

Un evento A è una proposizione logica non ambigua, cioè ben definita, che può essere vera o falsa. Ad ogni evento si può associare il suo indicatore e si indica $|A| = \begin{cases} 1 & \text{se } A \text{ è vero} \\ 0 & \text{se } A \text{ è falso.} \end{cases}$

Con il simbolo A^c = negazione di A , si indica l'evento A contrario. Esso è vero quando A è falso ed è falso quando A è vero. Esempio:

A = "In 5 lanci di un dado il 6 ~~non~~ esce almeno una volta"

A^c = "Il 6 non esce mai"

Dati due eventi si definisce loro unione $A \dot{\vee} B$, ^{simbolo} simbolo logico $A \cup B$, ^{simbolo} simbolo insiemistico

L'evento $A \cup B$ è vero quando almeno uno dei due è vero ed è falso quando sono entrambi falsi. Più semplicemente l'evento $A \cup B$ può essere individuato mediante una nuova proposizione ottenuta collegando le proposizioni che individuano gli eventi A e B con la congiunzione "o".

Dati due eventi si definisce loro intersezione e si indica con i simboli $A \cap B$, $A \wedge B$, AB . L'evento che è vero quando sono veri entrambi ed è falso quando almeno uno dei due è falso. L'evento intersezione è individuato dalla proposizione che si ottiene collegando le proposizioni di A e di B con la congiunzione "e".

Si può verificare, attraverso le definizioni di unione e intersezione, che valgono le seguenti proprietà sugli indicatori:

$$|A \cup B| = |A| + |B| - |AB|$$

$$|AB| = |A| \cdot |B|$$

Valgono le seguenti proprietà tra gli eventi:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A, \quad \text{proprietà commutativa.}$$

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (A \cap B) \cap C = A \cap (B \cap C), \quad \text{proprietà associativa.}$$

$$(A \cup B) \cap C = A \cap C \cup B \cap C, \quad \text{proprietà distributiva}$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C), \quad \text{" "}$$

Mediante i diagrammi di Venn aiutano la comprensione di queste proprietà.

Diremo che A implica B , $A \Rightarrow B$ (insiemisticamente ciò equivale a $A \subseteq B$), quando se è vero A allora anche B è vero e si ha $|A| \leq |B|$.

Una proposizione sempre vera sarà detta evento certo Ω e si indica con Ω ed ovviamente $|\Omega| = 1$, mentre una proposizione sempre falsa sarà detta evento impossibile, $\emptyset$, e si ha $|\emptyset| = 0$.

Due eventi si dicono incompatibili quando la loro intersezione è impossibile, cioè quando $A \cap B = \emptyset$, evento impossibile.

Osserviamo che valgono le seguenti relazioni:

$$A \vee \Omega = \Omega$$

$$A \cup A = A$$

$$A A = A$$

$$A \cup A^c = \Omega$$

$$A \Omega = A$$

$$A A^c = \emptyset$$

$$A \cap \emptyset = \emptyset$$

Si dice che un insieme finito di eventi incompatibili costituisce una classe completa o una partizione di Ω se la loro somma logica è l'evento certo, cioè se

$$1) H_i H_j = \emptyset, i \neq j, i, j = 1, 2, \dots, n \quad (\text{incompatibilità})$$

$$2) \bigcup_{i=1}^n H_i = \Omega$$

o in maniera del tutto analoga se gli H_i formano una partizione se $\sum_{i=1}^n |H_i| = 1$.

Consideriamo n eventi, $A_1, A_2, \dots, A_n$

$$\text{Osserviamo che } (A_1 \cup A_1^c) \cap (A_2 \cup A_2^c) \cap \dots \cap (A_n \cup A_n^c) = \Omega$$

Sviluppando questa espressione si ottengono tutte le possibili disposizioni di n elementi dove al posto i -esimo la scelta è tra A_i e A_i^c . Infatti, ad esempio si ha:

$$(A \cup A^c) \cap (B \cup B^c) = (A \cap (B \cup B^c)) \cup (A^c \cap (B \cup B^c)) = AB \cup AB^c \cup A^c B \cup A^c B^c$$

Indichiamo tutte queste possibili intersezioni di intersezioni con C_j , dette costituenti: (formano una partizione)

$$C_j = A_1^{e_1} A_2^{e_2} \dots A_n^{e_n}, \quad A_i^{e_i} = \begin{cases} A_i & \text{se } e_i = 1 \\ A_i^c & \text{se } e_i = 0 \end{cases}, \quad j = 1, 2, \dots, 2^n$$

se sono meno di 2^n vuol dire che qualche intersezione è impossibile.

