Un esperimento si dice <u>aleatorio</u> se non siamo in grado di prevederne il risultato. Si chiamano <u>eventualità</u> i possibili risultati di un esperimento. L'insieme di tutte le eventualità è Ω.

<u>Esempi</u> (1) Lancio di un dado: $\Omega = \{1,2,3,4,5,6\}$.

(2) Estrazione del lotto: $\Omega = \{(n_1, n_2, n_3, n_4, n_5) : n_i \in \mathbb{N},\ 1 \leq n_i \leq 90,\ n_i \neq n_j \text{ per } i \neq j\}$.

(3) Svuotamento di un'urna piena di palline: $\Omega = \mathbb{N}$.
L'esperimento in questo caso non è ripetibile!

(4) Corsa di cavalli: se i concorrenti sono 9, $\Omega = \{$possibili ordini d'arrivo$\} = \{$permutazioni su 9 elementi$\}$.

Se ci interessa non il singolo risultato dell'esperimento, ma il fatto che si avveri un certo evento legato ad esso, o un certo insieme di risultati, saremo indotti a studiare un opportuno sottoinsieme $A \subseteq \Omega$. L'<u>evento</u> A si realizza se e solo se il risultato dell'esperimento "cade in A", cioè appartiene ad A.

<u>Esempi</u> (1) Uscita di una faccia pari: $A = \{2,4,6\}$.

(2) Uscita del 17: $A = \{(n_1, n_2, n_3, n_4, n_5) : \exists i : n_i = 17\}$.

(4) Corsa dei cavalli: se il cavallo su cui vogliamo puntare ha il numero 5, sarà $A = \{$permutazioni che scambiano 1 con 5$\}$.

In generale, l'insieme degli eventi è $\mathcal{P}(\Omega) = \{A : A \subseteq \Omega\}$.

Talvolta però è utile considerare una classe più ristretta di eventi, $\mathcal{A} \subset \mathcal{P}(\Omega)$, quando certi eventi non si ritengono interessanti per i nostri scopi. Però è opportuno che $\mathcal{A}$ sia chiusa per unioni e intersezioni (finite o numerabili) e per passaggio al complementare: cioè $\mathcal{A}$ deve essere una tribù $\subseteq \mathcal{P}(\Omega)$.

La coppia $(\Omega, \mathcal{A})$ si chiama spazio probabilizzabile, Ω è l'insieme delle eventualità, $\mathcal{A}$ è la tribù degli eventi.

Se A è un evento, A^c è l'evento contrario;

Se $\omega \in A$, ω realizza l'evento A;

$A \cup B$ è l'evento "A o B";

$A \cap B$ è l'evento "A e B"

A e B si dicono incompatibili se $A \cap B = \phi$.

Esempio Scegliamo a caso un punto in un segmento I. Con un'opportuna unità di misura possiamo supporre $I = [0,1]$. Se ci interessano gli eventi del tipo "il punto scelto cade fra a e b", dovremo inserire ogni $[a,b] \subseteq [0,1]$ fra gli eventi. La minima tribù che contiene tutti questi intervalli è la tribù boreliana $\mathcal{B}[0,1]$. Dunque $\Omega = [0,1]$, $\mathcal{A} = \mathcal{B}[0,1]$.

Adesso si deve scegliere una misura di probabilità, cioè una funzione che misuri il nostro grado di fiducia nel realizzarsi di un certo evento, dunque una funzione $P : \mathcal{A} \to [0,1]$ (per convenzione il "grado di fiducia" è un numero in percentuale, ovvero una quantità fra 0 e 1).

Definizione Una misura di probabilità su $(\Omega, \mathcal{A})$ è una funzione $P : \mathcal{A} \to [0,1]$ tale che

(i) $P(\phi) = 0$, $P(\Omega) = 1$,

(ii) $P(A \cup B) = P(A) + P(B)$ per ogni $A, B \in \mathcal{A}$ che siano fra loro incompatibili, e più in generale

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \quad \text{per ogni } \{A_n\}_{n \in \mathbb{N}^+} \subseteq \mathcal{A}$$

con gli A_n a due a due incompatibili.

(additività)

Si dice trascurabile ogni evento A tale che $P(A) = 0$; si dice quasi certo ogni evento A tale che $P(A) = 1$.

Si noti che esistono varie misure di probabilità sullo stesso spazio $(\Omega, \mathcal{A})$. La scelta di una P può essere anche soggettiva. Tuttavia in molti casi ci sono condizioni naturali che riducono a una scelta univoca della P.

