La probabilità in ventiquattr'ore

Appunti delle lezioni tenute dal Prof. A. Fonda

Università di Trieste, CdL Informatica, a.a. 2009/2010

1^a e 2^a ora di lezione

Cercheremo di introdurre il concetto di "probabilità" con alcuni esempi.

Primo esempio. Cominciamo con il lancio di un dado. Consideriamo uno dei seguenti possibili "eventi":

1) esce il 6;
2) esce un numero pari;
3) esce un numero dispari;
4) esce un numero ≤ 4;
5) esce un numero ≥ 7;
6) esce un numero ≤ 6.

Notiamo che i possibili risultati di un lancio costituiscono l'insieme

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

Ad ogni evento corrisponde un sottoinsieme A di Ω. Precisamente:

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

Ad ognuno di questi sottoinsiemi A possiamo associare una "probabilità" $P(A)$: si tratta di un numero reale, compreso tra 0 e 1, che dovrebbe servire a misurare il grado di fiducia che assegnamo a ciacun evento. Notiamo che il caso 5) corrisponde all'"evento impossibile" e il caso 6) corrisponde all'"evento certo". In questi casi si pone, per convenzione,

$$P(\emptyset) = 0\,, \qquad P(\Omega) = 1\,.$$

Risulta inoltre naturale richiedere che, se A e B sono due sottoinsiemi disgiunti di Ω, allora

$$P(A \cup B) = P(A) + P(B)\,.$$

Questa sarà chiamata "proprietà di additività".

Supponiamo che il nostro dado non sia truccato e che il lancio venga effettuato senza che alcuno dei sei numeri sia privilegiato rispetto agli altri. Ecco allora che, guardando i due casi complementari 2) e 3), non essendoci motivo di credere che uno dei due sia più "probabile" dell'altro, ed essendo la loro unione tutto Ω, possiamo concludere che

$$P(\{2,4,6\}) = P(\{1,3,5\}) = \tfrac{1}{2}\,.$$

Affrontiamo ora il caso 1). Possiamo osservare che, essendo

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

ognuno dei sei eventi essendo ugualmente "probabile", dovrà essere

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \tfrac{1}{6}\,.$$

Resta infine da analizzare il caso 4). Essendo

$$\{1,2,3,4\} = \{1\} \cup \{2\} \cup \{3\} \cup \{4\}\,,$$

ne deduciamo che

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

Secondo esempio. Consideriamo ora il lancio di due dadi. I possibili risultati formano l'insieme

$$\begin{aligned}\Omega = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}\,.\end{aligned}$$

Cerchiamo di determinare le probabilità dei seguenti eventi:

1) esce un doppio 6;
2) escono due numeri uguali;
3) la somma dei due numeri fa 8;
4) almeno uno dei due numeri è pari.

Per quanto riguarda l'evento 1), con semplici considerazioni, similmente a quanto visto nel primo esempio, vediamo che, siccome l'insieme Ω è costituito da 36 elementi,

$$P(\{(6,6)\}) = \tfrac{1}{36}\,.$$

All'evento 2) è associato l'insieme

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

costituito da sei elementi. Quindi, non essendoci alcuno che sia più "probabile" di un altro,

$$P(A) = P(\{(1,1)\}) + P(\{(2,2)\}) + P(\{(3,3)\}) + \\ +P(\{(4,4)\}) + P(\{(5,5)\}) + P(\{(6,6)\}) = \tfrac{6}{36}\,.$$

Per quanto riguarda l'evento 3), abbiamo l'insieme

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

costituito da cinque elementi, per cui $P(A) = \frac{5}{36}$. Infine, per l'evento 4) abbiamo l'insieme

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

costituito da 27 elementi, per cui $P(A) = \frac{27}{36}$.

Terzo esempio. Scegliamo "a caso" un punto nel quadrato

$$\Omega = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1\,,\ 0 \leq y \leq 1\}\,.$$

Qual è la probabilità che questo punto stia anche nel cerchio

$$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}\,?$$

Una breve riflessione fa pensare che la risposta giusta venga dalla valutazione dell'area della regione

$$A = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1\,,\ 0 \leq y \leq 1\,,\ x^2 + y^2 \leq 1\}\,,$$

in rapporto con l'area totale del quadrato, che vale 1. Proponiamo quindi la risposta:

$$P(A) = \tfrac{\pi}{4}\,.$$

3^a ora di lezione

Prendiamo ora due sottoinsiemi A e B di Ω, non necessariamente disgiunti. Scriviamo

$$A = (A \backslash B) \cup (A \cap B)\,, \qquad B = (B \backslash A) \cup (A \cap B)\,,$$

$$A \cup B = (A \backslash B) \cup (A \cap B) \cup (B \backslash A)\,,$$

in cui tutti gli insiemi tra parentesi risultano disgiunti. Usando la proprietà di additività, si ha

$$P(A) = P(A \backslash B) + P(A \cap B)\,, \qquad P(B) = P(B \backslash A) + P(A \cap B)\,,$$

$$P(A \cup B) = P(A \backslash B) + P(A \cap B) + P(B \backslash A) ,$$

e si trova la formula generale

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

Esempio. Consideriamo l'estrazione di un numero della tombola. Come è noto, i numeri possibili sono 90, per cui

$$\Omega = \{1, 2, 3, \ldots, 90\} .$$

Prendiamo in esame i due eventi

$$A = \{j \in \Omega : j \text{ è pari}\} , \qquad B = \{j \in \Omega : j \text{ è multiplo di tre}\} .$$

Troviamo facilmente che

$$P(A) = \tfrac{45}{90} = \tfrac{1}{2} , \qquad P(B) = \tfrac{30}{90} = \tfrac{1}{3} .$$

Notiamo che

$$A \cap B = \{j \in \Omega : j \text{ è multiplo di sei}\} ,$$

per cui

$$P(A \cap B) = \tfrac{15}{90} = \tfrac{1}{6} .$$

Ne segue che

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{6} = \tfrac{2}{3} .$$

Esempio modificato. Consideriamo lo stesso problema, nel caso i cui ci siano 100 numeri invece di 90. Abbiamo quindi

$$\Omega = \{1, 2, 3, \ldots, 100\} .$$

Vediamo che, in questo caso, le probabilità cambiano. Avremo infatti

$$P(A) = \tfrac{50}{100} , \qquad P(B) = \tfrac{33}{100} , \qquad P(A \cap B) = \tfrac{16}{100} ,$$

da cui

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \tfrac{50+33-16}{100} = \tfrac{67}{100} \neq \tfrac{2}{3} .$$

Introduciamo ora l'insieme complementare

$$\mathcal{C}A = \Omega \backslash A = \{j \in \Omega : j \notin A\} .$$

Essendo A e $\mathcal{C}A$ disgiunti e $A \cup \mathcal{C}A = \Omega$, per la proprietà di additività si ha che $P(A) + P(\mathcal{C}A) = P(\Omega)$, ed essendo $P(\Omega) = 1$, si ha la formula

$$P(\mathcal{C}A) = 1 - P(A) .$$

Esempio. Lanciamo un dado e ci chiediamo qual è la probabilità che non esca il numero 4. Sappiamo che $P(\{4\}) = \frac{1}{6}$, quindi

$$P(\mathcal{C}\{4\}) = 1 - \tfrac{1}{6} = \tfrac{5}{6}\,.$$

Si noti che avremmo anche potuto calcolare direttamente la probabilità

$$P(\{1,2,3,5,6\}) = \tfrac{5}{6}\,.$$

Dagli esempi precedenti, in cui tutti i risultati possibili sono "equiprobabili" e in numero finito, risulta chiaro che la formula per calcolare la probabilità di un sottoinsieme A di Ω è:

$$P(A) = \frac{\text{numero degli elementi di } A}{\text{numero degli elementi di } \Omega}\,.$$

In questi casi, il calcolo della probabilità si riduce al contare il numero di elementi di certi insiemi, cosa non sempre semplicissima. Ne vedremo alcuni esempi.

Lanci di una moneta. Consideriamo il seguente problema. Lanciamo dieci volte una moneta e ci chiediamo: qual è la probabilità di ottenere cinque volte testa e cinque volte croce?

Inizieremo considerando un problema semplificato. Lanciamo quattro volte una moneta e ci chiediamo: qual è la probabilità di ottenere due volte testa e due volte croce? Indicando con T il risultato "Testa" e con C il rsultato "Croce", abbiamo

$$\begin{aligned}\Omega = \{\, &(TTTT),(TTTC),(TTCT),(TTCC),\\ &(TCTT),(TCTC),(TCCT),(TCCC),\\ &(CTTT),(CTTC),(CTCT),(CTCC),\\ &(CCTT),(CCTC),(CCCT),(CCCC)\}\,.\end{aligned}$$

L'insieme che ci interessa è

$$A = \{(TTCC),(TCTC),(TCCT),(CTTC),(CTCT),(CCTT)\}\,.$$

La probabilità è quindi

$$P(A) = \tfrac{6}{16} = \tfrac{3}{8}\,.$$

Analizzando quanto fatto sopra, ci accorgiamo che l'insieme Ω è costituito dalle stringhe di quattro elementi, ognuno dei quali può assumere uno dei due valori, T o C. Esso ha quindi $2^4 = 16$ elementi. L'insieme A si ottiene invece scegliendo due T in una stringa di quattro elementi, per cui i rimanenti risultano automaticamente due C. Il numero di elementi di A corrisponde quindi al numero di sottoinsiemi di due elementi scelti da un inseme di quattro elementi.