Def. Dati due eventi A e B si dicono logicamente indipendenti quando conosciuto il valore di verità dell'uno il valore di verità dell'altro rimane indeterminato (arbitrario) cioè nessuno dei quattro eventi $AB, AB^c, A^c B, A^c B^c$ (i costituenti) è impossibile. Nel caso di n eventi essi si dicono logicamente indipendenti quando il numero s dei costituenti è uguale a 2^n .

°Esempio

Si ha un mazzo di 40 carte e facciamo due estrazioni senza restituzione. Consideriamo gli

eventi: $A =$ "la 1ª carta estratta è un asso"

$B =$ "la 2ª carta estratta è un asso"

$$P(A) = \frac{\text{casi favorevoli}}{\text{casi possibili}} = \frac{4}{40} = \frac{1}{10}$$

$$A \cup A^c = \Omega$$

$$B = B \cap \Omega = B \cap (A \cup A^c) = BA \cup BA^c \Rightarrow$$

$$P(B) = P(BA \cup BA^c) = P(BA) + P(BA^c) = \frac{4}{40} \cdot \frac{3}{39} + \frac{36}{40} \cdot \frac{4}{39} = \frac{1}{10} \cdot \frac{1}{13} + \frac{1}{10} \cdot \frac{12}{13} = \frac{1}{10}$$

Def. di probabilità classica

gli eventi sono visti come i risultati di un ipotetico esperimento. Se m è il numero di risultati possibili e dato l'evento A , r_A è il numero dei casi favorevoli, cioè il numero di volte che si verifica l'evento, allora $P(A) = \frac{r_A}{m}$. Da notare che secondo questa definizione i risultati dell'esperimento devono essere ugualmente possibili, cioè si devono verificare

con la stessa probabilità. Questo è un difetto di questa definizione perché ancora la (3) probabilità non è stata definita, però si richiede che i risultati siano equiprobabili.

Siamo tre proprietà di cui gode la probabilità secondo questa definizione:

1) $P(A) \geq 0$ perché rapporto di numeri positivi.

2) $P(\Omega) = 1$ perché $P(\Omega) = \frac{m}{m} = 1$

3) Se $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

Infatti se $A \cap B = \emptyset \Rightarrow \varepsilon_{A \cup B} = \varepsilon_A + \varepsilon_B \Rightarrow P(A \cup B) = \frac{\varepsilon_{A \cup B}}{m} = \frac{\varepsilon_A}{m} + \frac{\varepsilon_B}{m} = P(A) + P(B)$.

OSS. Queste tre proprietà valgono per tutte le definizioni di probabilità.

Def. frequentista della probabilità

Supponiamo di poter effettuare un esperimento un numero infinito di volte nelle medesime condizioni. Dopo aver ripetuto l'esperimento m volte osserviamo che l'evento A si è verificato ε_A volte, allora definiamo $\frac{\varepsilon_A}{m}$ = frequenza relativa di successo e la probabilità viene definita come il suo limite $\Rightarrow P(A) = \lim_{m \rightarrow \infty} \frac{\varepsilon_A}{m}$, detta probabilità relativa di successo. Il difetto di questa definizione, come prima, sta nel fatto che quando si dice di ripetere l'esperimento un numero n sempre nelle stesse condizioni si intende sempre con la stessa probabilità di riuscita, ma ancora la probabilità non è stata definita.

22-10-07
LUKE SI.

Sia S un insieme finito di n elementi. Si dice disposizione di classe ε estratta da S , una qualunque ε -pla ordinata di elementi, non necessariamente distinti di S . Due disposizioni differiscono o quando un elemento è presente in una e non nell'altra oppure se sono costituite dai medesimi elementi quando è diverso l'ordine secondo cui essi sono disposti.

Disposizioni con ripetizione di n oggetti su ε posti, $D'_{n,\varepsilon} = \underbrace{n \cdot n \cdot n \cdot \dots \cdot n}_{\varepsilon \text{ volte}} = n^\varepsilon$

Disposizioni semplici (senza ripetizione), $D_{n,\varepsilon} = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-\varepsilon+1)$

Le disposizioni semplici di n oggetti su ε posti prendono il nome di permutazioni,

$$D_{n,n} = P_n = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1 = n!$$

Per convenzione si pone $0! = 1$, è come dire che c'è una sola permutazione dell'insieme vuoto, quella banale cioè l'insieme vuoto stesso.

