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

1 Introduzione **2**
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 Proposizione 1 - Proprietà degli spazi di probabilità finiti 6
1.4.2 Costruzione di spazi di probabilità finiti 8

2 Serie **11**
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

3 Spazi di probabilità generali **17**
3.1 $\boldsymbol{\sigma}$-algebra 17
3.2 Probabilità generale 18
3.3 Spazio di probabilità generale 18
3.4 Proposizione 2 - Proprietà degli spazi di probabilità generali 22
3.5 Formula di inclusione-esclusione 28

4 Calcolo combinatorio **29**
4.1 Disposizione con ripetizione 29
4.2 Disposizione senza ripetizione 30
4.3 Combinazione 32
4.4 Coefficiente binomiale 32
4.4.1 Proprietà del coefficiente binomiale 33
4.5 Permutazione con ripetizione 34
4.6 Principio fondamentale del calcolo combinatorio 35
4.7 Combinazione con ripetizione 39
4.8 Estrazione 40

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 inziale, 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 un 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 conidizione 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 ancio 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 ≤ 6
4. esce un numero ≤ 4
5. esce un numero ≥ 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 Ω 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 Ω 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. $A = \varnothing$

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, ovverosia 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

$$p(\varnothing) = 0$$

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 Ω 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 sottinsiemi A e B disgiunti, tali per cui $A \cap B = \varnothing$, 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 elemantari, interviene la probabilità: in particolare, la struttura matematica alla base del calcolo della probabilità prevede tre importanti elementi

1. un insieme Ω (che per il momento si considera finito);
2. una famiglia $\mathcal{A}$ (non vuota) di sottoinsiemi di Ω (spesso la famiglia di tutti i sottoinsieme); mentre gli elementi di Ω sono chiamati risultati elementari, gli elementi sottoinsiemi di Ω 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 Ω sono detti **eventi elementari**, per cui Ω 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**:
$$\varnothing \in \mathcal{A}$$

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 Ω (essendo Ω **finito**, l'insieme dei sottoinsiemi di Ω è ancora una famiglia finita con cui si può lavorare). Quando, invece, Ω è infinito, non si potrà lavorare con tutti i sottoinsiemi di Ω, 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 quatto 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
$$p(\varnothing) = 0$$

3. se due eventi A e B sono mutualmente esclusivi, ovvero tali che i sottoinsiemi di Ω associati siano disgiunti, ovvero $A \cap B = \varnothing$, 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 Ω è finito):

SPAZIO DI PROBABILITÀ FINITO

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

- Ω è un insieme **finito** che identifica tutti e soli gli eventi elementari;
- $\mathcal{A}$ è un'algebra di sottoinsiemi di Ω (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 Proposizione 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 $A_1, ..., A_n \in \mathcal{A}$ e sono a due a due **disgiunti** (ovvero mutualmente esclusivi), allora

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

ciò significa che la probabilità è **additiva**.
Se gli eventi non sono a due a due digiunti, allora è possibile solo affermare che

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

In particolare, è sempre verificato che

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

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 $A_1, ..., A_n \in \mathcal{A}$ sono disgiunti, deve essere che

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

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

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

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

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

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

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

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

PROPRIETÀ FONDAMENTALE

Si verifica che

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

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

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

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

PROPRIETÀ DI MONOTONIA

Si verifica che

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

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 disgiutni e che $A \cup A^c = \Omega$. Pertanto, sfruttando la **proprietà di additività della probabilità rispetto ad eventi digiunti** 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** $\Omega = \{\omega_1, ..., \omega_n\}$ di eventi elementari (che prenderà il nome di **spazio campione**) e un insieme $\Delta = \{\alpha_1, ..., \alpha_n\}$ di **numeri reali positivi** la cui somma deve produrre il valore 1, ovvero

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

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

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

considerando i valori $\alpha_1, ..., \alpha_n$ come i valori di probabilità che verranno attribuiti a ciascun evento elementare di Ω, ovverosia

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

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

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

ovvero un sottoinsieme non vuoto di Ω. 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, $p(\varnothing) = 0$. In altre parole, **la probabilità di A è la somma delle probabilità degli eventi elemntari 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 Ω, 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:

$$p(\varnothing) = 0$$

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

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

   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

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

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

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

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

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

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

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

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

$p(A)$	$=$	somma delle probabilità degli eventi elementari che appartengono ad $\mathcal{A}$	$=$	numero degli elementi di $\mathcal{A}$ moltiplicato per $\frac{1}{36}$

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

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

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

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

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

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

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

$$p\left(\{(1,1)\}\right) = \frac{1}{9}$$
$$p\left(\{(1,2)\}\right) = p\left(\{(1,6)\}\right) = p\left(\{(2,1)\}\right) = ... = p\left(\{(6,1)\}\right) = \frac{1}{18}$$

e sui restanti 25 eventi elementari porre

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

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

$$\begin{cases} s_1 = a_1 \\ s_2 = a_1 + a_2 \\ ... \\ s_n = a_1 + ... + a_n \\ ... \end{cases}$$

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

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

Tale successione si indica con il simbolo

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

e si impiega lo stesso simbolo per indicare il limite per $n \to \infty$ (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 inzia a partire dal termine a_1. Il simbolo

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

invece, è atto a indicare che la somma inzia 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

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

per cui il limite a destra esste **se e solo se** esiste quello di sinistra, in quanto la differenza tra le due successione è costituita soltanto un numero finito di elementi (dalla costante $a_1 + a_2 + ... + a_{j-1}$) 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**:

$$s_1 = -1, s_2 = 0, s_3 = -1, ...$$

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

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

cioé ***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

$$\lim_{k\to\infty} = 0$$

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

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

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

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

in quanto, ovviamente

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

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

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

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 $n \to \infty$ $s_n \to +\infty$, $s_n \leq s_{2n}$ e si verifica che

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

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

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

dividendo ambo i mebri 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:

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

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

- se $q \in (-1, 1)$, allora ovviamente
$$\lim_{n\to\infty} q^{n+1} = 0$$
per cui
$$\lim_{n\to\infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}$$
- se $q > 1$, allora ovviamente
$$\lim_{n\to\infty} q^{n+1} = +\infty$$
per cui
$$\lim_{n\to\infty} \frac{1 - q^{n+1}}{1 - q} = +\infty$$
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 geoemtrica 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 dela 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
$$s_n = a_1 + ... + a_n$$
- si indichi con r_n la successione delle somme parziali relativa ai termini b_n, ovvero
$$r_n = b_1 + ... + b_n$$

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

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

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 diminati 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

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

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

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

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

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

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

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

in quanto tutti gli altri termini si semplificano. Applicando il criteriore 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 σ-algebra

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

σ-ALGEBRA

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

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

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

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 σ-algebra si chiede che se si considera una successione di sottoinsiemi A_n all'interno della σ-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 σ-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 **σ-algebra** è anche un'algebra, e i due concetti coincidono se Ω 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 σ-algebra:

PROBABILITÀ PER SPAZI DI PROBABILITÀ GENERALI

Siano Ω un insieme e la famiglia $\mathcal{A}$ una **σ-algebra** su Ω. 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

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

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

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

ovvero la probabilità deve essere **σ-additiva**. Da notare che

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

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

3.3 Spazio di probabilità generale

Avendo esteso il concetto di alebra e probabilità al caso generale, è posibile 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

- Ω è un insieme, ovvero lo spazio degli eventi elementari (spazio campione)
- $\mathcal{A}$ è una σ-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 ripetutatemente, si vuole determinare la proabilità 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 $(0, 0, 1, 0, 1, 1, ...)$ 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:

$$\begin{aligned}\omega_n = &\text{ la prima volta che esce testa è al lancio } n\text{-esimo}\\ = &\text{ la famiglia delle successione che hanno "1" neiprimi } n-1 \text{ elementi e "0" all'evento } n\text{-esimo.}\end{aligned}$$

