, è sempre verificato che

$$p(A) + p(B) \geq p(A \cup B), \quad \forall A, B \in \mathcal{A}$$

ciò significa che la probabilità è **subadditiva**.

**Dimostrazione 1:** La proprietà di additività e subadditività si dimostra per **induzione**: si chiami  $B$  l'evento così ottenuto

$$B = \bigcup_{k=1}^{n-1} A_k$$

Dal momento che per ipotesi gli eventi  $A_1, \dots, A_n \in \mathcal{A}$  sono disgiunti, deve essere che

$$A \cap \bigcup_{k=1}^{n-1} A_k = A \cap B = \emptyset$$

e quindi si ha che

$$p\left(\bigcup_{k=1}^n A_k\right) = p(A_n) + p\left(\bigcup_{k=1}^{n-1} A_k\right)$$

sfruttando la proprietà di **additività della probabilità rispetto ad eventi disgiunti** (ossia la proprietà 3 della probabilità).

Si procede analogamente, per induzione a scendere su  $n$  fino a ottenere che

$$p\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n p(A_k)$$

Per dimostrare anche la seconda implicazione, si deve ricorrere alla proprietà fondamentale dimostrata di seguito, la quale permette di affermare che, siccome  $p(A \cap B) \geq 0$ , allora

$$p(A \cup B) + p(A \cap B) = p(A) + p(B) \longrightarrow p(A) + p(B) \geq p(A \cup B)$$

Ancora una volta, per dimostrare la proprietà di cui sopra si procede per **induzione**: avendo dimostrato il passo base con  $n = 2$ , si suppone che tale proprietà sia verificata per  $n - 1$  e la si dimostri per  $n$ ; appare evidente che se si considerano  $n$  eventi non a due a due disgiunti, tali per cui

$$A_n \cap \bigcup_{k=1}^{n-1} A_k \neq \emptyset$$

allora segue che

$$p\left(\bigcup_{k=1}^n A_k\right) \leq p(A_n) + p\left(\bigcup_{k=1}^{n-1} A_k\right)$$

ma usando l'ipotesi induttiva si perviene al risultato seguente

$$\sum_{k=1}^n p(A_k) \geq p\left(\bigcup_{k=1}^n A_k\right)$$

2. **Proprietà fondamentale**: di seguito si espone una relazione fondamentale che riguarda il calcolo della probabilità:

### PROPRIETÀ FONDAMENTALE

Si verifica che

$$p(A \cup B) + p(A \cap B) = p(A) + p(B), \quad \forall A, B \in \mathcal{A}$$

**Dimostrazione 2**: Per dimostrare la proprietà fondamentale è opportuno osservare che  $A \cup B$  può essere vista come l'unione di tre **insiemi disgiunti**:

$$A - (A \cap B) \quad A \cap B \quad B - (A \cap B)$$

ovvero

$$A \cup B = (A - (A \cap B)) \cup (A \cap B) \cup (B - (A \cap B))$$

sfruttando ancora la proprietà di **additività della probabilità rispetto ad eventi disgiunti** nel caso  $n = 3$  e la proprietà di **monotonia** dimostrata di seguito, osservando che  $A \cap B \subset A$  e  $A \cap B \subset B$  si può concludere che

$$\begin{aligned} p(A \cup B) &= p(A - (A \cap B)) + p(A \cap B) + p(B - (A \cap B)) \\ &= p(A) - p(A \cap B) + p(A \cap B) + p(B) - p(A \cap B) = p(A) + p(B) - p(A \cap B) \end{aligned}$$

3. **Proprietà di monotonia:** di seguito si espone la proprietà che giustifica la monotonia della probabilità:

### PROPRIETÀ DI MONOTONIA

Si verifica che

$$p(A - B) = p(A) - p(B), \quad \forall A, B \in \mathcal{A} : B \subset A$$

In particolare, dal momento che  $p(A - B) \geq 0$ , allora deve essere che

$$B \subset A \longrightarrow p(B) \leq p(A)$$

ciò significa che la probabilità è **monotona (crescente)**, come intuitivamente ci si poteva aspettare.

**Dimostrazione 3:** È facile osservare che gli eventi  $A - B$  e l'evento  $B$  sono, per costruzione, due eventi mutualmente esclusivi (e quindi disgiunti); inoltre deve essere  $(A - B) \cup B = A$ , dal momento che  $B \subset A$ . Ciò permette di concludere che

$$p(A - B) + p(B) = p((A - B) \cup B) = p(A)$$

sfruttando la proprietà di **additività della probabilità rispetto ad eventi disgiunti**.

**Corollario 1.0.1 Osservazione:** Se si considera un evento  $A \in \mathcal{A}$ , allora si ha che

$$p(A^c) = 1 - p(A)$$

**Dimostrazione:** La dimostrazione è presto fatta e segue dai risultati precedenti. Infatti è ovvio che  $A$  e  $A^c$  sono due eventi disgiunti e che  $A \cup A^c = \Omega$ . Pertanto, sfruttando la **proprietà di additività della probabilità rispetto ad eventi disgiunti** e la probabilità dell'evento certo si ottiene che

$$p(A \cup A^c) = p(\Omega) = p(A) + p(A^c) \longrightarrow p(A) = 1 - p(A^c)$$

#### 1.4.2 Costruzione di spazi di probabilità finiti

Si consideri un **insieme finito**  $\Omega = \{\omega_1, \dots, \omega_n\}$  di eventi elementari (che prenderà il nome di **spazio campionario**) e un insieme  $\Delta = \{\alpha_1, \dots, \alpha_n\}$  di **numeri reali positivi** la cui somma deve produrre il valore 1, ovvero

$$\sum_{k=1}^n \alpha_k = 1 \quad \text{con } \alpha_k \in \mathbb{R}, \alpha_k \geq 0$$