Si dice combinazione semplice di n oggetti su ε posti una ε -upla di S . Due combinazioni semplici sono distinte quando differiscono almeno per un elemento ma non per l'ordine.

Osserviamo che: $D_{n,\varepsilon} = C_{n,\varepsilon} \cdot \varepsilon!$

Infatti una disposizione ^{semplice} è costituita da una ε -upla ordinata. Se consideriamo tutte le sue permutazioni otteniamo $\varepsilon!$ disposizioni semplici costituite dai medesimi elementi e quindi rappresentano tutte la stessa combinazione perché le combinazioni sono distinte a meno dell'ordine.

$$\text{Si pone } C_{n,\varepsilon} = \binom{n}{\varepsilon} = \frac{D_{n,\varepsilon}}{\varepsilon!} = \frac{n(n-1) \cdot \dots \cdot (n-\varepsilon+1)}{\varepsilon!} = \frac{n!}{\varepsilon!(n-\varepsilon)!} \quad (\text{moltep. e div. per } (n-\varepsilon)!)$$

Il numero di combinazioni con ripetizione di n oggetti su r posti, è: $\binom{n+r-1}{r}$

È ovvio che il numero cercato non dipende dalla natura degli n oggetti che costituiscono l'insieme S ; perciò, si può tranquillamente supporre che $S = \{1, 2, \dots, n\}$. Allora possiamo pensare una r -combinazione $[a_1, a_2, \dots, a_r]$ scritta in modo che:

$$a_1 \leq a_2 \leq \dots \leq a_r \quad (1),$$

dal momento che, alterando l'ordine dei suoi elementi, si ha sempre la stessa r -combinazione. Nelle disuguaglianze (1) può aversi anche il segno uguale, non essendo escluso il verificarsi di ripetizioni. Consideriamo ora gli r numeri:

$$b_1 = a_1 + 0, \quad b_2 = a_2 + 1, \quad b_3 = a_3 + 2, \quad \dots, \quad b_r = a_r + r - 1 \quad (2),$$

i quali, per la (1), sono tutti distinti ed inoltre $1 \leq b_i \leq n+r-1$, $i=1, \dots, r$.

Ha allora, data una r -combinazione estratta da S , ad essa si può fare corrispondere, mediante la (2), l'insieme $\{b_1, b_2, \dots, b_r\}$, che è una r -combinazione semplice estratta dall'insieme $\{1, 2, \dots, n+r-1\}$. Ciò significa che le r -combinazioni (con ripetizioni) di un n -insieme sono tante quante le r -combinazioni semplici di un $(n+r-1)$ -insieme $\Rightarrow$ segue la Tesi.

Osserviamo che il numero di sottoinsiemi di r elementi che si possono formare con un insieme di $n \geq r$ elementi è dato da $\binom{n}{r}$.

Osserviamo che si pone $\binom{n}{0} = 1$; è come dire che c'è un solo sottoinsieme vuoto in un insieme di n elementi.

Calcoliamo $(a+b)^n$. Se lo sviluppiamo con le solite regole dell'algebra, dobbiamo ^{considerare} tutti i

termini del tipo $a^{n-r} \cdot b^r$ con $0 \leq r \leq n$. Si può dimostrare facilmente che,

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

Infatti, fissato r , i termini del tipo $a^{n-r} \cdot b^r$ si ottengono scegliendo in tutti i modi possibili gli r fattori del prodotto $(a+b)(a+b) \dots (a+b)$ da cui si prefera b , mentre dai rimanenti $n-r$

si prefera a . Per la proprietà commutativa del prodotto di numeri reali, tali scelte danno luogo allo stesso risultato e sono in numero di $\binom{n}{r}$.

NOTA Se scegliamo $a=b=1 \Rightarrow (1+1)^n = \sum_{r=0}^n \binom{n}{r} 1^{n-r} \cdot 1^r = \sum_{r=0}^n \binom{n}{r} = 2^n \quad (*)$

Osserviamo che $\binom{n}{r}$ è il numero di sottoinsiemi di cardinalità r di un insieme di cardinalità $n \Rightarrow \sum_{r=0}^n \binom{n}{r} =$ il numero di tutti i sottoinsiemi, di ~~ogni~~ cardinalità r , $0 \leq r \leq n$, di un insieme di cardinalità n ; dall'algebra sappiamo che tale numero è 2^n , come abbiamo visto in $(*)$.