4^a e 5^a ora di lezione

In generale ricordiamo che, dati due numeri naturali n e k, con $k \leq n$, il numero di sottoinsiemi di k elementi scelti da un insieme di n elementi è dato dal "coefficiente binomiale"

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

In particolare, come abbiamo visto sopra, il numero di elementi di A è $\binom{4}{2} = 6$.

Torniamo ora al nostro problema iniziale. In dieci lanci, l'insieme Ω sarà costituito dalle stringhe di dieci elementi, ognuno dei quali può assumere uno dei due valori, T o C. Esso ha quindi $2^{10} = 1024$ elementi. L'insieme A si ottiene scegliendo cinque T in una stringa di dieci elementi. Esso ha quindi $\binom{10}{5} = 252$ elementi. La probabilità cercata è quindi

$$P(A) = \tfrac{252}{1024} = \tfrac{63}{256}\,.$$

Estrazioni da un'urna. Ci sono quattro palline in un'urna, numerate da 1 a 4. Ne estraiamo due, contemporaneamente. Qual è la probabilità di trovare la numero 3?

Vediamo che

$$\begin{aligned}\Omega = \{&(1,2),(1,3),(1,4),(2,1),(2,3),(2,4),\\ &(3,1),(3,2),(3,4),(4,1),(4,2),(4,3)\}\,,\end{aligned}$$

mentre l'insieme che ci interessa è

$$A = \{(1,3),(2,3),(3,1),(3,2),(3,4),(4,3)\}\,.$$

Quindi, $P(A) = \frac{6}{12} = \frac{1}{2}$. Si noti che lo stesso risultato può essere ottenuto con un ragionamento più semplice: una volta effettuata l'estrazione delle due palline, la numero 3 può essere tra le due estratte o tra le due rimaste nell'urna, con uguale probabilità...

Supponiamo ora che, invece di estrarre le due palline assieme, ne estraiamo una, prendiamo nota del suo numero, la rimettiamo nell'urna ed estraiamo di nuovo. Qual è la probabilità che in almeno una delle due estrazioni sia stata trovata la numero 3?

Qui abbiamo che

$$\begin{aligned}\Omega = \{&(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(2,4),\\ &(3,1),(3,2),(3,3),(3,4),(4,1),(4,2),(4,3),(4,4)\}\,,\end{aligned}$$

mentre l'insieme che ci interessa è

$$A = \{(1,3),(2,3),(3,1),(3,2),(3,3),(3,4),(4,3)\}\,.$$

Quindi, $P(A) = \frac{7}{16}$.

I due problemi considerati differiscono per il fatto che, nel primo si sono effettuate due estrazioni "senza reimbussolamento", mentre nel secondo esse sono state eseguite "con reimbussolamento". La differenza sarebbe stata ancora più evidente se, invece di due estrazioni, se ne fossero effettuate quattro. Nel caso senza reimbussolamento, la probabilità di trovare la numero 3 è uguale a 1 (evento certo). Nel caso con reimbussolamento, l'insieme Ω, costituito da tutte le quaterne di numeri da 1 a 4, ha $4^4 = 256$ elementi, mentre un computo accurato mostra che tra questi ce ne sono esattamente 175 che contengono il numero 3. La probabilità di trovare la numero 3 almeno una volta è quindi uguale a $\frac{175}{256}$.

Si noti che il problema dei lanci di una moneta è equivalente al problema delle estrazioni con reimbussolamento da un'urna contenente due palline, che possiamo marcare con T e C.

Il problema delle estrazioni da un'urna verrà ripreso in seguito, in situazioni di maggiore generalità.

Sia $P(A)$ la probabilità di un evento A come sottoinsieme di Ω. Supponiamo che, ad un tratto, veniamo a conoscenza del fatto che i possibili risultati staranno sicuramente in un insieme più piccolo, un sottoinsieme $\tilde{\Omega}$ di Ω, e che A sia contenuto in $\tilde{\Omega}$. Ecco allora che A ha una nuova probabilità, come sottoinsieme di $\tilde{\Omega}$, che denoteremo con $\tilde{P}(A)$. Come calcolarla? Se pensiamo a:

$$\begin{array}{ll} P(A), & \text{come “misura” di } A \text{ rispetto a } \Omega, \\ \tilde{P}(A), & \text{come “misura” di } A \text{ rispetto a } \tilde{\Omega}, \\ P(\tilde{\Omega}), & \text{come “misura” di } \tilde{\Omega} \text{ rispetto a } \Omega, \end{array}$$

ricordando che $P(\Omega) = 1$, si deve avere

$$\tilde{P}(A) = \frac{P(A)}{P(\tilde{\Omega})}\,.$$

Esempio. Supponiamo di avere davanti a noi cinque scatole, di cui una sola contiene un'ambita vincita. Scegliendo a caso, avremo quindi una probabilità di $\frac{1}{5}$ di vincere. Numerando le cinque scatole, si ha $\Omega = \{1, 2, 3, 4, 5\}$. Supponiamo ora che un informatore fidato ci dica che la vincita non si trova nè nella prima nè nell'ultima scatola. Ecco allora che restano tre possibili scatole in cui si trova la vincita. La nuova probabilità di vincere è ora $\frac{1}{3}$. Infatti, il nuovo insieme in cui cercare la scatola vincente è $\tilde{\Omega} = \{2, 3, 4\}$, e

$$\tilde{P}(A) = \frac{P(A)}{P(\tilde{\Omega})} = \frac{\frac{1}{5}}{\frac{3}{5}} = \frac{1}{3}\,.$$

Spesso, invece di scrivere $\tilde{P}(A)$, si scrive

$$P(A \,|\, \tilde{\Omega})\,,$$

e si legge "probabilità (condizionale) di A rispetto a $\tilde{\Omega}$".

6^a ora di lezione

Un altro esempio. Supponiamo di sapere che un ragazzo di 20 anni ha una probabilità del 90% di vivere almeno fino a 47 anni, e una probabilità del 50% di vivere almeno fino a 80 anni. Che probabilità ha un uomo di 47 anni di vivere almeno fino a 80 anni? Abbiamo

$$\begin{aligned}\Omega &= \{\text{uomini di età } \geq \text{ 20 anni}\}\,,\\ \tilde{\Omega} &= \{\text{uomini di età } \geq \text{ 47 anni}\}\,,\\ A &= \{\text{uomini di età } \geq \text{ 80 anni}\}\,.\end{aligned}$$

Notiamo che $A \subseteq \tilde{\Omega} \subseteq \Omega$. Allora

$$P(A\,|\,\tilde{\Omega}) = \frac{P(A)}{P(\tilde{\Omega})} = \frac{\frac{50}{100}}{\frac{90}{100}} = \frac{5}{9}\,.$$

La probabilità di vivere fino agli 80 anni è quindi maggiore per un uomo di 47 anni che per un ragazzo di 20 anni. Magra consolazione!

Quanto visto sopra si può generalizzare anche nel caso in cui A non sia un sottoinsieme di $\tilde{\Omega}$. In tal caso, la "probabilità (condizionale) di A rispetto a $\tilde{\Omega}$" è definita da

$$P(A\,|\,\tilde{\Omega}) = \frac{P(A \cap \tilde{\Omega})}{P(\tilde{\Omega})}\,.$$

Esempio. Giochiamo alla roulette i numeri 7, 13 e 22. Siccome, com'è ben noto,

$$\Omega = \{0, 1, 2, 3, \ldots, 36\}\,,$$

la probabilità di vincere è di $\frac{3}{37}$. Ma ad un tratto veniamo a sapere da un amico che la roulette è truccata ed escono sempre solo numeri dispari. Allora, essendo

$$A = \{7, 13, 22\}\,, \qquad \tilde{\Omega} = \{1, 3, 5, \ldots, 35\}\,,$$

si ha che $A \cap \tilde{\Omega} = \{7, 13\}$ e quindi la nuova probabilità è:

$$P(A\,|\,\tilde{\Omega}) = \frac{P(A \cap \tilde{\Omega})}{P(\tilde{\Omega})} = \frac{\frac{2}{37}}{\frac{18}{37}} = \frac{1}{9}\,.$$

A questo punto, dati due sottoinsiemi A e B di Ω, si possono considerare entrambe le formule

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

Se $P(A\,|\,B) = P(A)$, si dice che "A è indipendente da B", e si ha che

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

In tal caso, si ha anche

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

per cui anche B è indipendente da A. Si dice allora che

“ A e B sono indipendenti ”.

Da quanto sopra, abbiamo quindi che

$$A \text{ e } B \text{ sono indipendenti} \qquad \Leftrightarrow \qquad P(A\cap B) = P(A)\,P(B)\,.$$

Esempio. Lanciamo contemporaneamente un dado e una moneta. Ho quindi il seguente insieme di 12 possibili risultati:

$$\begin{aligned}\Omega = \{\,& (1,T),(2,T),(3,T),(4,T),(5,T),(6,T),\\ & (1,C),(2,C),(3,C),(4,C),(5,C),(6,C)\}\,.\end{aligned}$$

Qual è la probabilità che dopo il lancio il dado mostri il numero 6? Essendo

$$A = \{(6,T),(6,C)\}\,,$$

la probabilità è $P(A) = \frac{2}{12} = \frac{1}{6}$. Supponiamo ora di venire a sapere che il lancio della moneta è truccato e che sicuramente il risultato della moneta sarà T. L'insieme dei possibili risultati è ora

$$B = \{(1,T),(2,T),(3,T),(4,T),(5,T),(6,T)\}\,.$$

Notiamo che $P(B) = \frac{6}{12} = \frac{1}{2}$ e, essendo $A\cap B = \{(6,T)\}$, abbiamo che $P(A\cap B) = \frac{1}{12}$. Quindi,

$$P(A\cap B) = \tfrac{1}{12} = P(A)\,P(B)\,,$$

per cui A (esce il 6) e B (esce “Testa”) sono eventi indipendenti.