La terna $(\Omega, \mathcal{A}, P)$ è detta spazio probabilizzato. La scelta di esso è il primo passo per studiare un esperimento aleatorio.

Proprietà generali degli spazi probabilizzati

Per la proprietà di additività, se $A, B \in \mathcal{A}$

$$(*) \quad P(B) = P(B \cap A) + P(B \cap A^c)$$

perché $B \cap A$ e $B \cap A^c$ sono incompatibili. Con $B = \Omega$,

$$1 = P(\Omega) = P(A) + P(A^c), \text{ ossia } P(A^c) = 1 - P(A).$$

Se poi $A \subseteq B$,

$$P(B) = P(A) + P(B \cap A^c) = P(A) + P(B \setminus A),$$

da cui

$$P(B \setminus A) = P(B) - P(A), \text{ se } A \subseteq B,$$

e in particolare

$$P(A) \leq P(B) \text{ se } A \subseteq B \quad (\underline{\text{isotonia}} \text{ di } P)$$

Poi, scrivendo la simmetrica di $(*)$

$$P(A) = P(B \cap A) + P(B^c \cap A),$$

e sommando questa alla $(*)$,

$$P(A) + P(B) = 2\,P(A \cap B) + P(B \cap A^c) + P(B^c \cap A).$$

Poiché $A \cap B$, $B \cap A^c$ e $B^c \cap A$ sono incompatibili, e la loro unione è $A \cup B$,

$$P(A) + P(B) = P(A \cap B) + P(A \cup B) \quad (\underline{\text{modularità}} \text{ di } P),$$

che generalizza l'additività quando A e B non sono incompatibili.

Se $\{A_n\}_{n\in\mathbb{N}^+} \subseteq \mathcal{A}$, si ha poi

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1 - P\left(\bigcap_{n=1}^{\infty} A_n^c\right),$$

formula utile perché è spesso più facile calcolare la probabilità di un'intersezione.

Inoltre:

<u>Proposizione</u> Se $A_n \subseteq A_{n+1}$ per ogni $n\in\mathbb{N}^+$, allora

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right);$$

se $A_n \supseteq A_{n+1}$ per ogni $n\in\mathbb{N}^+$, allora

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right). \quad \square$$

<u>SPAZI PROBABILIZZATI DISCRETI</u>

Supponiamo Ω finito o numerabile, dunque $\Omega \subseteq \mathbb{N}^+$, e $\mathcal{A} = \mathcal{P}(\Omega)$. Si dice allora che $(\Omega, \mathcal{A}, P)$ è <u>discreto</u>.

La classe di tutte le probabilità $P: \mathcal{A} \to [0,1]$ è in corrispondenza biunivoca con l'insieme di tutte le funzioni (successioni) $f: \Omega \to [0,1]$ tali che $\sum_{n\in\Omega} f(n) = 1$.

Infatti, fissata P, basta porre $f(n) = P(\{n\})$ per ogni $n\in\Omega$; fissata f, basta porre $P(A) = \sum_{n\in A} f(n)$ per ogni $A\in\mathcal{A}$. La funzione f associata a P si dice <u>densità discreta</u> di P.

Supponiamo Ω finito, con $\#\Omega = n$ (il simbolo $\#\Omega$ denota la cardinalità, ossia il numero di elementi, di Ω).
Se si ritiene che tutte le eventualità di Ω siano equi-probabili, allora è naturale scegliere una densità f costante:
$f(k) = \frac{1}{n}$ per ogni $k \in \Omega$ (visto che deve essere $\sum_{k \in \Omega} f(k) = 1$).
La corrispondente misura di probabilità P si dice ripartizione uniforme: se $A \in \mathcal{A}$ e $\#A = m$, allora si ha

$$P(A) = \sum_{m \in A} f(m) = \sum_{m \in A} \frac{1}{n} = \frac{m}{n} = \frac{\#A}{\#\Omega}.$$

Dunque $P(A)$ è il rapporto fra i casi favorevoli ed i casi possibili.

Esempio Se si lancia un dado equilibrato, allora
$\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $P(\{m\}) = \frac{1}{6}$ $\forall m \in \Omega$.

Esempio Non sempre conviene scegliere la ripartizione uniforme: se a una corsa di cavalli con 10 partenti ce n'è uno nettamente più forte, che ha il numero 5, converrà attribuire alla sua vittoria una probabilità maggiore. Se puntiamo solo sul vincente, avremo $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ e potremmo ad esempio scegliere

$$P(5) = \frac{1}{2}, \quad P(m) = \frac{1}{18} \quad \forall m \in \Omega \setminus \{5\}.$$