Dati n interi $r_i \geq 0$ tali che $r_1 + r_2 + \dots + r_n = r$, il numero di modi in cui r palline si possono ripartire in n scatole in modo che la scatola k contenga r_k palline, è dato dal coefficiente

multinomiale $\binom{r}{r_1, r_2, \dots, r_n} = \frac{r!}{r_1! \cdot r_2! \cdot \dots \cdot r_n!}$

Infatti, $\binom{r}{r_1}$ sono r_1 palline da mettere nella scatola 1, $\binom{r-r_1}{r_2}$ sono r_2 palline da mettere nella seconda, ..., allora

$$\binom{n}{r_1} \binom{r-r_1}{r_2} \binom{r-r_1-r_2}{r_3} \dots \binom{r-r_1-r_2-\dots-r_{m-1}}{r_m} = \frac{n!}{r_1! r_2! \dots r_m!} = \binom{r}{r_1, r_2, \dots, r_m} \quad (5)$$

Potenza di un polinomio:

$$(x_1 + x_2 + \dots + x_r)^n = \sum \binom{n}{n_1, n_2, \dots, n_r} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}, \text{ tale che } n_1 + n_2 + \dots + n_r = n$$

Formule di De Morgan

$$(A \cup B)^c = A^c \cap B^c$$



$$(A \cap B)^c = A^c \cup B^c$$



• Proprietà additiva

$$\text{Se } A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

Teorema generale

$$① P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



$$② P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Sinfatti, la ② si può dimostrare utilizzando la ①:

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

• PARADOSSO DI BERNSTEIN

Consideriamo i due eventi:

"A" = "In 4 lanci la faccia 6 esce almeno una volta"

B = "In 4 lanci di una coppia di dadi, la coppia (6,6) esce almeno una volta"

Bernstein sosteneva che $P(A) = P(B)$, perché:

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

$$P(B) = \frac{1}{36} + \frac{1}{36} + \dots + \frac{1}{36} = \frac{24}{36} = \frac{2}{3}$$

d'errore sta nel fatto che se facessi 8 lanci $\Rightarrow P(A) = \frac{8}{6} > 1$ assurdo, cioè non sono eventi incompatibili. Infatti:

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

$$A^c = \text{"la faccia 6 non esce mai in 4 lanci"} \Rightarrow P(A) = 1 - \frac{5^4}{6^4}$$

$$P(B) = 1 - P(B^c) = 1 - \frac{35^{24}}{36^{24}}$$

Calcoliamo le stesse probabilità come unione di eventi incompatibili:

$A_c =$ "In 4 lanci il 6 esce esattamente r volte", $r = 0, 1, \dots, 4$

A_0, A_1, A_2, A_3, A_4 sono eventi incompatibili $\Rightarrow A = A_0 \cup A_1 \cup A_2 \cup A_3 \cup A_4 \Rightarrow$

$$P(A) = P(A_0) + P(A_1) + P(A_2) + P(A_3) + P(A_4) = \binom{4}{1} \frac{5^3}{6^4} + \binom{4}{2} \frac{5^2}{6^4} + \binom{4}{3} \frac{5}{6^4} + \binom{4}{4} \frac{5^0}{6^4} =$$

$$= \sum_{r=1}^4 \binom{4}{r} \frac{5^{4-r}}{6^4} = \sum_{r=1}^4 \binom{4}{r} \frac{5^{4-r}}{6^4} + \frac{\binom{4}{0} 5^4}{6^4} - \frac{\binom{4}{0} 5^4}{6^4} = \sum_{r=0}^4 \binom{4}{r} \frac{5^{4-r}}{6^4} - \frac{5^4}{6^4} =$$

$$= \frac{(5+1)^4}{6^4} - \frac{5^4}{6^4} = 1 - \frac{5^4}{6^4}$$

Una famiglia ^{discreti, E} si dice Algebra se su di essa sono definite due operazioni di unione e negazione ed è chiusa rispetto ad esse, cioè:

$$1) A, B \in \mathcal{E} \Rightarrow A \cup B \in \mathcal{E}$$

$$2) A \in \mathcal{E} \Rightarrow A^c \in \mathcal{E}$$

Da questo segue:

$$1) \Omega \in \mathcal{E} \text{ Infatti } A \in \mathcal{E} \Rightarrow A^c \in \mathcal{E} \Rightarrow A \cup A^c \in \mathcal{E} \parallel \Rightarrow \Omega \in \mathcal{E}$$

$$2) \emptyset \in \mathcal{E} \text{ Infatti } \Omega^c = \emptyset$$

$$3) \text{ Se } A, B \in \mathcal{E} \Rightarrow A \cap B \in \mathcal{E}$$

$$\text{Infatti: } A, B \in \mathcal{E} \Rightarrow A^c, B^c \in \mathcal{E} \Rightarrow A^c \cup B^c \in \mathcal{E} \Rightarrow (A \cap B)^c \in \mathcal{E}$$

$$\text{Per De Morgan } (A^c \cup B^c)^c = A \cap B \parallel \Rightarrow A \cap B \in \mathcal{E}$$