7^a e 8^a ora di lezione

Dati due insiemi A e B, scrivendo $A = (A\cap B)\cup(A\cap \complement B)$, abbiamo che

$$P(A) = P(A\cap B) + P(A\cap \complement B)\,,$$

da cui la formula

$$P(A) = P(A\,|\,B)\,P(B) + P(A\,|\,\complement B)\,P(\complement B)\,.$$

Esempio. Consideriamo questo gioco: in un'urna ci sono nove palline bianche e una rossa. Estrae il primo giocatore: se trova la rossa, ha vinto. Se la rossa non è stata trovata, si mette da parte la pallina estratta ed estrae il secondo giocatore: se trova la rossa, ha vinto. Altrimenti, la partita finisce in parità. La domanda è: conviene estrarre per primi? Per secondi? O è lo stesso?

Chiaramente, il primo giocatore ha una probabilità di $\frac{1}{10}$ di trovare la pallina rossa e vincere la partita. Vogliamo ora calcolare la probabilità che ha il secondo giocatore di vincere. Consideriamo quindi l'evento

$$A: \text{ la pallina rossa esce alla seconda estrazione.}$$

Per poter usare la formula scritta sopra, consideriamo l'evento

$$B: \text{ la pallina rossa esce alla prima estrazione.}$$

Come abbiamo già detto, si ha $P(B) = \frac{1}{10}$; quindi, $P(\mathcal{C}B) = \frac{9}{10}$. Chiaramente, se la pallina rossa è uscita alla prima estrazione, la probabilità che esca alla seconda estrazione è uguale a 0. Quindi, $P(A \,|\, B) = 0$. D'altra parte, se la pallina rossa non è uscita alla prima estrazione, essendoci rimaste nove palline nell'urna, la probabilità che esca alla seconda estrazione è di $\frac{1}{9}$. Quindi, $P(A \,|\, \mathcal{C}B) = \frac{1}{9}$. Pertanto,

$$P(A) = P(A \,|\, B)\, P(B) + P(A \,|\, \mathcal{C}B)\, P(\mathcal{C}B) = 0 \cdot \frac{1}{10} + \frac{1}{9} \cdot \frac{9}{10} = \frac{1}{10}\,.$$

Si ha quindi che $P(A) = P(B)$, per cui i due giocatori hanno uguale probabilità di vittoria.

Prima variante. Supponiamo ora che ci siano tre giocatori. Le regole sono le stesse, con la novità che, se i primi due giocatori non hanno trovato la pallina rossa, estrae il terzo giocatore: se trova la rossa, ha vinto. Altrimenti, la partita finisce in parità.

Per calcolare la probabilità che ha il terzo giocatore di vincere, consideriamo l'evento

$$A: \text{ la pallina rossa esce alla terza estrazione.}$$

Per poter usare la formula scritta sopra, consideriamo l'evento

$$B: \text{ la pallina rossa esce alla prima o alla seconda estrazione.}$$

Allora, con considerazioni simili,

$$P(A) = P(A \,|\, B)\, P(B) + P(A \,|\, \mathcal{C}B)\, P(\mathcal{C}B) = 0 \cdot \frac{2}{10} + \frac{1}{8} \cdot \frac{8}{10} = \frac{1}{10}\,.$$

Di nuovo la stessa probabilità!

Naturalmente potremmo ora aggiungere altri giocatori, e troveremmo sempre le stesse probabilità. Il caso di dieci giocatori è simile al "gioco degli stecchetti": di dieci stecchetti uguali, uno viene spezzato a metà. Si dispongono poi gli stecchetti parzialmente nascosti in modo da non distinguere qual è quello dimezzato, e si procede all'estrazione a turno. Come abbiamo visto, tutti avranno la stessa probabilità di vincere (o di perdere).

Seconda variante. Supponiamo ora che nell'urna ci siano otto palline bianche e due rosse. Ricordiamo le regole del gioco: se il primo giocatore trova una rossa, ha vinto. In caso contrario, estrae il secondo giocatore: se trova una rossa, ha vinto. Altrimenti, la partita finisce in parità.

In questo caso, il primo giocatore ha una probabilità di $\frac{2}{10}$ di trovare una pallina rossa e vincere la partita. Consideriamo ora l'evento

$$A: \text{ vince il secondo giocatore.}$$

Per poter usare la formula scritta sopra, consideriamo l'evento

$$B: \text{ una pallina rossa esce alla prima estrazione.}$$

Come abbiamo già detto, si ha $P(B) = \frac{1}{5}$; quindi, $P(\mathcal{C}B) = \frac{4}{5}$. Se una pallina rossa è uscita alla prima estrazione, il secondo giocatore non può vincere la partita. Quindi, $P(A \,|\, B) = 0$. D'altra parte, se alla prima estrazione non è uscita alcuna pallina rossa, essendoci rimaste nove palline nell'urna, di cui due rosse, la probabilità che esca alla seconda estrazione è di $\frac{2}{9}$. Quindi, $P(A \,|\, \mathcal{C}B) = \frac{2}{9}$. Pertanto,

$$P(A) = P(A \,|\, B)\, P(B) + P(A \,|\, \mathcal{C}B)\, P(\mathcal{C}B) = 0 \cdot \tfrac{1}{5} + \tfrac{2}{9} \cdot \tfrac{4}{5} = \tfrac{8}{45}\,.$$

Si ha quindi che $P(A) < P(B)$, per cui il primo giocatore ha una probabilità maggiore di vincere.

Terza variante. Come nella seconda variante, supponiamo che nell'urna ci siano otto palline bianche e due rosse. Le regole del gioco sono però diverse. Anche se il primo giocatore estrae una pallina rossa, si dà al secondo giocatore la possibilità di estrarre a sua volta, e se anche lui trova una pallina rossa, la partita finisce in parità.

Il primo giocatore ha una probabilità di $\frac{1}{5}$ di trovare una pallina rossa. Consideriamo ora gli eventi

$$A: \text{ una pallina rossa esce alla seconda estrazione,}$$

e, come sopra,

$$B: \text{ una pallina rossa esce alla prima estrazione.}$$

Si ha che $P(B) = \frac{1}{5}$ e $P(\mathcal{C}B) = \frac{4}{5}$. Se una pallina rossa è uscita alla prima estrazione, nell'urna sono rimaste nove palline, di cui una rossa. Quindi, $P(A \,|\, B) = \frac{1}{9}$. D'altra parte, se alla prima estrazione non è uscita alcuna pallina rossa, nell'urna ci sono ancora due rosse e quindi $P(A \,|\, \mathcal{C}B) = \frac{2}{9}$. Pertanto,

$$P(A) = P(A \,|\, B)\, P(B) + P(A \,|\, \mathcal{C}B)\, P(\mathcal{C}B) = \tfrac{1}{9} \cdot \tfrac{1}{5} + \tfrac{2}{9} \cdot \tfrac{4}{5} = \tfrac{9}{45} = \tfrac{1}{5}\,.$$

Si ha quindi che $P(A) = P(B)$, per cui i due giocatori hanno uguale probabilità di vittoria.

Si può facilmente controllare che, con queste nuove regole, qualunque sia il numero delle palline rosse i due giocatori hanno la stessa probabilità di vincere. Questo gioco ricorda la "pesca di beneficienza", dove ognuno compera un biglietto che viene estratto da un'urna, ma solo alcuni biglietti attribuiscono un premio a chi li ha comprati. Da quanto visto qui, non c'è bisogno di affrettarsi!

Torniamo alle formule

$$P(A \cap B) = P(A \mid B)\, P(B) = P(B \mid A)\, P(A)\,.$$

Se ne deduce la "formula di Bayes",

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)} = \frac{P(A \mid B)\, P(B)}{P(A \mid B)\, P(B) + P(A \mid \mathcal{C}B)\, P(\mathcal{C}B)}\,,$$

che permette i calcolare la probabilità condizionale $P(B \mid A)$, note che siano $P(A \mid B)$, $P(A \mid \mathcal{C}B)$ e $P(B)$ (per cui anche $P(\mathcal{C}B) = 1 - P(B)$).

Esempio. Abbiamo due monete, di cui una truccata. La prima, quella normale, presenta le facce $\{T, C\}$, la seconda invece $\{C, C\}$. Ne prendiamo una a caso, la lanciamo e si verifica l'evento

$$A: \text{ esce "Croce".}$$

Qual è la probabilità che abbiamo scelto la moneta truccata? Usiamo la formula di Bayes, con

$$B: \text{ viene lanciata la moneta truccata.}$$

Il problema è di trovare $P(B \mid A)$. Essendo stata scelta una moneta a caso, abbiamo che $P(B) = P(\mathcal{C}B) = \frac{1}{2}$. Se la moneta scelta è quella truccata, è certo che uscirà "Croce", quindi $P(A \mid B) = 1$. Se invece la moneta scelta è quella normale, c'è uguale probabilità che esca "Testa" o "Croce", quindi $P(A \mid \mathcal{C}B) = \frac{1}{2}$. Allora,

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A \mid B)\, P(B) + P(A \mid \mathcal{C}B)\, P(\mathcal{C}B)} = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}} = \frac{2}{3}\,.$$

La formula di Bayes, che talvolta si chiama anche "formula della probabilità delle cause", può essere così generalizzata: se $B_1, B_2, \dots, B_n$ sono a due a due disgiunti e $B_1 \cup B_2 \cup \ldots \cup B_n = \Omega$, allora

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{P(A \mid B_1)\, P(B_1) + P(A \mid B_2)\, P(B_2) + \ldots + P(A \mid B_n)\, P(B_n)}\,,$$

per ogni $i = 1, 2, \dots, n$.