Per la costruzione dello spazio di probabilità si considera come **algebra**  $\mathcal{A}$  degli eventi complessi (o aggregati) l'**insieme delle parti** di  $\Omega$  e si definisce una funzione

$$p : \mathcal{A} \longrightarrow [0, 1]$$

considerando i valori  $\alpha_1, \dots, \alpha_n$  come i valori di probabilità che verranno attribuiti a ciascun evento elementare di  $\Omega$ , ovvero sia

$$p(\{\omega_k\}) = \alpha_k, \quad \text{con } k = 1, \dots, n$$

Tale concetto deve poi essere esteso a qualsiasi evento  $A \in \mathcal{A}$  non vuoto, quale il seguente

$$A = \{\omega_{k,1}, \dots, \omega_{k,m}\}$$

ovvero un sottoinsieme non vuoto di  $\Omega$ . Allora la probabilità dell'evento  $A$  sarà definita come segue

$$p(A) = \sum_{j=1}^m p(\{\omega_{k,j}\}) = \sum_{j=1}^m \alpha_{k,j} = \sum_{\omega \in A} p(\{\omega\})$$



ponendo, per semplicità,  $p(\emptyset) = 0$ . In altre parole, **la probabilità di  $A$  è la somma delle probabilità degli eventi elementari che costituiscono  $A$ .**

Mentre la famiglia  $\mathcal{A}$  è necessariamente un'algebra (in quanto soddisfa le 4 proprietà precedentemente esposte), in quanto l'insieme delle parti di  $\Omega$ , è opportuno verificare che la funzione  $p$  presa in considerazione sia, effettivamente, una probabilità:

1. La probabilità dell'evento certo è naturalmente 1 per costruzione, in quanto

$$p(\Omega) = \sum_{k=1}^n p(\{w_k\}) = \sum_{k=1}^n \alpha_k = 1$$

2. Per costruzione, la probabilità dell'evento impossibile è nulla, ovvero:

$$p(\emptyset) = 0$$

3. Si verifica immediatamente, infine, che la funzione  $p$  è effettivamente additiva rispetto a eventi disgiunti, ovvero

$$p(A \cup B) = \sum_{\omega \in A \cup B} p(\{\omega\}) = \sum_{\omega \in A} p(\{\omega\}) + \sum_{\omega \in B} p(\{\omega\})$$

essendo  $A$  e  $B$  disgiunti, ovvero  $A \cup B = A + B$ .

**Esempio 1:** Nel caso di un dado regolare a sei facce, si è posto

$$\omega_k = k \text{ e } \alpha_k = \frac{1}{6}$$

ovverosia

$$\Omega = \{1, \dots, 6\} \text{ e } \Delta = \left\{ \frac{1}{6}, \dots, \frac{1}{6} \right\}$$

**Esempio 2:** Si consideri il lancio di due dadi regolari, ambedue a sei facce. Lo spazio degli eventi elementari è

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), (2, 2), \dots, (6, 1), (6, 2), \dots, (6, 6)\}$$

ovvero un insieme finito costituito da 36 elementi. Assumendo che i dadi siano uguali e con facce uniformi, è naturale pensare che la probabilità che esca uno degli eventi elementari sia la stessa, ovvero

$$p(\{(1, 1)\}) = p(\{(1, 2)\}) = \dots = p(\{(6, 6)\})$$

Il valore della probabilità di ciascuno degli eventi elementari è  $\frac{1}{36}$ : questo in quanto la somma delle probabilità di tutti gli eventi elementari deve essere 1. In questo caso, si sta quindi ponendo

$$\alpha_k = \frac{1}{36} \text{ per } k = 1, \dots, 36$$

Pertanto, considerando come algebra  $\mathcal{A}$  degli eventi l'insieme delle parti di  $\Omega$  e definendo una probabilità  $p$  su  $\mathcal{A}$  come esposto in precedenza, se  $A \in \mathcal{A}$ , allora

$$p(A) = \begin{array}{l} \text{somma delle probabilità degli} \\ \text{eventi elementari che apparten-} \\ \text{gono ad } \mathcal{A} \end{array} = \begin{array}{l} \text{numero degli elementi di } \mathcal{A} \\ \text{moltiplicato per } \frac{1}{36} \end{array}$$

**Esempio:** Si calcoli, ora, la probabilità dei seguenti eventi complessi:

1. escono due numeri uguali
2. la somma dei due numeri fa 8
3. almeno uno dei due numeri è pari

**Esempio 1:** Al primo evento è associato il sottoinsieme

$$A = \{(1, 1), (2, 2), \dots, (6, 6)\}$$

costituito da 6 elementi. Quindi si evince che

$$p(A) = \frac{6}{36} = \frac{1}{6}$$

Infatti, dopo aver assegnato in maniera arbitraria i valori di probabilità agli eventi elementari, si adopera l'insieme delle regole del calcolo della probabilità al fine di determinare la probabilità degli eventi complessi (o aggregati).

**Esempio 2:** Al secondo evento è associato il sottoinsieme

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

costituito da 5 elementi. Quindi si evince che

$$p(A) = \frac{5}{36}$$

**Esempio 1:** Al primo evento è associato il sottoinsieme

$$A = \left\{ \begin{array}{l} (2, 1), \dots, (2, 6), (4, 1), \dots, (4, 6), (6, 1), \dots, (6, 6), \\ (1, 2), (1, 4), (1, 6), (3, 2), (3, 4), (3, 6), (5, 2), (5, 4), (5, 6) \end{array} \right\}$$

costituito da 27 elementi. Quindi si evince che

$$p(A) = \frac{27}{36} = \frac{3}{4}$$

**Osservazione:** Si osservi, ancora una volta, che la scelta di porre

$$\alpha_k = \frac{1}{36} \text{ per } k = 1, \dots, 36$$

è di **natura totalmente arbitraria** e basata su considerazioni empiriche (dadi uguali, facce uniformi, etc.). Se i dadi avessero delle anomalie (fossero, per esempio, truccati), sarebbe più ragionevole pesare in modo diverso le probabilità degli eventi elementari.

Per esempio, se ci fosse uno squilibrio di peso sulle facce 1 dei dadi, si potrebbe decidere di porre

$$p(\{(1, 1)\}) = \frac{1}{9}$$

$$p(\{(1, 2)\}) = p(\{(1, 6)\}) = p(\{(2, 1)\}) = \dots = p(\{(6, 1)\}) = \frac{1}{18}$$

e sui restanti 25 eventi elementari porre

$$p(\{(\dots, \dots)\}) = \left(1 - \frac{1}{9} - \frac{10}{18}\right) \cdot \frac{1}{25} = \frac{1}{75}$$

## 2 Serie

Di seguito si espone la definizione di **serie numerica**:

### SERIE NUMERICA

Data una successione di numeri reali  $\{a_n\}_{n \in \mathbb{N}}$ , si chiama **serie numerica** la successione delle somme parziali

$$\left\{ s_n := \sum_{k=1}^n a_k \right\}_{n \in \mathbb{N}}$$

Più esplicitamente

$$\left\{ \begin{array}{l} s_1 = a_1 \\ s_2 = a_1 + a_2 \\ \dots \\ s_n = a_1 + \dots + a_n \\ \dots \end{array} \right.$$

Pertanto, effettuare la sommatoria di infinito termini  $a_n$  equivale ad effettuare

$$\lim_{n \rightarrow \infty} s_n$$

Tale successione si indica con il simbolo

$$\sum_{k=1}^{\infty} a_k$$

e si impiega lo stesso simbolo per indicare il limite per  $n \rightarrow \infty$  (ammesso che esista).