Lo spazio degli eventi elementari è dunque $\Omega = \{\omega_n, n \in \mathbb{N}\}$, ovvero l'insimee delle fammiglie ω_1, ω_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 ragionebvole 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 as usicre più volte rispetto a croce.
Per stabilire la probabilità di verificarsi dell'evento ω_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 ω_2 prevede che non sia uscita testa al primo lancio e serve che al secono lancio esca proprio testa. Quindi la probabilità è $\frac{1}{2}$ di $\frac{1}{2}$, cioé $\frac{1}{4}$, ossia il 25%.
È così via per l'evento ω_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 σ-algebra $\mathcal{A}$, come già visto, l'insieme di tutti i sottoinsiemi di Ω: questo è possibile farlo nonostante Ω sia infinito, in quanto è costituita da una **quntità 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

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

La funzione così definita è ben posta, poiché Ω 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 utti gli effetti una probabilità, in quanto la somma delle probabilità di tutti gli eventi elementari è proprio 1

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

essendo una serie geometrica di ragione $\frac{1}{2}$. Si può facilmente osservare che vale la proprietà di **σ-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 definzione, ovvero:

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

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:

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

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 Ω 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 Ω è al più numerabile (ovvero o finito o numberabile) e $\mathcal{A}$ è la famiglia di tutti i sottoinsieme di Ω. Naturalmente, per quanto già visto, la probabilità sulla famiglia $\mathcal{A}$ viene definita come

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

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

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

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 elementi 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 ω_n invece che come eventi complessi A_n.

2. Inoltre, ci si deve chidere 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 Ω 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 ω_n;
- L'elastiticità 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 modellizzaione, 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:

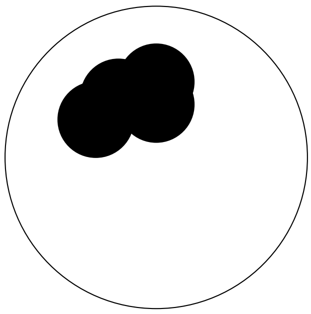


Figura 1: Esempio 4

Lanciando una freccetta, qual'è la probabilità di colpire la macchia scura?
In questo caso, ha senso considerare l'insieme degli eventi elementari Ω l'insieme di tutti i punti del cerchio

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

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 Ω, 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 del 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 Ω è un insieme al più numerabile, è sempre possibile scegliere $\mathcal{A}$ come la famiglia di tutti i sottoinsiemi di Ω; tuttavia, in generale questo non è possibile farlo, o non risulta conveniente farlo. Infatti, in questo caso, lavorando con Ω non finito e non numerabile, la σ-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 σ-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 σ-algebra prende il nome di **$\boldsymbol{\sigma}$-algebra di Borel**, ovvero la **famiglia di tutti i sottoinsiemi misurabili di $\boldsymbol{\Omega}$**, o in termini tecnici, la più piccola σ-algebra che contiene gli aperti di Ω.

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.

3.4 Proposizione 2 - Proprietà degli spazi di probabilità generali

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, cioé $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$, allora

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

  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, cioé $A_{n+1} \subset A_n, \forall n \in \mathbb{N}$, allora

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

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 infintio di insiemi a due a due disgiunti, usando la σ-additività si ottiene

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

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

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

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

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

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

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

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

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

Per cui si è ottenuto, alla fine, che

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

11 marzo 2022

È noto, come dimostrato nella **Proposizione 1** (§ 1.4.1), che la probabilità è **subadditiva**; più un generale, se $A_1, ..., A_n$ sono eventi (non necessariamente disgiunti), allora

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

E la **Proposizione 2** (§ 3.4 precedentemente esposta permette di generalizzare la *subadditività* alle successioni di eventi.

Corollario 3.0.1 *Sia* $\{A_n\}_{n\in\mathbb{N}}$ *una successione di eventi in uno spazio di probabilità* $(\Omega, \mathcal{A}, p)$. *Allora*

$$p\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} p(A_k) \quad [\boldsymbol{\sigma}\textbf{-}\textbf{\textit{subadditività}}]$$

Tale proprietà prende il nome di σ-subadditività.

Dimostrazione: Si consideri

$$B_n := \bigcup_{k=1}^{n} A_k$$

ovvero B_n sotituisce un aggregato di n successioni. Naturalmente, appare evidente come

$$B_{n+1} = B_n + A_{n+1} \geq B_n$$

per cui $\{B_n\}_{n\in\mathbb{N}}$ è una successione crescente di eventi e quindi $B_n \subset B_{n+1}$: ciò permettera di impiegare la **continuità dall'alto** dimostrata nella **Proposizione 2** (§ 3.4); inoltre, appare evidente come

$$\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k$$

per come è stato costruito l'oggetto B_k. Per la **continuità dall'alto** si ottiene che

$$p\left(\bigcup_{k=1}^{\infty} A_k\right) = p\left(\bigcup_{k=1}^{\infty} B_k\right) = \lim_{n\to+\infty} p\left(B_n\right)$$

ovvero la probabilità dell'aggregato, in una successione crescente, è stimata dal comportamento asintotico dell'elemento B_n. Riscrivendo B_n secondo la sua definizione si ha che

$$\lim_{n\to+\infty} p\left(B_n\right) = \lim_{n\to+\infty} p\left(\bigcup_{k=1}^{n} A_k\right)$$

sfruttando, ora, quanto è noto per il caso finito, ovvero che la probabilità dell'unione è minore della somma delle probabilità dei singoli eventi, si ha che

$$\lim_{n\to+\infty} p\left(\bigcup_{k=1}^{n} A_k\right) \leq \lim_{n\to+\infty} \sum_{k=1}^{n} p\left(A_k\right)$$

sfrutando, ora, il concetto di serie, si ha che il limite della serie è uguale al limite delle somme parziali, per cui

$$\lim_{n\to+\infty} \sum_{k=1}^{n} p\left(A_k\right) = \sum_{k=1}^{\infty} p\left(A_k\right)$$

in cui tale limite esiste sempre in quanto si sta considerando una serie a termini positivi. Pertanto tale limite può essere finito o $+\infty$, a seconda che la serie sia convergente o divergente.

Osservazione: È noto che, in generale, la proprietà

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

se gli eventi $A_1, ..., A_n$ non sono digunti, non è un'uguaglianza. È noto, però, nel caso di due eventi A e B (così chiamati in luogo di A_1 e A_2) disgiunti, che

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

in cui è facile osservare che affinché la disuguaglianza

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

sia una uguaglianza, deve essere considerato il termine $p(A \cap B)$. Ciò è ravvisabile ancora meglio se tale configurazione si analizza dal punto di vista geometrico:

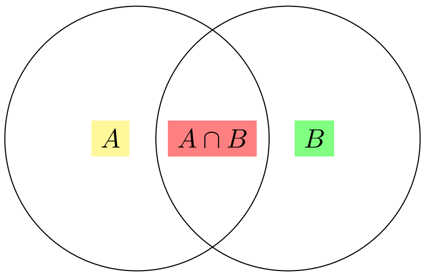


Figura 2: Intersezione tra due insiemi

Considerando la probabilità una **misura** (in questo caso dell'area dei due insiemi), quando viene eseguita la somma $p(A)+p(B)$, si sta considerando due volte la probabilità della loro intersezione, ovvero $p(A \cap B)$; pertanto, al fine di ottenere $p(A \cup B)$, è necessario sottrarre a $p(A)+p(B)$ proprio la probabilità dell'intersezione $p(A \cap B)$, ottenendo l'identità