9^a ora di lezione

La legge binomiale. In un'urna ci sono delle palline di due colori, bianche e rosse. Supponiamo di conoscere la probabilità di estrarre una pallina rossa: sia essa p. Pertanto, la probabilità di estrarre una pallina bianca è $q = 1 - p$. Siano k e n due numeri naturali, con $k \leq n$.

<u>Domanda</u>. Operando n estrazioni, con reimbussolamento, qual è la probabilità di aver estratto, in tutto, esattamente k palline rosse?

Per rispondere alla domanda, indichiamo con "0" l'uscita di una pallina bianca e con "1" l'uscita di una pallina rossa. L'insieme Ω sarà allora costituito dalle stringhe di n elementi, ognuno dei quali può valere 0 oppure 1:

$$\Omega = \{(\omega_1, \omega_2, \ldots, \omega_n) : \omega_i \in \{0, 1\}\}\,.$$

Consideriamo, per cominciare, l'elemento

$$\omega = (1, 1, \ldots, 1, 0, 0, \ldots, 0)\,,$$

dove 1 compare ai primi k posti, mentre nei rimanenti $n - k$ posti c'è lo 0. Inoltre, per $i = 1, 2, \ldots, n$, sia

$$A_i = \{(\omega_1, \omega_2, \ldots, \omega_n) : \omega_i = 1\}\,.$$

Pertanto, A_1 corrisponde all'evento "il primo risultato è 1", A_2 all'evento "il secondo risultato è 1", ecc. Abbiamo quindi che, per quanto riguarda il nostro $\omega = (1, 1, \ldots, 1, 0, 0, \ldots, 0)$,

$$\{\omega\} = A_1 \cap A_2 \cap \ldots \cap A_k \cap \mathcal{C}A_{k+1} \cap \mathcal{C}A_{k+2} \cap \ldots \cap \mathcal{C}A_n\,.$$

Siccome questi eventi sono a due a due indipendenti, avremo che

$$\begin{aligned} P(\{\omega\}) &= P(A_1)\, P(A_2)\, \ldots\, P(A_k)\, P(\mathcal{C}A_{k+1})\, P(\mathcal{C}A_{k+2})\, \ldots\, P(\mathcal{C}A_n) \\ &= p \cdot p \cdot \ldots \cdot p \cdot q \cdot q \cdot \ldots \cdot q = p^k q^{n-k}\,. \end{aligned}$$

Ora, osserviamo che ogni stringa di n elementi di cui k sono uguali a 1 e $n-k$ sono uguali a 0 ha la stessa probabilità di comparire, e che tale probabilità vale, per quanto visto sopra, esattamente $p^k q^{n-k}$. Quante sono in tutto queste stringhe? Esattamente quanti sono i sottoinsiemi di k elementi di un insieme di n elementi, cioè $\binom{n}{k}$. Abbiamo quindi trovato la

Risposta. La probabilità cercata è: $\binom{n}{k} p^k q^{n-k}$.

In questo caso, si parla di "legge binomiale" in quanto il risultato trovato è uno degli addendi dello sviluppo nella formula di Newton

$$(p + q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}\,.$$

Questa situazione si può presentare se nell'urna ci sono m palline, di cui r palline rosse e b palline bianche. In tal caso, $p = \frac{r}{m}$ e $q = \frac{b}{m}$. È interessante notare, però, che una situazione analoga si presenta ogni qualvolta ci si trovi di fronte a un problema di tipo successo - insuccesso, di n prove, una volta nota la probabilità p di successo: è il cosiddetto "schema di Bernoulli" $B(n, p)$.

La legge ipergeometrica. Supponiamo ora che nell'urna ci siano m palline, di cui r palline rosse e b palline bianche, per cui la probabilità di estrarre una pallina rossa è $p = \frac{r}{m}$ e la probabilità di estrarre una pallina bianca è $q = \frac{b}{m}$. Si noti che $m = r + b$ e che $q = 1 - p$. Siano k e n due numeri naturali, con $k \leq n \leq m$ e $k \leq r$.

Domanda. Operando n estrazioni, senza reimbussolamento, qual è la probabilità di aver estratto, in tutto, esattamente k palline rosse?

Conviene numerare le palline: le rosse da 1 a r e le bianche da $r+1$ a $r+b$, cioè m. Sia Ω l'insieme costituito dai sottoinsiemi di n elementi dell'insieme $\{1, 2, \ldots, m\}$. Esso ha $\binom{m}{n}$ elementi, ognuno dei quali rappresenta una serie di n estrazioni, senza reimbussolamento, dalla nostra urna di m palline. Il problema è quello di determinare quanti sono gli elementi di Ω che contengono esattamente k numeri tra 1 e r.

Abbiamo $\binom{r}{k}$ modi di scegliere questi k numeri tra 1 e r. Una volta scelti, abbiamo ancora $\binom{b}{n-k}$ modi di scegliere i rimanenti $n-k$ numeri tra $r+1$ e $r+b$. In totale abbiamo quindi $\binom{r}{k}\binom{b}{n-k}$ modi per scegliere k numeri tra 1 e r e $n-k$ numeri tra $r+1$ e $r+b$. Ricordando che Ω ha in tutto $\binom{m}{n}$ elementi, abbiamo la

Risposta. La probabilità cercata è: $\dfrac{\binom{r}{k}\binom{b}{n-k}}{\binom{m}{n}}$.

10^a e 11^a ora di lezione

Esempio. Qual è la probabilità di vincere un terno al lotto sulla ruota di Napoli, giocando i tre numeri 4, 13 e 85? Sappiamo che, nell'urna del lotto, ci sono $m = 90$ palline, e ne vengono estratte $n = 5$, senza reimbussolamento. Giocando tre numeri, posso immaginare di aver colorato di rosso tre palline, mentre tutte le altre 87 resteranno bianche. Ho così $r = 3$ e $b = 87$. Vorrei sapere qual è la probabilità di trovare $k = 3$ palline rosse (cioè tutte e tre le palline rosse). Usando la formula della legge ipergeometrica, la risposta è:

$$\frac{\binom{r}{k}\binom{b}{n-k}}{\binom{m}{n}} = \frac{\binom{3}{3}\binom{87}{2}}{\binom{90}{5}} = \frac{1}{11748} \,.$$

Quando i numeri diventano grandi, risulta difficile calcolare i coefficienti binomiali e conviene avvalersi dell'aiuto di un computer. Ci sono diversi programmi che permettono un calcolo agevole delle probabilità relative alla legge binomiale e alla legge ipergeometrica. A lezione, abbiamo visto alcuni esempi di utilizzo del programma Excel. Esso presenta il vantaggio di aver già programmate le distribuzioni binomiale (DISTRIB.BINOM) e ipergeometrica (DISTRIB.IPERGEOM). Impostando i parametri, si possono ottenere dei grafici interessanti.

Ad esempio, per la distribuzione binomiale, fissati n e p, si può far disegnare al programma un istogramma della funzione

$$k \mapsto \binom{n}{k} p^k (1-p)^{n-k} \,,$$

con k che varia da 0 a n. È interessante osservare che, se n è scelto abbastanza grande, il grafico assume una forma "a campana", che assomiglia a quello di una "funzione gaussiana". Approfondiremo in seguito questo fenomeno.

12^a ora di lezione

Introdurremo ora il concetto di "variabile aleatoria". Cominciamo con alcuni esempi.

Il dado del bevitore. Talvolta, sulle bancarelle, si può trovare in vendita un dado che, al posto dei numeri 1, 2, ..., 6, possiede delle scritte, ad esempio:

$$
\begin{aligned}
1 &\mapsto \text{whisky} \\
2 &\mapsto \text{vodka} \\
3 &\mapsto \text{champagne} \\
4 &\mapsto \text{rhum} \\
5 &\mapsto \text{gin} \\
6 &\mapsto \text{go home}
\end{aligned}
$$

Il messaggio è chiaro: il bevitore delega alla sorte la scelta sul da farsi, contemplando anche la possibilità, con probabilità $\frac{1}{6}$, di smettere di bere e tornare a casa.

Il gioco d'azzardo. Ci viene proposta la seguente sfida: lanciando un dado, vinciamo o perdiamo del denaro, con le seguenti regole:

$$
\begin{aligned}
1 &\mapsto \text{vinciamo 10 Euro} \\
2 &\mapsto \text{perdiamo 3 Euro} \\
3 &\mapsto \text{vinciamo 5 Euro} \\
4 &\mapsto \text{vinciamo 2 Euro} \\
5 &\mapsto \text{vinciamo 1 Euro} \\
6 &\mapsto \text{perdiamo 12 Euro}
\end{aligned}
$$

Conviene giocare? Facciamo un semplice conto: se ogni risultato ha probabilità $\frac{1}{6}$ di uscire, ad ogni giocata dovremmo guadagnare, "in media",

$$\tfrac{1}{6}\,(10 - 3 + 5 + 2 + 1 - 12) = 0,50 \text{ Euro}\,.$$

Quindi, conviene! Si guadagnano, "in media", 50 centesimi ad ogni giocata.