**Osservazione 1:** Se la successione  $\{a_n\}$  è una successione di numeri reali non negativi, allora, naturalmente,  $\{s_n\}$  è crescente. Quindi, come ben noto,  $\{s_n\}$  converge, oppure diverge a  $+\infty$ .

**Osservazione 2:** È chiaro che il simbolo

$$\sum_{k=1}^{\infty} a_k$$

viene utilizzato per indicare che la somma inizia a partire dal termine  $a_1$ . Il simbolo

$$\sum_{k=0}^{\infty} a_k$$

invece, è atto a indicare che la somma inizia a partire dal termine  $a_0$ . Similmente, si impiega il simbolo

$$\sum_{k=3}^{\infty} a_k$$

quando si inizia a sommare a partire dal termine  $a_3$ .

**Osservazione 3:** Si osservi che, naturalmente

$$\sum_{k=1}^{\infty} a_k$$

converge **se e solo se** converge

$$\sum_{k=3}^{\infty} a_k$$

Questo in quanto

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \dots + a_{j-1} + \sum_{k=0}^{\infty} a_k$$

per cui il limite a destra esiste **se e solo se** esiste quello di sinistra, in quanto la differenza tra le due successioni è costituita soltanto un numero finito di elementi (dalla costante  $a_1 + a_2 + \dots + a_{j-1}$ ) che non influiscono sulla convergenza della successione stessa.

**Osservazione 4:** Si osservi che

**Lemma 2.1** *Non è vero che ogni serie converge o diverge a  $\pm\infty$ .*

**Esempio:** Si consideri la seguente serie

$$\sum_{k=1}^{\infty} (-1)^k$$

tale serie non converge e non diverge: infatti la somma parziale con  $k$  pari converge a 0, mentre la somma parziale con  $k$  dispari converge a  $-1$ : pertanto la serie di partenza non converge e non diverge, in quanto non esiste il limite della somma di successione.

Si parla, in questo caso, di una **successione oscillante**:

$$s_1 = -1, s_2 = 0, s_3 = -1, \dots$$

## 2.1 Condizione necessaria per la convergenza di una serie

In genere non è facile stabilire se una serie sia convergente o meno (e se anche lo fosse, non è sempre facile stabilire il limite della stessa). Si consideri, a tal proposito, il seguente lemma:

**Lemma 2.2** *Si consideri la serie seguente*

$$\sum_{k=1}^{\infty} a_k$$

*Se essa converge, allora deve essere che*

$$\lim_{n \rightarrow \infty} a_n = 0$$

*cioè **condizione necessaria** affinché una serie converga è che i termini da sommare progressivamente diventino indefinitivamente piccoli.*

### 2.1.1 Serie armonica

**Osservazione:** Si osservi che tale condizione è necessaria, ma **non sufficiente**. E ciò è facilmente verificabile con l'esempio seguente: si consideri la **serie armonica** seguente

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

Tale serie diverge a  $+\infty$ , anche se

$$\lim_{k \rightarrow \infty} \frac{1}{k} = 0$$

in quanto essa va a 0 troppo lentamente e quindi non ha sufficienza per garantire la convergenza.

**Dimostrazione:** Si dimostri per assurdo che la **serie armonica** diverge a  $+\infty$ . Dal momento che taluna è una serie ottenuta come sommatoria di quantità positive o converge, oppure diverge. Si supponga per assurdo che non diverga: allora converge ad un certo  $l$ , ovvero dovrebbe essere che

$$\lim_{n \rightarrow \infty} s_n = l$$

Si può facilmente osservare che la differenza tra le due somme parziali  $s_{2n}$  e  $s_n$  si può scrivere come

$$s_{2n} - s_n = \sum_{k=n+1}^{2n} \frac{1}{k}$$

questo perché i primi  $n$  termini vengono eliminati nella differenza. Inoltre si ha che

$$s_{2n} - s_n = \sum_{k=n+1}^{2n} \frac{1}{k} = \underbrace{\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n}}_{n \text{ termini}} \geq n \cdot \frac{1}{2n} = \frac{1}{2}$$

in quanto, ovviamente

$$\frac{1}{n+1} > \frac{1}{n+2} > \dots > \frac{1}{2n}$$

ed essi sono proprio  $n$  termini. Ciò permette di affermare che

$$\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} \geq n \cdot \frac{1}{2n} = \frac{1}{2}$$

Per la definizione di limite si ha che definitivamente, ossia per  $n$  abbastanza grande, si ha che

$$\forall \epsilon > 0, \exists n_\epsilon \in \mathbb{N} : \forall n \geq n_\epsilon \longrightarrow |s_n - l| < \epsilon$$

Ponendo  $\epsilon = \frac{1}{4}$  si ottiene che

$$|s_n - l| < \frac{1}{4}$$

Da ciò segue che, per la disuguaglianza triangolare e ricordando che  $|w| = |-w|$

$$\frac{1}{2} \leq |s_{2n} - s_n| = |s_{2n} - l + l - s_n| \leq |s_{2n} - l| + |s_n - l| < \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

ricordando che il limite delle due somme parziali è identico, dal momento che per  $n \rightarrow \infty$   $s_n \rightarrow +\infty$ ,  $s_n \leq s_{2n}$  e si verifica che

$$\lim_{k \rightarrow \infty} \frac{1}{k} = 0$$

Dal momento che una quantità non può essere minore di se stessa si ottiene l'assurdo.

## 2.2 Serie armonica generalizzata

Di seguito si fornisce la definizione di **serie armonica generalizzata**:

### SERIE ARMONICA GENERALIZZATA

Dato  $\alpha \in \mathbb{R}$ , la serie

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$$

è detta **serie armonica generalizzata**.