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

Nel caso di $n = 3$ insiemi, si considerino A,B e C tre eventi

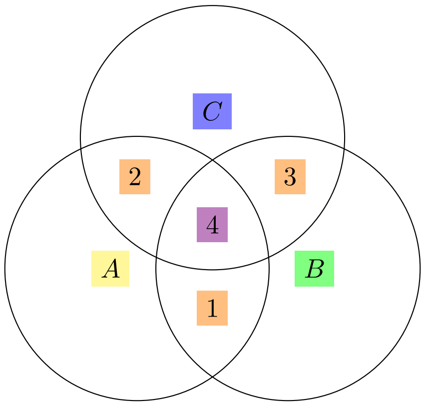


Figura 3: Intersezione tra tre insiemi

In questo secondo caso, quando si sommano $p(A) + p(B) + p(C)$ vengono misurate **due volte** le intersezioni a due a due degli insiemi (ovvero 1, 2, 3), mentre l'intersezione dei tre insiemi (4) viene misurata **tre volte**. Pertanto, se si vuole ottenere $p(A \cup B \cup C)$, alla somma delle tre probabilità $p(A) + p(B) + p(C)$ bisogna sottrarre $p(A \cap B) + p(A \cap C) + p(B \cap C)$. Tuttavia, in questo modo, è stata rimossa completamente l'intersezione tra i tre insiemi, che quindi deve essere riaggiunta, ottenendo:

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

Dimostrazione: Per la dimostrazione di tale proprietà, nel caso di 3 eventi, si impiegano le nozioni già note per 2 eventi, come mostrato di seguito:

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

Ora, l'unione degli eventi $A \cup B$ e C può essere scritta come somma delle probabilità degli eventi distinti meno la probabilità dell'intersezione, ovvero

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

Tuttavia, ora, $A \cup B$ è ancora l'unione di due eventi, per cui, secondo la formula, si ha che

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

in cui, semplicemente, per la proprietà distributiva degli insiemi, si osserva che $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$. Pertanto, ora, sfruttando ancora una volta la proprietà di unione di due insiemi si può scrivere che

$$p((A \cap C) \cup (B \cap C)) = P(A \cap C) + p(B \cap C) - p((A \cap C) \cap (B \cap C))$$

da cui si ottiene la seguente formula

$$= p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$$

tenuto conto che $(A \cap C) \cap (B \cap C) = A \cap B \cap C$.

Osservazione: Tale formula può essere riscritta in modo più compatto: se si indicano i tre insiemi con A_1, A_2, A_3 è possibile scrivere

$$p\left(\bigcup_{k=1}^{3} A_k\right) = \sum_{k=1}^{3} \sum_{j \subset \{1,2,3\} \text{ tale che } |j|=k} (-1)^{k+1} \cdot p\left(\bigcap_{i \in j} A_i\right)$$

in cui la prima sommatoria serve solo per fissare la lunghezza dell'indice, mentre la seconda sommatoria considera tutte le famiglie costituite dal numero di elementi indicati dall'indice. Infatti

- i sottoinsiemi $j \subset \{1,2,3\}$ per i quali $|j| = 1$ (ovvero che hanno cardinalità 1) sono $\{1\}, \{2\}, \{3\}$, quindi

$$\sum_{j \subset \{1,2,3\} \text{ tale che } |j|=1} (-1)^2 \cdot p\left(\bigcap_{i \in j} A_i\right) = p(A_1) + p(A_2) + p(A_3)$$

- i sottoinsiemi $I \subset \{1,2,3\}$ per i quali $|I| = 2$ sono $\{1,2\}, \{1,3\}, \{1,3\}$, quindi

$$\sum_{j \subset \{1,2,3\} \text{ tale che } |j|=2} (-1)^3 \cdot p\left(\bigcap_{i \in j} A_i\right) = -p(A_1 \cap A_2) - p(A_1 \cap A_3) - p(A_2 \cap A_3)$$

- I sottoinsiemi $I \subset \{1,2,3\}$ per i quali $|I| = 3$ sono $\{1,2,3\}$, quindi

$$\sum_{j \subset \{1,2,3\} \text{ tale che } |j|=3} (-1)^4 \cdot p\left(\bigcap_{i \in j} A_i\right) = p(A_1 \cap A_2 \cap A_3)$$

Eseguendo, ora, la somma, di ciascuno di tali casi si ottiene proprio la formula di partenza.

3.5 Formula di inclusione-esclusione

Procedendo, ora, per induzione, si può generalizzare tale formula al caso n, grazie al fatto che essa è stato scritta per mezzo di sommatorie, come illustrato di seguito:

FORMULA DI INCLUSIONE-ESCLUSIONE

Sia $(\omega, \mathcal{A}, p)$ uno spazio di probabilità. Presi n eventi $A_1, A_2, ..., A_n \in \mathcal{A}$ si ha che

$$p\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \sum_{j \subset \{1,...,n\} \text{ tale che } |j|=k} (-1)^{k+1} \cdot p\left(\bigcap_{i \in j} A_i\right)$$

Si chiama formula di inclusione-esclusione in quanto per i termini pari si aggiunge, mentre per quelli dispari si toglie, a causa delle sovrapposizioni.

4 Calcolo combinatorio

Volendo procedere al calcolo probabilistico, risulta fondamentale conoscere il numero di eventi elementari che vanno a costituire lo spazio campione Ω. Si espongono di seguito, per tale ragione, alcune fondamentali nozioni di calcolo combinatorio:

4.1 Disposizione con ripetizione

Si espone di seguito la definizione di **dispozione con ripetizione**, in cui si considera una scelta che non condiziona quelle successive:

DISPOSIZIONE CON RIPETIZIONE

Dato $k \in \mathbb{N}$ e un insieme finito A, si chiama **disposizione con ripetizione** di k elementi estratti da A (o anche di **classe** k) una funzione

$$f : \{1, ..., k\} \longrightarrow A$$

In altre parole, considerando un insieme A, una disposizione con ripetizione di k elementi estratti da A non è altro che una k-upla di elementi dell'insieme A.
Ciò che è importante osservare, in tal senso, è che è **possibile scegliere uno stesso elemento più volte**. Si descriva, quindi, tale scelta, con una funzione

$$f : \{1, ..., k\} \longrightarrow A$$

ponendo

$$\begin{aligned} &f(1) \text{ il primo elemento scelto} \\ &f(2) \text{ il secondo elemento scelto} \\ &f(3) \text{ il terzo elemento scelto} \\ &\quad ... \; ... \\ &f(k) \text{ il } k\text{-esimo elemento scelto} \end{aligned}$$

In cui vi è una corrispondenza biunivoca tra la famiglia delle possibili disposizioni e l'insieme prodotto A^k (ossia l'insieme delle k-uple con coordinate nell'insieme A): ad ogni disposizione f è possibile associare la k-upla $(f(1), ..., f(k))$, mentre ad ogni k-upla $(a_1, ..., a_k)$ è possibile associare la disposizione

$$f(j) := a_j \quad j = 1, ..., k$$

Osservazione: Si definisca quante disposizioni con ripetizione esistono di k elementi su un insieme A con n elementi; ovviamente, vista la corrispondenza biunivoca ve ne sono

$$\boxed{|A^k| = n^k}$$

in quanto si hanno n alternative per la scelta del primo elemento, n alternative per la scelta del secondo elemento e così via fino al k-esimo elemento.

Esempio: Per compilare la colonna di una schedina del totocalcio occorre scegliere, per ciascuna delle 13 partite considerate, un numero $A = \{1, \times, 2\}$ (1 vittoria interna, $\times$ pareggio, 2 vittoria esterna). Cioé bisogna effettuare una disposizione con ripetizione di $k = 13$ elementi dall'insieme A. Quindi ci sono $3^{13} \cong 1,6$ milioni di modi per compilare la schedina.