Ma un amico fidato ci informa che il dado è truccato. Ci sono probabilità diverse di uscita dei singoli numeri, e precisamente:

$$
\begin{aligned}
1 &\mapsto \text{probabilità } 15\% \\
2 &\mapsto \text{probabilità } 5\% \\
3 &\mapsto \text{probabilità } 10\% \\
4 &\mapsto \text{probabilità } 20\% \\
5 &\mapsto \text{probabilità } 28\% \\
6 &\mapsto \text{probabilità } 22\%
\end{aligned}
$$

Conviene ancora giocare? Questa volta, nel fare il conto del "guadagno medio" a ogni giocata, dobbiamo "pesare" le vincite e le perdite possibili con le loro probabilità:

$$10 \cdot \tfrac{15}{100} - 3 \cdot \tfrac{5}{100} + 5 \cdot \tfrac{10}{100} + 2 \cdot \tfrac{20}{100} + 1 \cdot \tfrac{28}{100} - 12 \cdot \tfrac{22}{100} = -\,0,11 \text{ Euro}\,.$$

Quindi, non conviene! Si perdono, "in media", 11 centesimi ad ogni giocata.

Analizziamo i due esempi presentati. Abbiamo un insieme Ω delle possibili eventualità, precisamente

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

Ogni eventualità ha una certa probabilità di verificarsi:

$$\begin{array}{lll} P(\{1\}) = p_1\,, & P(\{2\}) = p_2\,, & P(\{3\}) = p_3\,, \\ P(\{4\}) = p_4\,, & P(\{5\}) = p_5\,, & P(\{6\}) = p_6\,. \end{array}$$

Inoltre, ad ogni eventualità è associato un elemento di un certo insieme E. Nel caso del dado del bevitore, E è l'insieme delle azioni da intraprendere

$$E = \{\text{bevo whisky, bevo vodka, } \ldots\}$$

Nel caso del gioco d'azzardo, possiamo prendere per E l'insieme $\mathbb{N}$ dei numeri naturali, o l'insieme $\mathbb{R}$ dei numeri reali. Viene a essere così definita una funzione da Ω a E, che a ogni numero $1, 2, \ldots, 6$ associa l'elemento corrispondente dell'insieme E. Una tale funzione $X : \Omega \to E$ si chiama "variabile aleatoria".

Nel seguito, saremo principalmente interessati al caso in cui E sia l'insieme $\mathbb{R}$ dei numeri reali. Accenneremo anche ai casi in cui E possa essere un insieme di vettori, $\mathbb{R}^2$ ad esempio. L'insieme Ω potrà avere un numero finito o anche un numero infinito di elementi. Questo fatto ci porterà a distinguere tra variabili aleatorie "discrete" e "continue".

Iniziamo con il considerare il caso in cui Ω abbia un numero finito N di elementi. Possiamo quindi numerarli e supporre, senza perdere generalità, che sia

$$\Omega = \{1, 2, \ldots, N\}\,.$$

Ad ogni evento di Ω è associata una probabilità:

$$P(\{1\}) = p_1\,, \quad P(\{2\}) = p_2\,, \quad \ldots \quad , P(\{N\}) = p_N\,.$$

(Ricordiamo che deve essere $p_1 + p_2 + \ldots + p_N = 1$.) La variabile aleatoria è una funzione

$$X : \Omega \to \mathbb{R}\,,$$

che a ogni numero $1, 2, \ldots, N$ associa un numero reale

$$\begin{array}{l} 1 \mapsto x_1 \\ 2 \mapsto x_2 \\ \ldots \\ N \mapsto x_N \end{array}$$

In altri termini,

$$X(1) = x_1\,, \quad X(2) = x_2\,, \quad \ldots \quad, X(N) = x_N\,.$$

Osservazioni. 1) Non è restrittivo supporre che gli elementi $x_1, x_2, \ldots, x_N$ siano tutti distinti. In effetti, se così non fosse, basterebbe identificare quelli uguali tra loro, avendo poi cura di sommare le probabilità ad essi associate.

2) Non sarebbe nemmeno necessario supporre che l'insieme Ω abbia un numero finito di elementi: basta infatti che l'insieme immagine della variabile aleatoria sia finito:

$$X(\Omega) = \{x_1, x_2, \ldots, x_N\}\,.$$

In tal caso, ad ogni x_i corrisponde un sottoinsieme A_i di Ω tale che

$$A_i = \{t \in \Omega : X(t) = x_i\}\,,$$

e questi insiemi $A_1, A_2, \ldots, A_N$ costituiscono gli elementi di un nuovo insieme $\hat{\Omega}$, che pertanto ha un numero finito di elementi, con il quale si può procedere come sopra.

13^a e 14^a ora di lezione

Definiamo la "media" della variabile aleatoria $X : \Omega \to \mathbb{R}$ considerata alla fine della lezione precedente:

$$E[X] = \sum_{j=1}^{N} x_j p_j = x_1 p_2 + x_2 p_2 + \ldots + x_N p_N\,.$$

Essa si chiama anche "speranza matematica" o "valore atteso" di X, e si indica anche con μ_X o semplicemente con μ.

Si definisce inoltre la "varianza" della variabile aleatoria X:

$$V[X] = \sum_{j=1}^{N} (x_j - \mu)^2 p_j = (x_1 - \mu)^2 p_1 + (x_2 - \mu)^2 p_2 + \ldots + (x_N - \mu)^2 p_N\,,$$

dove $\mu = E[X]$. Infine, definiamo la "deviazione standard" semplicemente come la radice quadrata della varianza:

$$\sigma_X = \sqrt{V[X]}\,.$$

Qualora non ci siano ambiguità, scriveremo σ invece di σ_X.

Esempio. Nel gioco d'azzardo introdotto la lezione precedente avevamo $\Omega = \{1, 2, 3, 4, 5, 6\}$ e $X : \Omega \to \mathbb{R}$ era la seguente variabile aleatoria:

$$\begin{array}{ll}
1 \mapsto 10 & \text{(con probabilità } 15\%) \\
2 \mapsto -3 & \text{(con probabilità } 5\%) \\
3 \mapsto 5 & \text{(con probabilità } 10\%) \\
4 \mapsto 2 & \text{(con probabilità } 20\%) \\
5 \mapsto 1 & \text{(con probabilità } 28\%) \\
6 \mapsto -12 & \text{(con probabilità } 22\%)
\end{array}$$

Abbiamo già calcolato la media:

$$\mu = 10 \cdot \tfrac{15}{100} - 3 \cdot \tfrac{5}{100} + 5 \cdot \tfrac{10}{100} + 2 \cdot \tfrac{20}{100} + 1 \cdot \tfrac{28}{100} - 12 \cdot \tfrac{22}{100} = -\,0,11\,.$$

Calcoliamo la varianza:

$$\begin{aligned} V[X] = {} & (10-\mu)^2 \cdot \frac{15}{100} + (-3-\mu)^2 \cdot \frac{5}{100} + (5-\mu)^2 \cdot \frac{10}{100} + \\ & + (2-\mu)^2 \cdot \frac{20}{100} + (1-\mu)^2 \cdot \frac{28}{100} + (-12-\mu)^2 \cdot \frac{22}{100} = 49,86269\,. \end{aligned}$$

La deviazione standard è quindi

$$\sigma = \sqrt{49,86269} \approx 7.06\,.$$

Date due variabili aleatorie $X : \Omega \to \mathbb{R}$ e $Y : \Omega \to \mathbb{R}$, si può considerare la loro somma: è la variabile aleatoria $X + Y : \Omega \to \mathbb{R}$ definita da

$$(X+Y)(\omega) = X(\omega) + Y(\omega)\,, \qquad \text{per ogni } \omega \in \Omega\,.$$

Analogamente, il prodotto $XY : \Omega \to \mathbb{R}$ è definito da

$$(XY)(\omega) = X(\omega)Y(\omega)\,, \qquad \text{per ogni } \omega \in \Omega\,.$$

Inoltre, data una costante $\alpha \in \mathbb{R}$, le variabili aleatorie $\alpha + X : \Omega \to \mathbb{R}$ e $\alpha X : \Omega \to \mathbb{R}$ sono definite da

$$(\alpha + X)(\omega) = \alpha + X(\omega)\,, \qquad (\alpha X)(\omega) = \alpha X(\omega)\,, \qquad \text{per ogni } \omega \in \Omega\,.$$

È come se identificassimo la costante $\alpha \in \mathbb{R}$ con la variabile aleatoria costante, che a ogni $t \in \Omega$ associa sempre α. Scriveremo anche $X + \alpha$ invece di $\alpha + X$, e $X\alpha$ invece di αX.

Si possono verificare senza troppe difficoltà le seguenti proprietà di linearità della media:

$$E[X+Y] = E[X] + E[Y]\,, \qquad E[\alpha X] = \alpha E[X]\,.$$

Notiamo inoltre che, per come è stata definita la varianza, si ha:

$$V[X] = E[(X-\mu)^2]\,,$$

dove $\mu = E[X]$. La varianza e la deviazione standard servono quindi per misurare quanto sono "dispersi", ossia quanto si allontanano dalla media, i valori di X.

15^a ora di lezione

Considereremo alcuni esempi di variabili aleatorie.