**Osservazione:** Anche se la serie armonica è divergente, la presenza di un esponente  $\alpha > 1$  sul termine  $\frac{1}{k}$  fa sì che le somme parziali crescano più lentamente e quindi ci sia convergenza. Se

$\alpha < 1$ , invece, la serie che non può che divergere ancora più velocemente. Omettendo la dimostrazione, si può facilmente capire che

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha}} \begin{cases} \text{diverge a } +\infty \text{ se } \alpha \leq 1 \\ \text{converge se } \alpha > 1 \end{cases}$$

## 2.3 Serie geometrica

Di seguito si fornisce la definizione di **serie geometrica**:

### SERIE GEOMETRICA

Dato  $Q \in \mathbb{R}$ , la serie

$$\sum_{k=0}^{\infty} q^k$$

è detta **serie geometrica** di ragione  $q$  (e parte con  $k = 0$ ).

**Osservazione:** Il comportamento della serie è il seguente:

$$\sum_{k=0}^{\infty} q^k \begin{cases} \text{converge a } \frac{1}{1-q} \text{ se } q \in (-1, 1) \\ \text{diverge a } +\infty \text{ se } q \geq 1 \\ \text{indeterminata se } q \leq -1 \end{cases}$$

ove per **indeterminata** si intende che non converge, né diverge a  $\pm\infty$ .

**Dimostrazione:** Se  $q = 1$  il comportamento è ovvio, in quanto si somma indefinitamente il valore 1, per cui la serie diverge a  $+\infty$ . Se  $q \neq 1$ , allora si deve osservare che è possibile scrivere

$$1 - q^{n+1} = (1 - q) \cdot (1 + q + q^2 + \dots + q^n)$$

dividendo ambo i membri per il termine  $1 - q \neq 0$  in quanto per ipotesi si è assunto  $q \neq 1$ , si ottiene il seguente termine generico della serie geometrica:

$$s_n = 1 + q + q^2 + \dots + q^n = \frac{1 - q^{n+1}}{1 - q}$$

e la tesi segue per quanto noto sul limite di  $q^n$ :

- se  $q \in (-1, 1)$ , allora ovviamente

$$\lim_{n \rightarrow \infty} q^{n+1} = 0$$

per cui

$$\lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}$$

- se  $q > 1$ , allora ovviamente

$$\lim_{n \rightarrow \infty} q^{n+1} = +\infty$$

per cui

$$\lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = +\infty$$

in quanto si può portare fuori il “−”, senza avere problemi di segno.

- se  $q < 1$ , allora ovviamente

$$q^{n+1} > 0$$

se  $n + 1$  è pari, mentre

$$q^{n+1} < 0$$

se  $n + 1$  è dispari, quindi la serie è oscillante con ampiezza a crescere.

**Esempio:** Si calcoli la seguente serie geometrica:

$$\sum_{k=1}^{\infty} \left(\frac{1}{3}\right)^k$$

ricordando sempre, però, che la serie geometrica parte con  $k = 0$ . Ovviamente è noto che, se  $k = 0$

$$\left(\frac{1}{3}\right)^0 = 1$$

per cui si può scrivere

$$\sum_{k=1}^{\infty} \left(\frac{1}{3}\right)^k = -1 + \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k$$

ma siccome  $q = \frac{1}{3} \in (-1, 1)$ , allora la serie converge a

$$\frac{1}{1-q} = \frac{1}{1-\frac{1}{3}}$$

e si può scrivere

$$\sum_{k=1}^{\infty} \left(\frac{1}{3}\right)^k = -1 + \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k = -1 + \frac{1}{1-\frac{1}{3}} = \frac{1}{2}$$

## 2.4 Criterio del confronto per serie a termini positivi

Non è sempre facile capire se una serie converge o meno: anche se la serie è a termini positivi, non è facile determinare se essa converga o diverga a  $\pm\infty$ . Tuttavia, esistono diversi criteri, uno dei più basilari ed efficaci è il criterio del confronto:

### CRITERIO DEL CONFRONTO PER SERIE A TERMINI POSITIVI

Siano date due successioni  $a_n, b_n \geq 0$  e si supponga che esse siano tali che  $a_n \leq b_n, \forall n$ . Allora si ha che

- se  $\sum_{k=1}^{\infty} a_n$  diverge, diverge anche  $\sum_{k=1}^{\infty} b_n$
- se  $\sum_{k=1}^{\infty} b_n$  converge, allora converge anche  $\sum_{k=1}^{\infty} a_n$

Inoltre, detto  $l_b$  il limite della prima e  $l_a$  il limite della seconda, risulta  $l_a = l_b$ .

Si propone una breve dimostrazione (che si basa sui teoremi del confronto di successioni monotone):

- si indichi con  $s_n$  la successione delle somme parziali relativa ai termini  $a_n$ , ovvero

$$s_n = a_1 + \dots + a_n$$

- si indichi con  $r_n$  la successione delle somme parziali relativa ai termini  $b_n$ , ovvero

$$r_n = b_1 + \dots + b_n$$

Si supponga che la successione  $r_n$  converga e si chiami con  $l_b$  il suo limite, ovvero

$$\lim_{n \rightarrow \infty} r_n = l_b$$

sapendo, però, che  $a_n \leq b_n, \forall n$ , si evince che

$$s_n \leq r_n$$

ma siccome  $r_n$  è una successione crescente e convergente, i suoi termini sono diminuiti dall'alto dal limite  $l_b$ , da cui

$$s_n \leq r_n \leq l_b$$

per il teorema di convergenza delle successioni monotone. Ma siccome anche la successione  $s_n$  è crescente ed è limitata superiormente da  $l_b$ , per il teorema di esistenza del limite delle successioni monotone segue che

$$\exists \lim_{n \rightarrow \infty} a_n = l_a$$

per il teorema del confronto tra successioni, sapendo che  $a_n \leq b_n, \forall n$ , si ha che  $l_a \leq l_b$ . Generalmente per il confronto vengono utilizzate la serie geometrica e la serie armonica generalizzata (visto che già se ne conosce il comportamento).