4.2 Disposizione senza ripetizione

Si espone di seguito la definizione di **dispozione senza ripetizione**, in cui si considera una scelta che condiziona quelle successive:

DISPOSIZIONE SENZA RIPETIZIONE

Dato $k \in \mathbb{N}$ e un insieme finito A, si chiama **disposizione senza ripetizione** di k elementi estratti da A una funzione **iniettiva**

$$f : \{1, ..., k\} \longrightarrow A$$

Le disposizioni senza ripetizione sono anche chiamate **disposizioni semplici**. Nel caso delle disposizioni con ripetizione, è assolutamente ininfluente la relazione tra k ed n, con $n = |A|$ numero di elementi dell'insieme A; nel caso di disposizioni semplici, deve essere che $k \leq n$, in quanto altrimenti la funzione non potrebbe essere iniettiva, visto che a valori distinti di k dovrebbe corrispondere uno stesso elemento.

PERMUTAZIONE

Data una disposizione senza ripetizione di k elementi estratti da A, se $k = |A|$, si parla di una **permutazione** dell'insieme A.

Dal punto di vista operativo, una disposizione senza ripetizione è una scelta di k elementi all'interno di un insieme A, con il vincolo che

- nello scegliere il secondo elemento, bisogna che esso sia diverso dal primo, cioé
$$f(2) \neq f(1)$$
- nello scegliere il terzo elemento bisogna fare in modo che esso sia diverso dal primo e dal secondo, ovvero
$$f(3) \neq f(1) \quad \wedge \quad f(3) \neq f(2)$$
- e così via.

Si determini quante disposizioni semplici esistono di k elementi su un insieme A con n elementi; naturalmente si hanno n possibilità per la prima scelta, $n-1$ per la seconda scelta (ovvero tutti gli elementi di A, escluso quello scelto precedentemente), e così via. Quindi il valore cercato è

$$\boxed{\underbrace{n \cdot (n-1) \cdot ... \cdot (n-k+1)}_{k \text{ scelte}} = \frac{n!}{(n-k)!}}$$

in cui l'uguaglianza si ottiene moltiplicando il lato sinistro per $(n-k)!$.

Esempio: Si consideri un mazzo di carte regolare, con $n = 52$ carte. Naturalmente vi sono

$$52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 \cong 312 \cdot 10^6$$

possibili modi per servire $k = 5$ carte.

Esercizio: Si determini la probabilità p_k che in un gruppo di k persone selezionate casualmente (nessuna delle quali sia però nata il 29 febbraio), almeno due di esse **compiono gli anni lo stesso giorno**.
Naturalmente, lo spazio degli eventi elementari Ω è costituito dalle possibili disposizioni con ripetizioni fi k elementi dall'insieme degli $n = 365$ giorni dell'anno (il 29/02 è escluso per ipotesi). Infatti per ognuna delle k persone viene selezionato un giorno di nascita. Si ha, quindi

$$|\Omega| = n^k = 365^k$$

La probabilità che una persona nasca in un dato giorno può essere assunta uniforme, e quindi può essere assunta uniforme anche la probabilità sulle disposizioni: ogni disposizione

$$f : \{1, ..., k\} \longrightarrow \{1, ..., 365\}$$

ha probabilità

$$\frac{1}{|\Omega|} = \frac{1}{365^k}$$

Per risolvere la domanda del problema bisogna provvedere ad effettuarne la traduzione dal punto di vista matematico; l'evento di cui si deve effettuare il calcolo della probabilità p_k è

$$A := \{f \in \Omega : f \text{ non è iniettiva}\}$$

Quindi

$$p_k = p(A) = 1 - p(A^c) = 1 - \frac{|A^c|}{|\Omega|}$$

ma siccome A^c è costituito dalle disposizioni semplici, si ha

$$|A^c| = \frac{365!}{(365-k)!} = \prod_{j=0}^{n-1}(365-j)$$

per cui, finalmente, si ottiene

$$p_k = 1 - \frac{\prod_{j=0}^{k-1}(365-j)}{365^k} = 1 - \prod_{j=0}^{k-1}\frac{365-j}{365} = 1 - \prod_{j=0}^{k-1}\left(1 - \frac{j}{365}\right)$$

Osservazione: Non appena $k \geq 23$, $p_k > \frac{1}{2}$. Cioé, in una classe di liceo vi è una probabilità superiore al 50% che due persone compiano il compleanno lo stesso giorno.

14 marzo 2022

4.3 Combinazione

Dopo aver esposto i concetti di disposizione con ripetizione e disposizione semplice, si espone di seguito il concetto di **combinazione**:

> **COMBINAZIONE**
>
> Dato $k \in \mathbb{N}$ e un insieme finito A, si chiama **combinazione** di k elementi estratti da A i sottoinsiemi di A di cardinalità k.

Osservazione: Si osservi che la differenza tra una disposizione e una combinazione è la rilevanza, nel primo, dell'ordine della disposizione degli elementi, fattore che nel secondo caso è ininfluente.
Infatti, ad una disposizione con ripetizione di k elementi è possibile associare una k-upla, e quindi ad una disposizione semplice una k-upla senza ripetizione. Pertanto, è possibile considerare una disposizione semplice come una **collezione ordinata**, mentre una combinazione è una **collezione non ordinata**.

Esempio: Si consideri l'insieme $A = \{\bigcirc, \square, \triangle\}$. Le disposizioni semplici di classe $k = 2$ elementi possono essere identificate con le coppie ordinate

$$(\bigcirc, \triangle) \quad (\bigcirc, \square) \quad (\triangle, \square) \quad (\triangle, \bigcirc) \quad (\square, \bigcirc) \quad (\square, \triangle)$$

Mentre le combinazioni, invece, sono

$$\{\bigcirc, \triangle\} \quad \{\bigcirc, \square\} \quad \{\triangle, \square\}$$

in quanto per le combinazioni non è rilevante l'ordine degli elementi dell'insieme, ma solamente gli elementi stessi.
Si determini, allora, il numero di combinazioni di k lementi relative ad un insieme con n elementi; al fine di rispondere a tale domanda, è sufficiente considerare una disposizione semplice di k elementi (collezione ordinata) e procedere, poi, con la permutazione di tutti i k elementi che la compongono, ottenendo una combinazione (collezione non ordinata). Da notare, infatti, che due collezioni ordinate (composte dagli stessi k elementi) restituiscono la stessa collezione non ordinata; viceversa, da una collezione non ordinata, procedendo al suo ordinamento tramite permutazione dei suoi k elementi si otterranno diverse collezioni ordinate, **tante quante sono le possibili permuazioni dei k elementi**.
Pertanto si ha

Numero di combinazioni di k elementi su n elementi	$\cdot$	Numero di permutazioni di k elementi	$=$	Numero di disposizioni semplici di k elementi su n elementi
$C_{n,k}$		$k!$		$\dfrac{n!}{(n-k)!}$

4.4 Coefficiente binomiale

Peranto si è ottenuto che il numero di combinazioni di k elementi su n elementi è proprio

$$\boxed{C_{n,k} = \frac{n!}{k! \cdot (n-k)!}}$$

Tale quantità prende il nome di **coefficiente binomiale**, il quale viene indicato con il simbolo

$$\binom{n}{k}$$

Tenendo conto che, naturalmente, $k > 0$ (in quanto non avrebbe senso una collezione di 0 elementi), dal punto di visto matematico, visto che per $k = 0$ si ha $k! = 1$, si pone anche

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

Anche nel caso delle combinazioni, naturalmente, ha senso supporre $k \leq n$, in quanto per $k > n$ non ci sono disposizioni semplici di k elementi, ovviamente. In generale, non esistono nemmeno combinazioni possibili con $k > n$, quindi si pone, per convenzione

$$\binom{n}{k} = 0 \quad \text{se} \quad k > n$$

Esempio: È noto che il numero dei possibili modi di servire $k = 5$ carte da un mazzo da $n = 52$ carte sono

$$\frac{52!}{(52-5)!} = \frac{52!}{47!}$$

ossia il numero di disposizioni semplici di 5 elementi su 52. Una volta, però, che le carte sono state distribuite, l'ordine con cui esse vengono disposte è ininfluente, pertanto vi sono

$$\binom{52}{5} = \frac{52!}{5! \cdot 47!}$$

possibili mani servite (ossia il numero di combinazioni di 5 elementi su 52).