Esempio 1. Dato un sottoinsieme A di Ω, sia $I_A : \Omega \to \mathbb{R}$ la variabile aleatoria così definita:

$$I_A(t) = \begin{cases} 1 & \text{se } t \in A\,, \\ 0 & \text{se } t \notin A\,. \end{cases}$$

Calcoliamone la media:

$$\mu = 1 \cdot P(A) + 0 \cdot P(\mathcal{C}A) = P(A)\,.$$

Tenendo conto che $P(\mathcal{C}A) = 1 - P(A)$, la varianza è invece

$$V[I_A] = (1-\mu)^2 \cdot P(A) + (0-\mu)^2 \cdot P(\mathcal{C}A) = P(A)(1-P(A))\,.$$

Esempio 2: lancio di una moneta. In questo caso, abbiamo $\Omega = \{T, C\}$. Sappiamo che la moneta non è necessariamente simmetrica, ma siamo comunque a conoscenza della probabilità p di ottenere T e, di conseguenza, della probabilità $q = 1-p$ di ottenere C. La variabile aleatoria $X : \Omega \to \mathbb{R}$ è definita da:

$$X(T) = 1\,, \qquad X(C) = 0\,.$$

Calcoliamone la media,

$$E[X] = 1 \cdot p + 0 \cdot q = p\,,$$

e la varianza,

$$V[X] = (1-p)^2 p + (0-p)^2(1-p) = p(1-p)\,.$$

Si noti l'analogia con l'esempio precedente. In entrambi i casi siamo infatti in uno schema successo - insuccesso di tipo $B(1,p)$, dove, per quanto riguarda il primo esempio, $p = P(A)$.

Esempio 3: n lanci di una moneta. Cominciamo con il caso di due lanci della moneta dell'esempio precedente. Abbiamo quindi che

$$\Omega = \{(T,T),\, (T,C),\, (C,T),\, (C,C)\}\,.$$

Consideriamo la variabile aleatoria $X : \Omega \to \mathbb{R}$ che conta il numero di T ottenute:

$$X(T,T) = 2\,, \quad X(T,C) = 1\,, \quad X(C,T) = 1\,, \quad X(C,C) = 0\,.$$

Calcoliamone la media:

$$E[X] = 2 \cdot p^2 + 1 \cdot p(1-p) + 1 \cdot p(1-p) + 0 \cdot (1-p)^2 = 2p\,.$$

Vediamo ora il caso $n = 3$. Si ha qui

$$\begin{aligned}\Omega = \{&(T,T,T),\, (T,T,C),\, (T,C,T),\, (C,T,T),\\ &(T,C,C),\, (C,T,C),\, (C,C,T),\, (C,C,C)\}\,.\end{aligned}$$

Consideriamo sempre la variabile aleatoria che conta il numero di T:

$$X(T,T,T) = 3\,, \quad X(T,T,C) = X(T,C,T) = X(C,T,T) = 2\,,$$
$$X(T,C,C) = X(C,T,C) = X(C,C,T) = 1\,, \quad X(C,C,C) = 0\,.$$

Calcoliamone la media:

$$E[X] = 3 \cdot p^3 + 3[2 \cdot p^2(1-p)] + 3[1 \cdot p(1-p)^2] + 0 \cdot (1-p)^3 = 3p\,.$$

Viene quindi naturale congetturare che, effettuando n lanci, se X è la variabile aleatoria che conta il numero di T ottenute, si abbia $E[X] = np$. Per verificare ciò, invece di usare la definizione, che comporterebbe il difficile calcolo della somma $\sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$, useremo un semplice trucco:

scriveremo 1 invece di T e 0 invece di C.

Torniamo a vedere allora cosa succede nel caso $n = 3$. Con questo trucco, la variabile aleatoria è definita semplicemente come

$$X(\omega_1, \omega_2, \omega_3) = \omega_1 + \omega_2 + \omega_3\,.$$

Risulta utile a questo punto considerare le variabili aleatorie X_1, X_2 e X_3 così definite:

$$X_1(\omega_1, \omega_2, \omega_3) = \omega_1\,, \quad X_2(\omega_1, \omega_2, \omega_3) = \omega_2\,, \quad X_3(\omega_1, \omega_2, \omega_3) = \omega_3\,,$$

per cui si ha che $X = X_1 + X_2 + X_3$. Come nell'esempio 2, si trova che

$$E[X_1] = 1 \cdot p + 0 \cdot (1-p) = p\,,$$

e analogamente anche $E[X_2] = p$ e $E[X_3] = p$. Per la proprietà di linearità della media, si ritrova quindi

$$E[X] = E[X_1] + E[X_2] + E[X_3] = p + p + p = 3p\,,$$

lo stesso risultato già trovato per via diretta. Il vantaggio di questo metodo è che si generalizza immediatamente al caso di n lanci. Infatti, in questo caso si ha che

$$X(\omega_1, \omega_2, \ldots, \omega_n) = \omega_1 + \omega_2 + \ldots + \omega_n\,.$$

Si considerano le variabili aleatorie X_1, X_2, $\ldots$, X_n definite da

$$X_k(\omega_1, \omega_2, \ldots, \omega_n) = \omega_k\,,$$

per ogni $k = 1, 2, \ldots, n$, per cui si ha che $X = X_1 + X_2 + \ldots + X_n$. Ciascuna delle X_k ha media uguale a p, per cui

$$E[X] = E[X_1] + E[X_2] + \ldots + E[X_n] = p + p + \ldots + p = np\,.$$

16^a e 17^a ora di lezione

Continuiamo a considerare il caso in cui Ω è costituito da un numero finito di elementi.

Diremo che due variabili aleatorie $X : \Omega \to \mathbb{R}$ e $Y : \Omega \to \mathbb{R}$ sono "indipendenti" se, comunque presi $[a, b]$ e $[c, d]$ in $\mathbb{R}$, si ha

$$\begin{aligned} &P(\{t : a \leq X(t) \leq b \ \text{ e } \ c \leq Y(t) \leq d\}) = \\ &\qquad\qquad = P(\{t : a \leq X(t) \leq b\}) \cdot P(\{t : c \leq Y(t) \leq d\}) \,. \end{aligned}$$

Si può dimostrare che:

$$\text{se } X \text{ e } Y \text{ sono indipendenti, allora } E[XY] = E[X]E[Y] \,.$$

La formula si generalizza poi per induzione nel caso di un numero finito qualsiasi di variabili aleatorie.

Definiamo la "covarianza" di X e Y:

$$\mathrm{Co}V[X, Y] = E[XY] - E[X]E[Y] \,.$$

Da quanto sopra, abbiamo che

$$\text{se } X \text{ e } Y \text{ sono indipendenti, allora } \mathrm{Co}V[X, Y] = 0 \,.$$

Vogliamo ora trovare una formula per la varianza di $X + Y$:

$$\begin{aligned} V[X + Y] &= E\Big[((X + Y) - E[X + Y])^2\Big] \\ &= E\Big[(X - E[X] + Y - E[Y])^2\Big] \\ &= E\Big[(X - E[X])^2 + (Y - E[Y])^2 + 2(X - E[X])(Y - E[Y])\Big] \\ &= E\Big[(X - E[X])^2\Big] + E\Big[(Y - E[Y])^2\Big] + \\ &\qquad\qquad + 2E\Big[XY - XE[Y] - YE[X] + E[X]E[Y]\Big] \\ &= V[X] + V[Y] + 2(E[XY] - E[X]E[Y]) \,, \end{aligned}$$

da cui

$$V[X + Y] = V[X] + V[Y] + 2\,\mathrm{Co}V[X, Y] \,.$$

Abbiamo quindi che:

$$\text{se } X \text{ e } Y \text{ sono indipendenti, allora } V[X + Y] = V[X] + V[Y] \,.$$

Esempio: ancora n lanci di una moneta. Torniamo alla nostra moneta, non necessariamente simmetrica, per cui conosciamo la probabilità p di ottenere T. Se X è la variabile aleatoria che conta il numero di T dopo n lanci, abbiamo già visto che $E[X] = np$. Vogliamo ora calcolare la varianza di X. Ricordiamo che si può scrivere $X = X_1 + X_2 + \ldots + X_n$, con

$$X_k(\omega_1, \omega_2, \ldots, \omega_n) = \omega_k \,,$$

e si ha $E[X_k] = p$, per ogni $k = 1, 2, \ldots, n$, mentre

$$V[X_k] = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = p(1-p)\,.$$

Si può inoltre dimostrare che, se $X_1, X_2, \ldots, X_n$ sono a due a due indipendenti, allora

$$V[X] = V[X_1] + VX_2] + \ldots + V[X_n] = np(1-p)\,.$$

Talvolta è utile considerare il "coefficiente di correlazione" di X e Y:

$$\rho_{X,Y} = \frac{\mathrm{Co}V[X,Y]}{\sigma_X \sigma_Y}\,,$$

dove σ_X è la deviazione standard di X e σ_Y è la deviazione standard di Y. Esso è definito se $\sigma_X > 0$ e $\sigma_Y > 0$. Si può dimostrare che si ha sempre $-1 \leq \rho_{X,Y} \leq 1$.

Si possono verificare piuttosto facilmente le seguenti proprietà della varianza:

$$V[\alpha X] = \alpha^2 V[X]\,, \qquad V[\alpha + X] = V[X]\,.$$

Per ogni numero positivo η, vale inoltre la seguente

Disuguaglianza di Cebicev:

$$P(\{j \in \Omega : |X(j) - E[X]| \geq \eta\}) \leq \frac{V[X]}{\eta^2}\,.$$

Dimostriamola: se $\mu = E[X]$,

$$\begin{aligned} V[X] &= \sum_{j=1}^{N} (x_j - \mu)^2 p_j \\ &\geq \sum_{\{j : |x_j - \mu| \geq \eta\}} (x_j - \mu)^2 p_j \\ &\geq \eta^2 \sum_{\{j : |x_j - \mu| \geq \eta\}} p_j \\ &= \eta^2 P(\{j : |x_j - \mu| \geq \eta\})\,, \end{aligned}$$

per cui la disuguaglianza è verificata.