**Dimostrazione:** Si consideri la seguente serie

$$\sum_{k=0}^{\infty} \frac{1}{k!}$$

e si dimostri che essa è **convergente**. Ricordando che  $0! = 1$ , per definizione, si ha che

$$\sum_{k=0}^n \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \dots + \frac{1}{n \cdot (n-1) \cdot \dots \cdot 2}$$

se ora, in ciascun denominatore, si procede ad eliminare tutti i prodotti tranne i primi due, ovvero lasciando solamente  $n \cdot (n-1)$ , è naturale ottenere una sommatoria maggiorata, per cui

$$\sum_{k=0}^n \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \dots + \frac{1}{n \cdot (n-1) \cdot \dots \cdot 2} \leq 1 + 1 + \frac{1}{2 \cdot 1} + \dots + \frac{1}{n \cdot (n-1)} = 2 + \sum_{k=2}^n \frac{1}{k \cdot (k-1)}$$

La serie a destra può essere così riscritta

$$\sum_{k=2}^{\infty} \frac{1}{k \cdot (k-1)} = \sum_{k=2}^{\infty} \left[ \frac{1}{k-1} - \frac{1}{k} \right]$$

e tale serie è **convergente** ad 1, in quanto  $s_n = 1 - \frac{1}{n}$ . Per verificarlo basta provare a scrivere esplicitamente la successione delle somme parziali

$$s_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots + \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 - \frac{1}{n}$$

in quanto tutti gli altri termini si semplificano. Applicando il criterio del confronto, è facile osservare che

$$\sum_{k=0}^{\infty} \frac{1}{k!} \leq 2 + \sum_{k=2}^{\infty} \frac{1}{k \cdot (k-1)} = 2 + 1 = 3$$

per cui si ottiene che anche la successione

$$\sum_{k=0}^{\infty} \frac{1}{k!}$$

è convergente e il suo limite è minore di 3. Tale limite prende il nome di **numero di Nepero** ed è indicato con  $e$ .



7 Marzo 2022

### 3 Spazi di probabilità generali

Finora sono stati considerati degli spazi degli eventi elementari finiti. Tuttavia, non è sempre possibile sapere a priori quanti elementi devono essere contenuti nello spazio campione su cui si andrà a lavorare, per cui si rende necessario operare con spazi arbitrariamente grandi, ovvero **non finiti**.

Tuttavia, ciò introduce una complicazione aggiuntiva: non sarà più possibile sommare un numero finito di probabilità, ma un numero arbitrariamente grande e quindi infinito.

#### 3.1 $\sigma$ -algebra

Per estensione con quanto esposto in precedenza, di seguito si espone la definizione di  $\sigma$ -algebra, ovvero un'algebra nella quale non si pone un limite sul numero di elementi:

##### $\sigma$ -ALGEBRA

Una famiglia  $\mathcal{A}$  di parti di un insieme  $\Omega$  è detta  **$\sigma$ -algebra** se

1. l'insieme vuoto e lo spazio campione stesso appartengono ad  $\mathcal{A}$ , ovvero

$$\emptyset, \Omega \in \mathcal{A}$$

2. se  $A \in \mathcal{A}$ , allora  $A^c \in \mathcal{A}$

3. mentre in un'algebra normale si chiedeva che dati  $n$  sottoinsiemi appartenenti all'algebra, anche la loro unione e intersezione di devono appartenere, nel caso di una  $\sigma$ -algebra si chiede che se si considera una successione di sottoinsiemi  $A_n$  all'interno della  $\sigma$ -algebra, ovvero  $\{A_n, n \in \mathbb{A}\} \subset \mathcal{A}$ , ossia una **quantità al più numerabile**, allora l'unione e l'intersezione di tali sottoinsiemi devono appartenere alla  $\sigma$ -algebra:

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \quad \text{e} \quad \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$$

**Osservazione:** Ovviamente, ogni  **$\sigma$ -algebra** è anche un'algebra, e i due concetti coincidono se  $\Omega$  ha un numero finito di elementi.

### 3.2 Probabilità generale

Così come su un'algebra è stato definito il concetto di probabilità, è opportuno estendere tale concetto anche ad una  $\sigma$ -algebra:

#### PROBABILITÀ PER SPAZI DI PROBABILITÀ GENERALI

Siano  $\Omega$  un insieme e la famiglia  $\mathcal{A}$  una  $\sigma$ -algebra su  $\Omega$ . Una (misura di) **probabilità** è una funzione

$$p : \mathcal{A} \longrightarrow [0, 1]$$

tale che

1. come per un'algebra normale, la probabilità dell'insieme vuoto sia nulla e quella dello spazio campione stesso sia massima, ovvero

$$p(\emptyset) = 0 \quad \text{e} \quad p(\Omega) = 1$$

2. se  $\{A_n, n \in \mathbb{N}\} \subset \mathcal{A}$  è una successione di insiemi appartenenti alla  $\sigma$ -algebra **a due a due disgiunti**, ovvero  $A_i \cap A_j = \emptyset, \forall i \neq j$ , allora

$$\sum_{n=1}^{\infty} p(A_n) = p\left(\bigcup_{n=1}^{\infty} A_n\right) \quad [\sigma\text{-additività}]$$

ovvero la probabilità deve essere  **$\sigma$ -additiva**. Da notare che

$$\sum_{n=1}^{\infty} p(A_n)$$

è una **serie numerica**, ovvero sia una **somma infinita** delle probabilità di ogni singolo insieme  $A_n$  considerato.

### 3.3 Spazio di probabilità generale

Avendo esteso il concetto di algebra e probabilità al caso generale, è possibile parlare di **spazio di probabilità generale**, la cui definizione viene di seguito esposta:

#### SPAZIO DI PROBABILITÀ GENERALE

Si chiama **spazio di probabilità** una terna  $(\Omega, \mathcal{A}, p)$  dove

- $\Omega$  è un insieme, ovvero lo spazio degli eventi elementari (spazio campione)
- $\mathcal{A}$  è una  $\sigma$ -algebra
- $p$  è una probabilità definita su  $\mathcal{A}$

Da notare, ovviamente, che i risultati visti in precedenza, nell'ambito di uno spazio campione finito, sono ancora validi, anche in questo ambito generale.

Si consideri, a tal proposito, il seguente esempio:

**Esempio:** Presa una moneta e lanciandola ripetutamente, si vuole determinare la probabilità che esca testa al lancio  $n$ -esimo, ma non prima.

Per la visualizzazione di tale problema dal punto di vista matematico, è possibile etichettare testa con il numero "0" e croce con il numero "1". È possibile, quindi, identificare i risultati di una successione di lanci con una successione che prenda valori in  $\{0, 1\}$ . Per esempio  $(0, 0, 1, 0, 1, 1, \dots)$  identifica una successione di lanci in cui al primo e al secondo lancio è uscita testa, al terzo croce,

al quarto di nuovo testa, etc...