4.4.1 Proprietà del coefficiente binomiale

1. La **prima proprietà** del coefficiente binomiale è di **simmetria**, ovvero

$$\binom{n}{k} = \frac{n!}{k| \cdot (n-k)!} = \binom{n}{n-k} = \frac{n!}{(n-k)! \cdot (n-(n-k))!} \qquad \forall n \in \mathbb{N}, \forall k \in \{0, ..., n\}$$

2. La **seconda proprietà** prende il nome di **Formula di Stifel** e permette di esprimere la somma tra due coefficienti binomiali, come mostrato di seguito, scrivendo esplicitamente i due coefficienti binomiali

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \frac{(n-1)!}{(k-1)! \cdot (n-k)!} + \frac{(n-1)!}{k! \cdot (n-k-1)!}$$

Ora, semplicemente, al fine di sommarli, è sufficiente impiegare le proprietà del fattoriale e scrivere

$$\frac{1}{(k-1)!} = \frac{k}{k!} \quad \text{e} \quad \frac{1}{(n-k-1)!} = \frac{(n-k)}{(n-k)!}$$

da cui si evince l'uguaglianza seguente

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

Ora che entrambi gli addendi presentano lo stesso denominatore, è possibile sommarli, ottenendo:

$$\frac{(n-1)! \cdot k}{k! \cdot (n-k)!} + \frac{(n-1)! \cdot (n-k)}{k! \cdot (n-k)!} = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k} \qquad \forall n \in \ltimes, \forall k \in \{0, ..., n\}$$

3. Infine, come **terza proprietà**, si espone la **formula del binomio di Newton**: siano $a, b \in \mathbb{R}$ e $n \in \mathbb{N}$, allora

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

In particolare è noto che il numero di tutti i possibili raggruppamenti di oggetti estratti da un insieme, ovvero il numero di tutti i sottoinsiemi di un insieme è

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

ponendo $a = b = 1$ nella formula del binomio di Newton.

4.5 Permutazione con ripetizione

Dopo aver introdotto il concetto di combinazione e di disposizione semplice e con ripetizione, é opportuno, ora, introdurre il significato della **permutazione con ripetizione**. Si consideri, a tal proposito, il seguente esempio: si supponga di disporre di 10 palline numerate e anche colorate (3 blue, 4 verdi e 3 rosse):

1 2 3 4 5 6 7 8 9 10

È possibile, ora, procedere all'ordinamento di tali oggetti, semplicemente provvedendo a permutarli, come mostrato di seguito

2 8 4 3 10 1 5 7 9 6

Questo cirrisponde alla permutazione

$$f : \{1, ..., 10\} \longrightarrow A$$

con A insieme delle palline, definita da

$$f(1) = 2 \;\; f(2) = 8 \;\; f(3) = 4 \;\; \text{e così via}$$

Oppure ancora

3 9 6 1 8 2 7 4 10 5

che corrisponde alla permutazione

$$g(1) = 3 \;\; g(2) = 9 \;\; g(3) = 6 \;\; \text{e così via}$$

In cui il numero delle permutazioni è, naturalmente, 10!, essendo una permutazione una disposizione semplice in cui $k = n$. Si osservi, però, che se nell'osservare le permutazioni si prende in considerazione esclusivamente il colore delle palline e non il loro valore, allora le due permutazioni precedenti sono perfettamente **equivalenti**. Formalmente, quindi, se l'insieme A viene diviso in 3 parti

$$\underbrace{A_1 = \{1, 2, 3\}}_{\text{BLU}} \quad \underbrace{A_2 = \{4, 5, 6, 7\}}_{\text{VERDE}} \quad \underbrace{A_3 = \{8, 9, 10\}}_{\text{ROSSO}}$$

e si osserva la controimmagine di ognuno di tali insimei attraverso f è g si ottiene

$$f^{-1}(A_1) = \{1, 3, 6\} = g^{-1}(A_1)$$
$$f^{-1}(A_2) = \{3, 7, 8, 10\} = g^{-1}(A_2)$$
$$f^{-1}(A_3) = \{2, 5, 9\} = g^{-1}(A_3)$$

cioé f e g spostano le palline di ogni colore sullo stesso blocco di posizioni.
Considerado, quindi, equivalenti due permutazioni che spostano i colori nello stesso modo, il numero di tali permutazioni non sarà più 10!, ma tale quantità dovrà essere divisa per tutte le permutazioni che si reputano identiche: quindi 3! nel caso del blu, 4! nel caso del verde e ancora 3! nel caso del rosso. Pertanto il numero di permutazioni che portano colori uguali in colori uguali, senza tenere conto dei è numeri è proprio

$$\frac{10!}{3! \cdot 4! \cdot 5!}$$

PERMUTAZIONE CON RIPETIZIONE

Dato un insieme finito A, sia $\{A_1, ..., A_q\}$ una sua partizione, cioé

$$\bigcup_{j=1}^{q} A_j = A$$

e gli A_j sono a due a due disgiunti. Si chiama **permutazione con ripetizione** una permutazione in cui si considerano uguali gli elementi di uno stesso sottoinsieme A_j. Ciò significa che si considerano equivalenti due permutazioni f e g su A quando

$$f^{-1}(A_j) = g^{-1}(A_j) \quad \forall j = 1, ..., q$$

ovvero le controimmagini di ciascuna permutazione coincidono.

Generalizzando, è possibile affermare che il numero di permutazioni con ripetuzione è

$$\boxed{\frac{n!}{\prod_{j=1}^{q} n_j!} \quad \text{dove} \quad \prod_{j=1}^{q} n_j! = (n_1)! \cdot ... \cdot (n_q)!}$$

in cui $n = |A|$, ossia il numero degli elementi dell'insieme A e $n_j = |A_j|$, ossia il numero degli elementi della partizione A_j. Questo è dovuto al fatto che permutando internamente gli elementi di uno stesso sottoinsieme A_j si ottiene una nuova permutazione, ma equivalente alla precedente.

Esempio: Sia $A = \{a, b, c, d\}$. Le possibili permutazioni di A sono $4! = 24$. Si considerino, ora, le permutazioni con ripetuzione relative alla partizione

$$A_1 = \{a\} \quad A_2 = \{b\} \quad A_3 = \{c, d\}$$

che, in pratica, equivale a non distinguere la lettera "*c*" dalla lettera "*d*"; pertanto

$$abcd \quad abdc \quad \text{sono equivalenti}$$

$$adbc \quad acbd \quad \text{sono equivalenti}$$

e così via. Il numero totale di permutazioni con ripetizione è, quindi

$$\frac{4!}{1! \cdot 1! \cdot 2!} = \frac{24}{2} = 12$$

4.6 Principio fondamentale del calcolo combinatorio

Si espone di seguito il **principio fondamentale del calcolo combinatorio**:

PRINICIPIO FONDAMENTALE DEL CALCOLO COMBINATORIO

Si supponga che gli elementi di una configurazione possano essere determinati mediante q scelte **successive** ed **indipendenti** (ovvero totalmente ininfluenti una sull'altra), in cui ogni scelta abbia un numero fissato di opzioni: la prima scelta ha n_1 opzioni, la seconda n_2 opzioni, ..., la q-esima scelta ha n_q opzioni.
Se sequenze distinte di opzioni nelle scelte determinano configurazioni distinte, allora il numero totale di configurazioni cche è possibile ottenere è pari a

$$\boxed{\prod_{j=1}^{q} n_q}$$

ossia dal prodotto delle possibili opzioni che si hanno a disposizione ad ogni scelta.