18^a ora di lezione

Torniamo ora agli n lanci della moneta. Questa volta, invece di considerare la variabile aleatoria X che conta il numero di T ottenute, consideriamo la variabile aleatoria $F_n = \frac{1}{n}X$, che ci fornisce la "frequenza" con la quale si ottiene T. Richiamandoci a quanto già visto, possiamo quindi scrivere

$$F_n = \tfrac{1}{n}(X_1 + X_2 + \ldots + X_n)\,.$$

Troviamo allora che

$$E[F_n] = \tfrac{1}{n} \cdot np = p\,, \qquad V[F_n] = \left(\tfrac{1}{n}\right)^2 \cdot np(1-p) = \frac{p(1-p)}{n}\,.$$

Usando la disuguaglianza di Cebicev, abbiamo che, fissato arbitrariamente un numero positivo η,

$$P(\{j \in \Omega : |F_n(j) - p| \geq \eta\}) \leq \frac{p(1-p)}{n\,\eta^2}\,.$$

Di conseguenza,

$$\lim_{n\to\infty} P(\{j \in \Omega : |F_n(j) - p| \geq \eta\}) = 0\,.$$

Abbiamo così trovato la

Legge dei grandi numeri. In uno schema di tipo successo - insuccesso, di n prove, preso un qualsiasi numero positivo η (piccolo quanto si vuole), si osserva il seguente fenomeno: uno scarto tra la frequenza dei successi e la probabilità p di successo, di entità superiore a η, ha una probabilità di verificarsi che può essere resa arbitrariamente piccola, pur di prendere il numero di prove n sufficientemente grande.

Consideriamo ora il caso di una variabile aleatoria i cui valori non siano numeri reali, ma vettori. Abbiamo quindi $X : \Omega \to \mathbb{R}^M$, per un certo $M \geq 2$. Possiamo scrivere

$$X(j) = (X^{(1)}(j), X^{(2)}(j), \ldots, X^{(M)}(j))\,,$$

dove $X^{(1)} : \Omega \to \mathbb{R}$, $X^{(2)} : \Omega \to \mathbb{R}$, $\ldots$, $X^{(M)} : \Omega \to \mathbb{R}$ sono variabili aleatorie a valori reali, le cosiddette "componenti" di X. La media di X si definisce in modo simile a quanto già visto: sia

$$\Omega = \{1, 2, \ldots, N\}\,,$$

con

$$P(\{1\}) = p_1\,, \quad P(\{2\}) = p_2\,, \quad \ldots \quad , P(\{N\}) = p_N\,;$$

allora, ponendo

$$X(1) = x_1\,, \quad X(2) = x_2\,, \quad \ldots \quad , X(N) = x_N\,,$$

la media di X è definita da

$$E[X] = \sum_{j=1}^{N} x_j p_j \,.$$

Si noti che $E[X]$ è ora un vettore di $\mathbb{R}^M$: si ha infatti

$$E[X] = (E[X^{(1)}], E[X^{(2)}], \ldots, E[X^{(M)}]) \in \mathbb{R}^M \,.$$

Nei corsi di Fisica, si incontra una situazione analoga in presenza di N particelle puntiformi, posizionate nei $x_1, x_2, \ldots, x_N$, di massa $p_1, p_2, \ldots, p_N$, rispettivamente. In questa situazione, il numero M rappresenta la dimensione dello spazio: normalmente, si ha $1 \le M \le 3$. Il vettore $E[X]$, come da noi definito, rappresenta allora la posizione del "baricentro" di tali masse.

Come definire, in questa situazione, la varianza? Ispirandoci alla Fisica, se $M = 2$, la formula

$$V[X] = E[(X - E[X]) \cdot (X - E[X])] \,,$$

rappresenta il "momento d'inerzia" del sistema di masse rispetto a un asse ortogonale al piano, passante per il baricentro. In questo caso, il prodotto tra il vettore $X - E[X]$ e sè stesso deve essere interpretato come un prodotto scalare in $\mathbb{R}^2$. Le cose si complicano se $M \ge 3$. Ad esempio, se $M = 3$, il momento d'inerzia è un vettore di $\mathbb{R}^3$, le cui coordinate sono i tre momenti d'inerzia rispetto agli assi passanti per il baricentro e paralleli ai tre assi cartesiani. Per prudenza, non ci addentreremo oltre in questo argomento.

19^a e 20^a ora di lezione

Consideriamo ora il caso in cui Ω sia un intervallo di $\mathbb{R}$. Quindi, Ω può essere di uno dei seguenti tipi:

$$]\alpha,\beta[\,,\quad [\alpha,\beta[\,,\quad]\alpha,\beta]\,,\quad [\alpha,\beta]\,,\quad]\alpha,+\infty[\,,\quad [\alpha,+\infty[\,,\quad]-\infty,\beta[\,,\quad]-\infty,\beta]\,,$$

oppure $\Omega = \mathbb{R}$, nel qual caso si scrive anche

$$]-\infty,+\infty[\,.$$

Per definire una probabilità su Ω, si può far uso di una "densità": si tratta di una funzione $f : \Omega \to \mathbb{R}$, che supporremo integrabile, con le seguenti proprietà:

1) $f(t) \ge 0$, per ogni $t \in \Omega$;

2) $\displaystyle\int_\Omega f(s)\,ds = 1$.

Facendo uso di tale funzione densità, si può definire, per ogni sottoinsieme "misurabile" A di Ω, la sua probabilità:[1]

$$P(A) = \int_A f(s)\,ds \,.$$

Si tratta di un numero reale, con le seguenti proprietà:

[1]Nel seguito, useremo talvolta il simbolo di integrale anche per insiemi che non siano degli intervalli: i cosiddetti insiemi "misurabili". Lo studente che non avesse familiarità con queste cose non si spaventi troppo: il loro uso verrà limitato al minimo necessario.

(a) $0 \leq P(A) \leq 1$;

(b) $P(\emptyset) = 0$, $P(\Omega) = 1$;

(c) se A e B sono due sottoinsiemi disgiunti di Ω, allora

$$P(A \cup B) = P(A) + P(B)\,.$$

Si noti che, in particolare, se $[a,b]$ è un sottointervallo di Ω, si ha

$$P(\,]a,b[\,) = P(\,[a,b[\,) = P(\,]a,b]\,) = P(\,[a,b]\,) = \int_a^b f(s)\,ds\,.$$

Come nel caso in cui Ω ha un numero finito di elementi, si dimostra che

$$P(\mathcal{C}A) = 1 - P(A)\,,$$

e, in generale, che per due sottoinsiemi qualsiasi A e B di Ω,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,.$$

Si può inoltre definire la "probabilità condizionale"

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

Si dice che A e B sono "indipendenti" se $P(A|B) = P(A)$, o equivalentemente $P(B|A) = P(B)$, nel qual caso si ha

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

Una funzione $X : \Omega \to \mathbb{R}$ si chiama "variabile aleatoria". Consideriamo il caso in cui X sia "assolutamente continua": esiste cioè una densità

$$f_X : \mathbb{R} \to \mathbb{R}\,,$$

ad essa associata, per cui si abbia che, per ogni intervallo $[a,b]$,

$$P(\{t \in \Omega : a \leq X(t) \leq b\}) = \int_a^b f_X(s)\,ds\,.$$

In tal caso, si dice che "X ha densità f_X". In seguito, considereremo sempre solo variabili aleatorie di queso tipo. Si definisce la "media" (o "speranza matematica", o "valore atteso") di X:

$$E[X] = \int_{\mathbb{R}} s f_X(s)\,ds\,.$$

Talvolta, invece di $E[X]$, si scrive μ_X o semplicemente μ.

21^a ora di lezione

Cercheremo di spiegare il legame tra la nozione di media di una variabile aleatoria vista nella lezione precedente e quella vista nel caso in cui Ω sia costituito da un numero finito di elementi. In entrambi i casi, data la variabile aleatoria $X : \Omega \to \mathbb{R}$, si introduce la sua "funzione di ripartizione"

$$\mathcal{F}_X : \mathbb{R} \to \mathbb{R}\,,$$

definita da

$$\mathcal{F}_X(x) = P(\{t \in \Omega : X(t) \leq x\})\,.$$

Se Ω è un intervallo e supponiamo che esista $f_X : \mathbb{R} \to \mathbb{R}$, la densità associata alla variabile aleatoria X, e che sia una funzione continua, allora

$$\mathcal{F}_X(x) = \int_{-\infty}^{x} f_X(s)\,ds\,,$$

e, per il Teorema Fondamentale del Calcolo Differenziale e Integrale, si ha

$$\mathcal{F}'_X(x) = f_X(x)\,, \qquad \text{per ogni } x \in \mathbb{R}\,.$$

Sia ora Ω un insieme avente un numero finito di elementi. Come al solito, possiamo scrivere $\Omega = \{1, 2, \ldots, N\}$. Consideriamo la variabile aleatoria $X : \{1, 2, \ldots, N\} \to \mathbb{R}$, con valori

$$X(1) = x_1\,,\ X(2) = x_2\,,\ \ldots\,,\ X(N) = x_N\,.$$

Scriviamo inoltre

$$P(\{1\}) = p_1\,,\ P(\{2\}) = p_2\,,\ \ldots\,,\ P(\{N\}) = p_N\,,$$

e ricordiamo che $p_1 + p_2 + \ldots + p_N = 1$. Riordiniamo i punti $x_1, x_2, \ldots, x_N$ in ordine crescente:

$$x_{i_1} \leq x_{i_2} \leq \ldots \leq x_{i_N}\,.$$

La funzione di ripartizione $\mathcal{F}_X : \mathbb{R} \to \mathbb{R}$ ha quindi i seguenti valori:

$$\mathcal{F}_X(x) = \begin{cases} 0 & \text{se } x < x_{i_1}\,, \\ p_{i_1} & \text{se } x_{i_1} \leq x < x_{i_2}\,, \\ p_{i_1} + p_{i_2} & \text{se } x_{i_2} \leq x < x_{i_3}\,, \\ \ldots & \\ p_{i_1} + p_{i_2} + \ldots + p_{i_{N-1}} & \text{se } x_{i_{N-1}} \leq x < x_{i_N}\,, \\ 1 & \text{se } x \geq x_{i_N}\,. \end{cases}$$