Con questa identificazione l'evento elementare è, fissato  $n$ , il seguente:

$\omega_n$  = la prima volta che esce testa è al lancio  $n$ -esimo

= la famiglia delle successione che hanno "1" nei primi  $n - 1$  elementi e "0" all'evento  $n$ -esimo.

Lo spazio degli eventi elementari è dunque  $\Omega = \{\omega_n, n \in \mathbb{N}\}$ , ovvero l'insieme delle famiglie  $\omega_1, \omega_2$ , etc., il quale, quindi, ha un numero infinito, ossia una quantità numerabile, di elementi.

Per rispondere alla domanda che è stata posta in principio, è necessario stabilire un criterio ragionevole per assegnare la probabilità. È possibile assumere, in maniera puramente **euristica**, che

- in ogni singolo lancio, la probabilità che esca testa o croce sia la stessa;
- il risultato di un lancio sia indipendente dai risultati dei lanci precedenti.

Queste ultime sono ipotesi ragionevoli, ma come per il lancio di un dado, sono assunzioni totalmente arbitrarie. Per esempio, la moneta potrebbe essere sbilanciata e quindi testa potrebbe tendere a uscire più volte rispetto a croce.

Per stabilire la probabilità di verificarsi dell'evento  $\omega_1$  è necessario capire che cosa accade nel primo lancio: se esce testa la probabilità è presto calcolata  $\frac{1}{2}$ , ossia il 50%.

La probabilità che si verifichi  $\omega_2$  prevede che non sia uscita testa al primo lancio e serve che al secondo lancio esca proprio testa. Quindi la probabilità è  $\frac{1}{2}$  di  $\frac{1}{2}$ , cioè  $\frac{1}{4}$ , ossia il 25%.

È così via per l'evento  $\omega_n$ :

$$p(\{\omega_n\}) = \left(\frac{1}{2}\right)^n$$

**Osservazione:** Nonostante si abbia risposto alla domanda, non è stato definito completamente lo spazio di probabilità che si sta considerando.

È opportuno considerare come  $\sigma$ -algebra  $\mathcal{A}$ , come già visto, l'insieme di tutti i sottoinsiemi di  $\Omega$ : questo è possibile farlo nonostante  $\Omega$  sia infinito, in quanto è costituita da una **quantità numerabile** di elementi.

Dopodiché la probabilità su  $\mathcal{A}$  si definisce come in precedenza, ossia come somma delle probabilità degli eventi elementari costituenti  $\mathcal{A}$ ; in particolare

$$p(A) = \begin{cases} \sum_{\omega \in A} p(\{\omega\}) & \forall A \subset \Omega \text{ non vuoto} \\ 0 & \text{se } A = \emptyset \end{cases}$$

La funzione così definita è ben posta, poiché  $\Omega$  è numerabile e quindi  $\mathcal{A}$  è al più numerabile. Quindi

$$\sum_{\omega \in \mathcal{A}}$$

è una sommatoria finita o una serie (se il numero di elementi costituenti  $\mathcal{A}$  è infinito). Si ha, inoltre, che tale funzione è tutt'al più una probabilità, in quanto la somma delle probabilità di tutti gli eventi elementari è proprio 1

$$p(\Omega) = \sum_{n=0}^{\infty} p(\{\omega_n\}) = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n - 1 = \frac{1}{1 - \frac{1}{2}} - 1 = 2 - 1 = 1$$

essendo una serie geometrica di ragione  $\frac{1}{2}$ . Si può facilmente osservare che vale la proprietà di  **$\sigma$ -additività**: si calcoli la probabilità dell'evento

$$A = \{\text{esce testa per la prima volta tra il terzo e il sesto lancio (compresi)}\} = \{\omega_1, \omega_4, \omega_5, \omega_6\}$$

Per calcolare la probabilità di tale evento complesso  $\mathcal{A}$  sarà sufficiente provvedere alla somma delle probabilità dei singoli eventi che appartengono ad  $\mathcal{A}$ , secondo la definizione, ovvero:

$$p(A) = \sum_{n=3}^6 p(\{\omega_n\}) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^6$$

Si consideri un altro evento complesso:

$$A = \{\text{esce testa per la prima volta in un lancio pari}\} = \{\omega_n : n \text{ è pari}\} = \{\omega_{2m} : m \in \mathbb{N}\}$$

La probabilità di tale evento è, quindi:

$$p(\mathcal{A}) = \sum_{m=0}^{\infty} p(\{\omega_{2m}\}) = \sum_{m=0}^{\infty} \left(\frac{1}{2}\right)^{2m} = \sum_{m=0}^{\infty} \left(\frac{1}{4}\right)^m = \sum_{m=1}^{\infty} \left(\frac{1}{4}\right)^m - 1 = \frac{1}{1 - \frac{1}{4}} - 1 = \frac{4}{3} - 1 = \frac{1}{3}$$

**Osservazione:** L'esempio appena visto è molto simile a quelli visti nel caso finito. Il motivo per cui tale estensione al caso generale è stata così naturale è che  $\Omega$  è numerabile, nonostante sia infinito, ma nulla di più. Esso merita, quindi, una definizione specifica:

### SPAZIO DI PROBABILITÀ DISCRETO

Si dirà che uno spazio di probabilità  $(\Omega, \mathcal{A}, p)$  è **discreto** se  $\Omega$  è al più numerabile (ovvero o finito o numerabile) e  $\mathcal{A}$  è la famiglia di tutti i sottoinsiemi di  $\Omega$ . Naturalmente, per quanto già visto, la probabilità sulla famiglia  $\mathcal{A}$  viene definita come

$$p(A) = \begin{cases} \sum_{\omega \in A} p(\{\omega\}) & \forall A \subset \Omega \text{ non vuoto} \\ 0 & \text{se } A = \emptyset \end{cases}$$

Quindi, in base a tale formula, si evince che il comportamento di  $p$  è **completamente determinato dalle probabilità degli eventi elementari**.