Esempio: Si supponga di entrare in una gelateria, in cui vi sono solamente tre gusti: cioccolato/crema/puffo. Considerando che è possibile scegliere anche tra cono e coppetta, si determinino quante configurazioni esistono per il proprio gelato se è possibile scegliere due palline.
Naturalmente si hanno $n_1 = 2$ opzioni per la base (cono/coppetta) e $n_2 = 6$ opzioni per i gusti (3 nel caso di palline uguali e 3 nel caso di palline diverse). Tenuto conto cche ad opzioni distinte corrispondono configurazioni finali distinte, si hanno $6 \cdot 2 = 12$ possibili assemblaggi.

Osservazione: Si osservi che il principio fondamentale del calcolo combinatorio prevede che le scelte da effettuare siano **indipendenti**, pertanto non è applicabile nel caso in cui da una scelta dipenda la scelta successiva.
Inoltre vi è il vincolo che **a scelte distinte corrispondano configurazioni finali distinte**: se, infatti, fosse sta individuata una prima scelta e una seconda scelta di gusto di gelato, si sarebbe potuto scegliere puffo e crema o crema e puffo ottenendo sempre la stessa configurazione finale, ma avendo prediletto opzioni distinte in principio.

Esercizio: Si impieghi, ora, il principio fondamentale del calcolo combinatorio per determinare il numero di possibili permutazioni con ripetizione, designato con c.
Si esamini, a tal proposito, il caso specifico delle palline colorate

2 8 4 3 10 1 5 7 9 6

e si supponga di voler "assemblare" una permutazione: è possibile decidere dapprima dove posizionare i colori e per farlo si dispone di c opzioni. Poi si deve decidere dove posizionare i numeri sulle palline blu, per cui si hanno $3!$ opzioni, sulle palline verdi, per cui si hanno $4!$ opzioni e sulle palline rosse, per cui si hanno $3!$ opzioni. Per il principio fondamentale del calcolo combinatorio, il numero di possibili permutazioni è adato da:

$$10! = c \cdot 3! \cdot 4! \cdot 3! \quad \text{e quindi} \quad c = \frac{10!}{3! \cdot 4! \cdot 3!}$$

Osservazione: Nel caso generale, per ottenere una permutazione dell'insieme A bisogna prima scegliere

$$\begin{cases} \text{le } n_1 \text{ posizioni su cui mettere gli elementi di } A_1 \\ \dots \\ \text{le } n_q \text{ posizioni su cui mettere gli elementi di } A_q \end{cases}$$

cioé si fissa una specifica permutazione con ripetizione; si dispone, poi

$$\begin{cases} \text{gli } n_1 \text{ elementi di } A_1 \text{ nelle posizioni prescelte} \\ \dots \\ \text{gli } n_q \text{ elementi di } A_q \text{ nelle posizioni prescelte} \end{cases}$$

Pertanto, nella prima scelta si hanno c possibili opzioni (ossia le permutazioni con ripetizione) e poi si hanno $n_1!$ opzioni nella seconda scelta (ossia le permutazioni su A_1), ..., ed infine si hanno $n_q!$ nell'ultima scelta (ossia le permutazioni su A_q). Pertanto si ha, per il principio fondamentale:

$$n! = c! \cdot n_1! \cdot \dots \cdot n_q! = c \cdot \prod_{j=1}^{q} n_j!$$

che restituisce il risultato già ottenuto

$$\boxed{\frac{n!}{\prod_{j=1}^{q} n_j!} \qquad \text{dove} \qquad \prod_{j=1}^{q} n_j! = (n_1)! \cdot \dots \cdot (n_q)!}$$

Osservazione: Si osservi che finora il concetto di disposizione con ripetizione, disposizione semplice e combinazione in modo matematico e preciso, mentre il concetto di permutazione con ripetizione è stato fornito in maniera intuitiva ed impirica.
Per essere più precisi, si rammenti che dato un insieme E, una **relazione** su E è detta di **equivalenza** se è

- **riflessiva**: $x \sim x \ \forall x \in E$
- **simmetrica**: $x \sim y$ allora $y \sim x \ \forall x, y \in E$
- **transitiva**: $x \sim y$ e $y \sim z$ allora $x \sim z \ \forall x, y, z \in E$

Due elementi x, y sono detti **equivalenti** se sono in relazione attraverso una relazione di equivalenza. La **classe di equivalenza** di un elemnto x è definita come

$$[x] := \{y \in E : x \sim y\}$$

Per la riflessività, $x \in [x]$; per la simmetria, se $y \in [x]$ allora anche $x \in [y]$, mentre per la transitività, se $y \in [x]$ allora $x \sim y$ e se $z \in [x]$ allora $z \sim x$ e quindi $z \sim y$, quindi $z \in [y]$, per ogni scelta di z (la stessa cosa vale se si parte considerando $z \in [y]$).
Dal primo risultato ($x \in [x]$), segue che ogni elementi di E apartiene alla propria classe di equivalenza, la famiglia delle classi di equivalenza copre completamente l'insieme E; non solo, ma se si considerano due classi di equivalenza, o queste coincidono o queste sono disgiunte, per cui le classi di equivalenza costituiscono una partizione di E.

Esempio 1: Sia $\mathbb{R} - \{0\}$ e si consideri la seguente relazione di equivalenza: $x \sim y$ se e solo se $x \cdot y > 0$ (cioé se x e y presentano lo stesso segno). È facile verificare che si tratta di una relazione di equivalenza, in quanto le classi di equivalenza sono due:

$$\{x \in \mathbb{R} : x > 0\} \quad \text{e} \quad \{x \in \mathbb{R} : x < 0\}$$

Esempio 2: Sia

$$A := \{f : [0, 1] \longrightarrow \mathbb{R}\}$$

e si ponga $f \sim g$ se e solo se $\exists c \in \mathbb{R} : f - g = c$ (cioé f e g sono una la traslata dell'altra in verticale). Anche in questo caso è facile verificare ce si tratta di una relazione di equivalenza.

Osservazione: Nella definzione di permutazione con ripetizione sono state considerate equivalenti due permutazioni f e g su A quando

$$f^{-1}(A_J) = g^{-1}(A_j) \quad \forall j = 1, ..., q$$

anche in questo caso si tratta di una relazione di equivalenza. Formalmente, dunque, **una permutazione con ripetizione è una classe di equivalenza** rispetto alla relazione

$$f^{-1}(A_J) = g^{-1}(A_j) \quad \forall j = 1, ..., q$$

nella famiglia delle permutazioni.