Tale funzione non è derivabile, essendo una funzione "a scalini". Procediamo allora approssimando il suo grafico con quello di una funzione derivabile. È possibile fare questo modificando la funzione in un piccolo intorno di ciascun punto x_{i_k}. Vediamo come si fa: si prende un intero m sufficientemente grande e si modifica nell'intervallo $[x_{i_k} - \frac{1}{m}, x_{i_k} + \frac{1}{m}]$ prendendo una funzione crescente e "liscia" che passi dal valore che $\mathcal{F}_X$ assume immediatamente prima di x_{i_k} al valore che $\mathcal{F}_X$ assume immediatamente dopo di x_{i_k}. Si ottiene così una nuova funzione $\mathcal{F}_{X,m} : \mathbb{R} \to \mathbb{R}$ che ora è derivabile. Sia dunque $f_{X,m} : \mathbb{R} \to \mathbb{R}$ definita da

$$f_{X,m}(x) = \mathcal{F}'_{X,m}(x)\,, \quad \text{per ogni } x \in \mathbb{R}\,.$$

Si può verificare che si tratta di una densità.

Il legame tra le diverse definizioni di media date per una variabile aleatoria è spiegato dalla seguente formula:

$$E[X] = x_1 p_1 + x_2 p_2 + \ldots + x_N p_N = \lim_{m\to\infty} \int_{\mathbb{R}} s f_{X,m}(s)\, ds\,.$$

22^a e 23^a ora di lezione

Si può dimostrare che, prese due variabili aleatorie $X : \Omega \to \mathbb{R}$, $Y : \Omega \to \mathbb{R}$ e una costante $\alpha \in \mathbb{R}$, si ha

$$E[X+Y] = E[X] + E[Y]\,, \qquad E[\alpha X] = \alpha E[X]\,.$$

Se $\mu = E[X]$, si definisce la "varianza" di X,

$$V[X] = E[(X-\mu)^2]\,.$$

Se f_X è la densità associata a X, si può scrivere anche

$$V[X] = \int_{\mathbb{R}} (s-\mu)^2 f_X(s)\, ds\,.$$

Si dimostra che, presa una costante $\alpha \in \mathbb{R}$, si ha

$$V[\alpha + X] = V[X]\,, \qquad V[\alpha X] = \alpha^2 V[X]\,.$$

Si definisce inoltre la "deviazione standard" di X:

$$\sigma_X = \sqrt{V[X]}\,.$$

Esempio. Usando la nota formula

$$\int_{\mathbb{R}} \exp(-s^2)\, ds = \sqrt{\pi}\,,$$

si può dimostrare che la funzione $g_{\mu,\sigma} : \mathbb{R} \to \mathbb{R}$, definita da

$$g_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2} \Big),$$

è una densità (qui prendiamo $\Omega = \mathbb{R}$) : si ha $g_{\mu,\sigma}(t) > 0$ per ogni $t \in \mathbb{R}$ e

$$\int_{\mathbb{R}} g_{\mu,\sigma}(s)\, ds = 1\,.$$

Si può inoltre dimostrare che, se $X : \mathbb{R} \to \mathbb{R}$ è una variabile aleatoria con densità $f_X = g_{\mu,\sigma}$, allora $E[X] = \mu$ (ossia $\mu_X = \mu$) e $V[X] = \sigma^2$, per cui $\sigma_X = \sigma$. Scriveremo in tal caso

$$X \sim N(\mu, \sigma^2)\,,$$

e diremo che X ha una distribuzione *normale* o *gaussiana* di parametri μ e σ^2. È utile considerare la funzione di ripartizione di $N(\mu, \sigma^2)$:

$$\mathcal{G}_{\mu,\sigma}(x) = \int_{-\infty}^{x} g_{\mu,\sigma}(s)\,ds\,.$$

In particolare, se $X \sim N(0,1)$, si ha $E[X] = 0$ e $V[X] = 1$. Ad essa corrisponde la funzione di ripartizione $\mathcal{G}_{0,1}$, che in seguito indicheremo con Φ. Si ha:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2}\,ds\,.$$

Torniamo a considerare la variabile aleatoria F_n che fornisce la frequenza di T nel caso di n lanci di una moneta. Se p è la probabilità di ottenere T, ricordiamo che

$$E[F_n] = p\,, \qquad V[F_n] = \frac{p\,(1-p)}{n}\,.$$

Ecco allora che la variabile aleatoria "rinormalizzata",

$$R_n = \sqrt{\frac{n}{p\,(1-p)}}\,(F_n - p)\,,$$

ha media $E[R_n] = 0$ e varianza $V[R_n] = 1$. Consideriamo la sua funzione di ripartizione:

$$\mathcal{F}_{R_n}(x) = P(\{t \in \Omega : R_n(t) \le x\})\,.$$

Il **Teorema Limite Centrale** permette di dimostrare il seguente notevole risultato:

$$\lim_{n\to\infty} \mathcal{F}_{R_n}(x) = \Phi(x)\,, \qquad \text{per ogni } x \in \mathbb{R}\,.$$

24^a ora di lezione

La "covarianza" di due variabili aleatorie $X, Y : \Omega \to \mathbb{R}$ è definita da

$$\mathrm{Co}V[X,Y] = E[XY] - E[X]E[Y]\,.$$

Vale la formula

$$V[X+Y] = V[X] + V[Y] + 2\,\mathrm{Co}V[X,Y]\,.$$

Si dice che X e Y sono "indipendenti" se, comunque presi $[a,b]$ e $[c,d]$ in $\mathbb{R}$, si ha

$$\begin{aligned} P(\{t : a \le X(t) \le b \ \text{ e } \ c \le Y(t) \le d\}) &= \\ &= P(\{t : a \le X(t) \le b\}) \cdot P(\{t : c \le Y(t) \le d\})\,. \end{aligned}$$

Si può dimostrare che, se X e Y sono indipendenti, allora $\mathrm{Co}V[X,Y] = 0$, per cui $V[X+Y] = V[X] + V[Y]$.

Se $\sigma_X > 0$ e $\sigma_Y > 0$, si definisce il "coefficiente di correlazione" di X e Y:

$$\rho_{X,Y} = \frac{\mathrm{Co}V[X,Y]}{\sigma_X \sigma_Y}\,.$$

Si può dimostrare che si ha sempre $-1 \leq \rho_{X,Y} \leq 1$.

Come nel caso in cui Ω era costituito da un numero finito di elementi, anche ora che Ω è un intervallo si può considerare il caso di una variabile aleatoria i cui valori siano non più numeri reali, ma vettori. Abbiamo quindi $X : \Omega \to \mathbb{R}^M$, per un certo $M \geq 2$. Possiamo scrivere

$$X(t) = (X^{(1)}(t), X^{(2)}(t), \dots, X^{(M)}(t))\,,$$

dove $X^{(1)} : \Omega \to \mathbb{R}$, $X^{(2)} : \Omega \to \mathbb{R}$, $\dots$, $X^{(M)} : \Omega \to \mathbb{R}$ sono le "componenti" di X. La media di X si definisce come

$$E[X] = (E[X^{(1)}], E[X^{(2)}], \dots, E[X^{(M)}]) \in \mathbb{R}^M\,.$$

Nei corsi di Fisica, si incontra una situazione analoga in presenza di un filo curvilineo con densità di massa possibilmente variabile. Il vettore $E[X]$, come da noi definito, rappresenta allora la posizione del "baricentro" di tale filo.

Considerazioni analoghe a quelle fatte nel caso dell'insieme Ω finito potrebbero portare a definire la varianza, nel caso $M = 2$, ispirandoci alla Fisica, in analogia con il "momento d'inerzia" di un filo curvilineo rispetto a un asse ortogonale al piano, passante per il baricentro. Come abbiamo visto, le cose si complicano se $M \geq 3$ e preferiamo non addentrarci oltre in questo argomento.

Come ulteriore estensione della teoria, si può considerare il caso in cui Ω sia non più un intervallo, ma un sottoinsieme di uno spazio $\mathbb{R}^K$, di dimensione K. La teoria generale dell'integrale permette infatti di trattare tali situazioni.

Un esempio di situazione di questo tipo è stato proposto all'inizio del corso, quando ci si chiedeva di valutare la probabilità che un punto, scelto a caso in un quadrato nel piano, appartenesse ad un cerchio.

Le diverse nozioni introdotte durante il corso si possono estendere anche a questo caso, con la dovuta cautela. Ma, per una trattazione completa di queste situazioni, ventiquattr'ore non sono sufficienti...

Libri consigliati:

[1] P. Baldi, Calcolo delle Probabiltà e Statistica, McGraw - Hill Ed., Milano, 1992.

[2] N. Pintacuda, Primo Corso di Probablità, Muzzio Ed., Padova, 1983.