**Osservazione:** Si osservi che l'espressione

$$\sum_{\omega \in A} p(\{\omega\})$$

è da interpretarsi come una serie. Tuttavia, nel calcolo di tale sommatoria è totalmente **ininfluente l'ordine con cui si vanno a sommare le probabilità** degli eventi costituenti  $\mathcal{A}$ , ma solamente perché taluna è una **serie convergente a termini positivi**. Pertanto, in questo caso, si parla di **convergenza assoluta** e, quindi il limite non cambia se i termini vengono riarrangiati.

Si considerino due questioni:

1. Si consideri l'alternativa di considerare nell'esempio appena trattato, lo spazio degli eventi elementari seguenti

$$\Omega = \{\text{successioni costituite da "0" e da "1"}\}$$

e, per  $n \in \mathbb{N}$ , gli eventi

$$A_n = \{\text{successioni che presentano "1" nei primi } n-1 \text{ elementi e "0" nell}'n\text{-esimo}\}$$

Ovvero ci si sta chiedendo se sia più conveniente considerare gli eventi oggetto di interesse come eventi elementari  $\omega_n$  invece che come eventi complessi  $A_n$ .

2. Inoltre, ci si deve chiedere per quale ragione non ci si limiti ad assegnare una probabilità esclusivamente sugli eventi elementari e si necessita, invece, di introdurre la famiglia  $\mathcal{A}$  e definire  $p$  su eventi complessi  $A \in \mathcal{A}$ .

La risposta alla seconda questione è presto detta: se  $\Omega$  non fosse un insieme finito o numerabile, i suoi sottoinsiemi potrebbero non essere finiti o numerabili, per cui non sarebbe possibile definire la probabilità dell'evento complesso semplicemente basandosi sulla sommatoria della probabilità degli eventi elementari, in quanto tale sommatoria potrebbe essere non finita, in quanto somma di quantità più che numerabili, perdendo di significato.

**Esempio:** Per rispondere alla prima questione, invece, si consideri un esempio pratico. Se si deve modellizzare la risposta elastica di un ponte di ferro costituito da delle travi di ferro; si compia il seguente parallelo

- Le travi sono gli eventi elementari  $\omega_n$ ;
- L'elasticità delle travi corrisponde alla probabilità di eventi elementari  $p(\{\omega_n\})$ ;
- Le regole con cui le travi interagiscono sono gli assiomi dello spazio di probabilità (quale l'additività);
- Il ponte è l'evento complesso  $\mathcal{A}$ .

In questo caso, per la modellizzazione, si sarebbe potuto certamente prendere come elemento base del ponte gli atomi di ferro e considerare sia le travi che il ponte strutture complesse; tuttavia, il concetto di elasticità sarebbe stato definito comunque a partire dalle travi e non sugli atomi, dove tale applicazione non avrebbe alcun senso. Similmente, si sarebbero potuto considerare come eventi elementari le successioni, però il concetto di probabilità va comunque definito a livello di evento  $A_n$  e non sulla singola successione, in cui non avrebbe avuto senso.

In effetti ogni successione avrebbe dovuto avere la stessa probabilità, ma ogni evento  $A_n$  ne contiene infinite e quindi  $A_n$  risulterebbe avere **probabilità infinita** (e ciò è impossibile).

In tale esempio, pertanto, dal punto di vista operativo, i due approcci non differiscono concettualmente; spesso, però, si usano spazi non discreti: il ponte potrebbe essere fatto da ben altro che semplici travi.

**Esercizio:** Si consideri un bersaglio con una macchia non omogenea al suo interno:

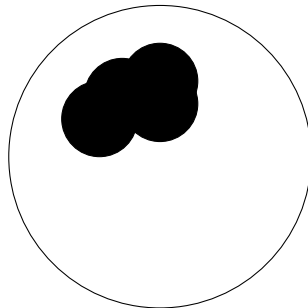


Figura 1: Esempio 4

Lanciando una freccetta, qual'è la probabilità di colpire la macchia scura?

In questo caso, ha senso considerare l'insieme degli eventi elementari  $\Omega$  l'insieme di tutti i punti del cerchio

$$\Omega := \{x \in \mathbb{R}^2 : x \in \text{cerchio}\}$$

Pertanto, sembra ragionevole assegnare ad una certa zona la probabilità di essere colpita in modo proporzionale alla sua area. In altre parole, se  $A$  è un sottoinsieme del cerchio  $\Omega$ , si ha che

$$p(A) = c \cdot |A|$$

dove  $|A|$  è l'area del sottoinsieme e  $c > 0$  è un opportuno fattore; questa è ovviamente un'**assunzione euristica**, totalmente arbitraria.

Affinché l'oggetto appena definito sia un'algebra (e  $p$  una probabilità), deve risultare  $p(\Omega) = 1$ , si può considerare come fattore rinormalizzante

$$c = \frac{1}{|\Omega|}$$

Nuovamente, la misura di probabilità così definita è una scelta arbitraria basata su considerazioni empiriche. Si sarebbe potuto ancora definirla in modo diverso, per esempio volendo tenere conto della gravità di zone poste in basso che, a parità di area, potrebbero avere più probabilità di essere colpite rispetto a zone poste in alto.

Si osservi, poi, che benché sia stato considerato come spazio degli eventi elementari l'insieme dei punti del bersaglio, ovvero uno spazio **né finito, né numerabile**, non ha senso definire prima la probabilità di colpire i singoli punti (che sarebbe sempre nulla) e poi la probabilità di colpire un'intera area del bersaglio, in quanto ad ogni area corrispondono infiniti punti e quindi la probabilità risultante sarebbe nulla.

Quando  $\Omega$  è un insieme al più numerabile, è sempre possibile scegliere  $\mathcal{A}$  come la famiglia di tutti i sottoinsiemi di  $\Omega$ ; tuttavia, in generale questo non è possibile farlo, o non risulta conveniente farlo. Infatti, in questo caso, lavorando con  $\Omega$  non finito e non numerabile, la  $\sigma$ -algebra  $\mathcal{A}$  non è

$$\mathcal{A} := \{\text{ogni sottoinsieme di } \Omega\}$$

in quanto poi dopo si avrebbe difficoltà a definire la probabilità  $p$ . In effetti, in questo caso, si è assunto

$$\mathcal{A} := \{\text{ogni sottoinsieme di } \Omega \text{ per cui abbia senso il concetto di area}\}$$

altrimenti non si sarebbe potuto scrivere  $|A|$ . In questo caso, la famiglia  $\mathcal{A}$  assunta è a tutti gli effetti una  $\sigma$ -algebra  $\mathcal{A}$ , in quanto l'unione di due eventi disgiunti di cui si conosce la misura (l'area) è ancora un evento di cui è nota la misura (l'area).