17 Marzo 2022

Dopo aver definito il concetto di permutazione con ripetizione rispetto ad una partizione, è necessario procedere ad osservare un'applicazione di tal concetto:

Esempio: Si supponga di disporre di 10 libri e di volerli disporre su una libreria avente 5 ripiani. Si vuole sapere in quanti modi è possibile farlo, senza tenere conto di quali siano i libri, ma solo di quanti ne si pone su ogni scaffale.
Per esempio, dati i 5 scaffali, vi si riporta quanti libri è stato deciso di porvisi, come mostrato di seguito

3	3	0	2	2

in cui, decidere di porre 3 manga al primo piano e i 3 libri di probabilità sul secondo o viceversa conta come una stessa configurazione.
Similmente, si sarebbe voluto sapere come suddividere 7 paia di calzini in 3 cassetti; se vengono disposti come segue

2	2	3

non è facile determinare il numero totale delle disposizioni. Più in generale, dati n oggetti e k contenitori, si calcoli il numero di possibili disposizioni.
Nel caso in cui $n = 10$ e $k = 5$, si considerino 10 tessere bianche (corrispondenti ai 10 libri) e $4 = k - 1$ tessere nere (corrispondenti ai divisori dei 5 scaffali) e si dispongano in fila come segue:

$$\underbrace{\square\square}_{2}\blacksquare\underbrace{\qquad}_{0}\blacksquare\underbrace{\square\square\square}_{3}\blacksquare\underbrace{\square\square}_{2}\blacksquare\underbrace{\square\square\square}_{3}$$

Naturalmente questa è una possibile configurazione: 2 tessere bianche prima della prima nera, 0 bianche tra la prima e la seconda nera, 3 bianche tra la seconda e la terza nera, 2 bianche tra la terza e la quarta nera e 3 bianche dopo la quarta nera. Si può associare, quindi, a questa configurazione di tessere la seguente configurazione di libri:

2	0	3	2	3

È ovvio che sussista una corrispondenza biunivoca tra le configurazioni delle tessere e le configurazioni dei libri. La configurazione di libri iniziale, per esempio, restituisce per le tessere la seguente:

$$\underbrace{\square\square\square}_{3}\blacksquare\underbrace{\square\square\square}_{3}\blacksquare\underbrace{\qquad}_{0}\blacksquare\underbrace{\square\square}_{2}\blacksquare\underbrace{\square\square}_{2}$$

Pertanto vi sono tante configurazioni di libri quante sono quelle per le tessere. Le configurazioni per le tessere sono permutazioni con ripetizione (rispetto alla partizione bianco/nero) e sono in numero di

$$\frac{14!}{10! \cdot 4!} = 1001$$

ovverosia 14 tessere, partizionate in 10 e 4. Più in generale, ripetendo lo stesso ragionamento con n oggetti e k contenitori, le configurazioni sono in numero uguale a quelle delle permutazione con ripetizione di $n + k - 1$ tessere partizionate in n e $k - 1$. Quindi

$$\boxed{\frac{(n+k-1)!}{n! \cdot (k-1)!} = \binom{n+k-1}{n}}$$

così facendo, nell'esempio dei calzini, con $n = 7$ e $k = 3$ si hanno

$$\frac{9!}{7! \cdot 2!} = 36$$

configurazioni. Da notare che se si considera un solo oggetto ($n = 1$), allora vi sono solo k configurazioni possibili, ovviamente.

Osservazione: Si osservi che i rudimenti di calcolo combinatorio fin ad ora analizzati (disposizioni semplici e con ripetizione, combinazioni, permutazioni con ripetzione rispetto ad una partizione) non sono semplici strumenti per risolvere problemi probabilistici, ma modi estremamente eficaci di contare possibili configurazioni.
Tale aspetto è essenziale quando si opera con probabilità finita, dal momento che è indispensabile per sapere quanti sono **tutti** i casi possibili e quanti sono quelli **favorevoli** (cioé quelli di interesse nell'evento considerato).

4.7 Combinazione con ripetizione

Si espone di seguito la definizione di **combinazione con ripetizione**:

COMBINAZIONE CON RIPETIZIONE

Dato $k \in \mathbb{N}$ e un insieme finito A, si chiama **combinazione con ripetizione** di classe k su A tutti i possibili raggruppamente di cardinalità k che si possono formare con gli elementi di A, includendo eventuali ripetizioni.
Da notare che due raggruppamenti vengono considerati differenti quando differiscono

- per qualche elemento;
- per il numero di volte in cui un dato elemento viene ripetuto.

Esempio: Sia $A = \{1, 2, 3, 4, 5\}$ e si determinino le combinazioni con 2 elementi e quali quelle con ripetizione. Naturalmente le combinazioni con 2 elementi sono:

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

ovvero sono

$$C_{5,2} = \binom{5}{2} = \frac{5!}{2! \cdot 3!} = 10$$

Se, invece, si possono ripetere qualche elemento, allora vanno anche considerate, oltre a quelle già esposte, le seguenti:

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

Allora, per determinare il numero di combinazioni con ripetizione di classe k su un insieme finito A di cardinalità n si consideri che gli elementi di $A = \{a_1, ..., a_n\}$ siano delle scatole, come mostrato di seguito:

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	$\cdots$	a_n

Si considerino, ora, k oggetti e si pongano questi ultimi all'interno delle scatole. Infine si formi una combinazione con ripetizione con gli elementi a_j che come scatola contengono almeno un ogetto, e si ripetano tante volte quanti sono gli oggetti che stanno nella scatola.
Per esempio, con $k = 5$ si possono porre 2 oggetti nella scatola a_3, 1 nella scatola a_5 e due nalla scatola a_8. La combinazione relativa sarà

$$\{a_3, a_3, a_5, a_8, a_8\}$$

Si è così creata una corrispondenza biunivoca tra le combinazioni con ripetizione di classe k su n elementi e i modi di mettere k oggetti in n contenitori. Detto $C^r_{n,k}$ il numero delle combinazioni con ripetizione, per quanto visto sui contenitori, è possibile affermare che

$$\boxed{C^r_{n,k} = \binom{n+k-1}{k}}$$

e si noti che sono k oggetti in n contenitori e non viceversa.

Esempio: Si scrivano tutte le combinazioni con ripetizione di classe 3 sull'insieme $A = \{a, b, c\}$. Allora si hanno:

$$\{a,a,a\}\ ,\ \{a,a,b\}\ ,\ \{a,a,c\}\ ,\ \{a,b,b\}\ ,\ \{a,b,c\}\ ,\ \{a,c,c\}\ ,\ \{b,b,b\}\ ,\ \{b,b,c\}\ ,\ \{b,c,c\}\ ,\ \{c,c,c\}$$

Nello scrivere aiuta sapere in anticipo che sono

$$C^r_{5,3} = \binom{5}{3} = \frac{5!}{3! \cdot 2!} = 10$$

Osservazione: Pertanto, si hanno le seguenti formule per calcolare il numero di configurazione di classe k di n oggetti:

CONFIGURAZIONI DI CLASSE k SU n OGGETTI

	SEMPLICI	CON RIPETIZIONE
DISPOSIZIONI (configurazioni ordinate)	$\frac{n!}{(n-k)!}$	n^k
COMBINAZIONI (configurazioni non ordinate)	$\binom{n}{k}$	$\binom{n+k-1}{k}$

4.8 Estrazione

Si supponga di disporre di un'urna contenente N palline di due colori: si considerino M palline bianche e $N - M$ palline nere (con $M \leq N$). Si supponga di eseguire n estrazioni successive. Naturalmente si hanno due cateogorie

- **Estrazione senza reimmissione**, cioé le palline estratte non vengono reinserite. In questo caso deve essere $n \leq N$;
- **Estrazione con reimmissione**, cioé ad ogni estrazione la pallina estratta viene reinserita nell'urna.

Per ciascuna delle due cateogorie bisogna calcolare la probabilità che k delle n palline estratte siano bianche.
Numerando da 1 ad M le palline bianche e da $M + 1$ a N le palline nere si consideri la prima cateogoria: l'evento "il numero di palline bianche estratte è k" non dipende dall'ordine di estrazione, quindi è possibile prendere

$$\Omega := \{\text{combinazioni di classe } n \text{ sull'insieme delle } N \text{ palline}\}$$

Da questo è immediato capire che

$$|\Omega| = \binom{N}{n}$$

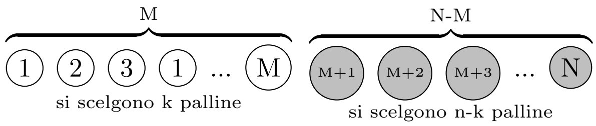
L'evento favorevole è

$$A := \{\omega \in \Omega : |\omega \cap \{1, ..., M\}| = k\} = \{\omega \in \Omega : |\omega \cap \{\omega \cap \{M + 1, ..., N\}| = n - k\}$$

cioé le combinazioni che hanno k elementi bianchi sono uguali alle combinazioni che hanno $n - k$ elementi neri.
È possibile determinare una configurazione $\omega \in A$ attraverso due scelte:

- scegliere la parte di ω in $\{1, ..., M\}$, cioé una combinazione di classe k. Naturalmente in questo caso si hanno
$$\binom{M}{k}$$
opzioni.