Ciò che in questo caso è rilevante è che in  $\mathcal{A}$  cadano tutti gli oggetti geometrici classici e le loro unioni ed intersezioni numerabili. Tale  $\sigma$ -algebra prende il nome di  **$\sigma$ -algebra di Borel**, ovvero la **famiglia di tutti i sottoinsiemi misurabili di  $\Omega$** , o in termini tecnici, la più piccola  $\sigma$ -algebra che contiene gli aperti di  $\Omega$ .

**Osservazione:** Si osservi che, innanzitutto, una probabilità è una **misura**, una misura della facilità con cui un evento si verifica e come qualsiasi altra misura presenta la seguente proprietà: se di tale oggetto se ne misurano le parti distinte (senza sovrapposizioni) per conoscere la misura dell'oggetto complessivo, è sufficiente sommare fra loro le singole misure distinte.

Volendo determinare ulteriori proprietà generali degli spazi di probabilità:

#### PROPRIETÀ DELLO SPAZIO DI PROBABILITÀ

Sia  $(\Omega, \mathcal{A}, p)$  uno spazio di probabilità

- se  $\{A_n\}$  è una successione crescente di eventi, cioè  $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$ , allora

$$p\left(\bigcup_{k=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} p(A_n) \quad [\text{continuità dal basso}]$$

ovvero la probabilità dell'unione di tali eventi dipende unicamente dal comportamento asintotico dell'ultimo evento, che contiene tutti gli altri, ovvero l'evento  $A_n$

- se  $\{A_n\}$  è una successione decrescente di eventi, cioè  $A_{n+1} \subset A_n, \forall n \in \mathbb{N}$ , allora

$$p\left(\bigcap_{k=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} p(A_n) \quad [\text{continuità dall'alto}]$$

**Dimostrazione 1:** È possibile scrivere  $\forall n \in \mathbb{N}$ , grazie al fatto che si sta lavorando con successioni crescenti

$$\bigcup_{k=1}^n A_k = A_1 \cup \bigcup_{k=2}^n (A_k - A_{k-1})$$

in modo tale da operare con insiemi disgiunti. Poiché

$$\bigcup_{k=1}^{\infty} A_k = A_1 \cup \bigcup_{k=2}^{\infty} (A_k - A_{k-1})$$

ovvero si sta considerando un insieme infinito di insiemi a due a due disgiunti, usando la  $\sigma$ -additività si ottiene

$$p\left(\bigcup_{k=1}^{\infty} A_k\right) = p(A_1) + \sum_{k=2}^{\infty} p(A_k - A_{k-1})$$

sfruttando il concetto di serie, una sommatoria infinita può essere considerata come il limite delle somme parziali, ovvero si ha che

$$p(A_1) + \sum_{k=2}^{\infty} p(A_k - A_{k-1}) = p(A_1) + \lim_{n \rightarrow +\infty} \sum_{k=2}^n p(A_k - A_{k-1})$$

Dal momento che  $p(A_1)$  è un numero reale, può essere tranquillamente portato dentro il limite, ottenendo

$$p(A_1) + \lim_{n \rightarrow +\infty} \sum_{k=2}^n p(A_k - A_{k-1}) = \lim_{n \rightarrow +\infty} \left[ p(A_1) + \sum_{k=2}^n p(A_k - A_{k-1}) \right]$$

Sfruttando ancora l'additività si possono rimettere nuovamente insieme tutti gli insiemi e calcolare il limite della probabilità dell'unione, come mostrato di seguito:

$$\lim_{n \rightarrow +\infty} \left[ p(A_1) + \sum_{k=2}^n p(A_k - A_{k-1}) \right] = \lim_{n \rightarrow +\infty} p\left(A_1 \cup \bigcup_{k=2}^n (A_k - A_{k-1})\right) = \lim_{n \rightarrow +\infty} p(A_n)$$

**Dimostrazione 2:** Per la dimostrazione della seconda proprietà, si deve osservare che  $\{A_1 - A_n\}$  è una successione necessariamente crescente. Dopo questa prima osservazione si può facilmente scrivere che

$$p(A_1) - p\left(\bigcap_{k=2}^{\infty} A_k\right) = p\left(A_1 - \bigcap_{k=2}^{\infty} A_k\right)$$

questo sfruttando il fatto che

$$\bigcap_{k=2}^{\infty} A_k \subset A_1$$

Inoltre, dal momento che l'intersezione degli  $A_k$  è uguale all'unione dei complementari si ottiene

$$p\left(A_1 - \bigcap_{k=2}^{\infty} A_k\right) = p\left(A_1 \cap \bigcup_{k=2}^{\infty} A_k^c\right)$$

Ora, invece, sfruttando la proprietà distributiva, si ha che  $A_1$  intersecato con l'unione è uguale all'unione delle intersezioni, per cui

$$p\left(A_1 \cap \bigcup_{k=2}^{\infty} A_k^c\right) = p\left(\bigcup_{k=2}^{\infty} (A_1 \cap A_k^c)\right)$$

Ma naturalmente si ha che  $A_1 \cap A_k^c = A_1 - A_k$ , ovvero l'intersezione tra  $A_1$  e il complementare di  $A_k$  è  $A_1$  a cui viene tolto  $A_k$ , per cui

$$p\left(\bigcup_{k=2}^{\infty} (A_1 \cap A_k^c)\right) = p\left(\bigcup_{k=2}^{\infty} (A_1 - A_k)\right)$$

Ma per l'osservazione iniziale, si ha che  $A_1 - A_k$  è una successione crescente, per cui per la **continuità dal basso** dimostrata in precedenza, segue che

$$p\left(\bigcup_{k=2}^{\infty} (A_1 - A_k)\right) = \lim_{n \rightarrow +\infty} p(A_1 - A_n) = p(A_1) - \lim_{n \rightarrow +\infty} p(A_n)$$

Per cui si è ottenuto, alla fine, che

$$p(A_1) - p\left(\bigcap_{k=2}^{\infty} A_k\right) = p(A_1) - \lim_{n \rightarrow +\infty} p(A_n) \longrightarrow p\left(\bigcap_{k=2}^{\infty} A_k\right) = \lim_{n \rightarrow +\infty} p(A_n)$$