- Scegliere la parte di ω in $\{M+1, ..., N\}$, cioè una combinazione di classe $n-k$. Pertanto si hanno
$$\binom{N-M}{n-k}$$
opzioni.



In questo modo si ottengono tutte le possibili configurazioni. Per il **teorema fondamentale del calcolo combinatorio**, il loro numero totale è dato da

$$|A| = \binom{M}{k} \cdot \binom{N-M}{n-k}$$

Una volta che è nota la cardinalità di Ω e di A, si può procedere ad assegnare le probabilità elementari: ancora una volta, si osserva che tale compito è affidato alla statistica, in quanto di natura prettamente euristica. Tuttavia, una configurazione significa infilare una mano nell'urna e tirare fuori n palline: è quindi naturale assumere che le configurazioni abbiano tutte la stessa probabilità.
La probabilità elementare sulla singola configurazione è data, ovviamente, da

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

in quanto la probabilità totale dello spazio Ω deve essere pari a 1. La probabilità dell'evento A è data, invece, da

$$p(A) = \frac{|A|}{|\Omega|} = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

Osservazione: Si osservi che non sarebbe cambiato nulla se invece di considerare le combinazioni fossero state considerate delle disposizioni semplici. In altre parole, se invece di estrare dall'urna un mucchio di n palline se ne fosse estratta una per volta e messa in ordine in fila il risultato sarebbe stato identico. Tuttavia, il problema è che ad ogni combinazione di classe n corrispondono $n!$ disposizioni semplici di classe n, per cui il problema si sarebbe inutilmente complicato.
Quindi prendendo come spazio degli eventi elementari

$$\hat{\Omega} = \{\text{disposizioni semplici di classe } n \text{ sull'insieme delle } N \text{ palline}\}$$

ed indicato con $\hat{A}$ l'evento favorevole, si hanno $|\hat{\Omega}| = n! \cdot |\Omega|$ e $\hat{A} = n! \cdot |A|$, cossicché

$$p(\hat{A}) = \frac{|\hat{A}|}{|\hat{\Omega}|} = \frac{|A|}{|\Omega|} = p(A)$$

Osservazione: Si proceda, ora, all'analisi dell'estrazione con reimmissione di 3 palline di un'urna con 3 palline. È possibile estrarre in modo ordinato, per cui vi sono $3^3 = 27$ disposizioni con ripetizione (di classe 3 su 3 oggetti):

$$(1,1,1)\ ,\ (2,2,2)\ ,\ (3,3,3)$$

$$\begin{gathered}(1,1,2)\ ,\ (1,2,1)\ ,\ (2,1,1)\ ,\ (1,1,3)\ ,\ (1,3,1)\ ,\ (3,1,1)\\ (2,2,1)\ ,\ (2,1,2)\ ,\ (1,2,2)\ ,\ (2,2,3)\ ,\ (2,3,2)\ ,\ (3,2,2)\\ (3,3,1)\ ,\ (3,1,3)\ ,\ (1,3,3)\ ,\ (3,3,2)\ ,\ (3,2,3)\ ,\ (2,3,3)\end{gathered}$$

$$(1,2,3)\ ,\ (1,3,2)\ ,\ (2,1,3)\ ,\ (2,3,1)\ ,\ (3,1,2)\ ,\ (3,2,1)$$

che sono state opportunamente suddivise in classi: il primo gruppo sono le disposizioni con 3 ripetizioni, a ciascuna delle quali corrisponde una combinazione, il secondo gruppo sono disposizioni con 2 ripetizioni che, a tre a tre, corrispondono ad una sola combinazione, mentre il terzo gruppo contiene sei disposizioni con 3 ripetizioni, corrispondenti ad una sola combinazione.
Alternativamente è possibile solamente segnare i risultati senza tenere conto dell'ordine di estrazione. In questo secondo caso si avrebbero

$$\binom{5}{3} = 10$$

combinazioni con ripetizione (di classe 3 su 3 oggetti):

$$\{1,1,1\}\ ,\ \{2,2,2\}\ ,\ \{3,3,3\}$$

$$\{1,1,2\}\ ,\ \{1,1,3\}\ ,\ \{2,2,1\}\ ,\ \{2,2,3\}\ ,\ \{3,3,1\}\ ,\ \{3,3,2\}$$

$$\{1,2,3\}$$

Come si osserva, nel caso di estrazioni con reimmissione, la cardinalità della corrispondenza tra disposizioni e combinazioni varia: 1 a 1, 3 a 1 e 6 a 1 (mentre nel caso di estrazioni senza reimmissione, la corrispondenza era costante: a 1 combinazione corrispondevano $n!$ disposizioni). Poichè sull'estrazione singola le tre palline sono equiprobabili, ammesso che siano palline uguali, appare ragionevole assegnare a tutte le disposizioni con ripetizione la stessa probabilità, mentre alle combinazioni con ripetizione l'assegnazione va, invece, pesata opportunamente.
Si esamini in senso generale la seconda categoria di estrazioni: estrazione con reimmissione. Sulla base di quanto appena detto, è meglio considerare anche l'ordine di estrazione e prendere

$$\Omega := \{\text{disposizioni con ripetizione di classe } n \text{ sull'insieme delle } N \text{ palline}\}$$

Naturalmene si considerano disposizioni con ripetizione perché le palline una volta estratte vengono reinserite. Per cui si ha che

$$|\Omega| = N^n$$

L'evento favorevole è

$$A := \{f : \{1, ..., n\} \longrightarrow \{1, ..., N\} \text{ tale che } \left|f^{-1}\left(\{1, ..., M\}\right)\right| = k\}$$

cioé le disposizioni con ripetizione che pescano k elementi bianchi. Per determinare $|A|$ si utilizza il principio fondamentale in maniera simile a come è stato usato per le permutazioni con ripetizione con ripetizione rispetto ad una partizione.
Per ottenere una disposizione f in A

- si scelgono le k posizioni, su n possibili, in cui mettere le palline bianche. Naturalmente si hanno

  $$\binom{n}{k}$$

  opzioni, essendo la scelta una combinazione di classe k su n elementi.

- si dispongono k palline bianche (eventualmente con ripetizione) nelle posizioni prescelte. Naturalmente si hanno M^k opzioni, essendo la scelta una disposizione.

- si dispongono $n-k$ palline nere (eventualmente con ripetizione) nelle restanti $n-k$ posizioni. Naturalmente si hanno $(N-M)^{n-k}$ opzioni, essendo la scelta una disposizione.

In questo modo si ottengono tutte le possibile configurazioni. Per il principio fondamentale il loro numero totale è dato da

$$|A| = \binom{n}{k} \cdot M^k \cdot (N-M)^{n-k}$$

È possibile, quindi, concludere che

$$\boxed{p(A) = \frac{|A|}{|\Omega|} = \binom{n}{k} \cdot \frac{M^k \cdot (N-M)^{n-k}}{N^n} = \binom{n}{k} \cdot S^k \cdot (1-S)^{n-k}}$$

ponendo

$$S = \frac{M}{N}$$