Una tale famiglia esist. costituita si dice σ -Algebra se è infinita $\rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{E}$ = unione numerabile

Def. di Probabilità Assiomatica dovuta a de Finetti o oggettiva dovuta a Kolmogorov, (perché secondo questa Teoria la probabilità è qualcosa di intrinseco all'evento).

Sia $\mathcal{E}$ una σ -algebra Allora una funzione $P: \mathcal{E} \rightarrow \mathbb{R}$ si dice probabilità se:

$$1) P(A) \geq 0, \forall A \in \mathcal{E}$$

$$2) P(\Omega) = 1$$

$$3) A_1, \dots, A_n, \dots \text{ tali che } A_i \cap A_j = \emptyset, i \neq j \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

NOTA. Secondo questa def. le proprietà della probabilità definita prima ora vengono date come assiomi.

29-10-07
LUNEDÌ

Def. di probabilità oggettiva

La probabilità di un evento è il grado di fiducia che ha l'individuo nel verificarsi dello stesso (De Finetti). Ci chiediamo se per De Finetti alla probabilità di un evento possiamo associare un numero qualsiasi. No infatti:

Sia dato un evento E , indichiamo con $p = P(E)$ il grado di fiducia. Si dice che si effettua una scommessa di quota p su un evento E se, versando una somma $p \cdot S$ (con $S \neq 0$ arbitrario), si ciere un importo S solo se E si verifica (scommessa vinta) e niente in caso contrario, perdendo la puntata $p \cdot S$.

Definiamo guadagno $G = S \cdot (IEI - p)$

$$\text{Nel caso in cui } E \text{ si verifica } \Rightarrow IEI = 1 \Rightarrow G_1 = S(1-p)$$

$$\text{Nel caso in cui } E \text{ non si verifica } \Rightarrow IEI = 0 \Rightarrow G_2 = -p \cdot S$$

$$\text{Coerenza: } \min_{G_1, S, p} G \cdot \max_{G_2, S, p} G \leq 0 \quad \forall S$$

Una definizione di probabilità si dice coerente quando i guadagni G_1 e G_2 sono di segno opposto, cioè se: $G_1 \cdot G_2 \leq 0, \forall S \in \mathbb{R}, 0 < p \leq 1$ se $\min \{G_1, i=1, \dots, n\} \cdot \max \{G_2, i=1, \dots, n\} \leq 0$

Se una def. di probabilità è coerente valgono le seguenti proprietà:

$$1) 0 \leq p \leq 1$$

$$\text{È coerente } \Rightarrow G_1 \cdot G_2 \leq 0 \Leftrightarrow -pS^2(1-p) \leq 0 \text{ . Se } S=0 \Rightarrow \text{non gioco} \Rightarrow S \neq 0 \Rightarrow$$

$$\Rightarrow -p(1-p) \leq 0 \Leftrightarrow p(1-p) \geq 0 \Leftrightarrow 0 \leq p \leq 1$$

2) $p = P(\Omega) = 1$

(2)

Ω si verifica sempre $\Rightarrow |\Omega| = 1$ sempre $\Rightarrow G = S(1-p)$ e l'unico guadagno possibile
 Per la coerenza $S(1-p) = 0, \forall S \in \mathbb{R} \Rightarrow 1-p = 0 \Rightarrow p = 1$

Def di probabilità soggettiva Si dice probabilità di un evento E la misura del grado di fiducia in E espressa da un numero reale $p = P(E)$ tale che una scommessa di quota p su E sia coerente.

3) Vale la proprietà di additività Se $AB = \emptyset \Rightarrow p(A \cup B) = p(A) + p(B)$

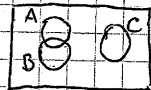
Questa definizione di probabilità è la più generale di tutte ed è quella che si avvicina di più al pensiero comune. Con questa def la probabilità si estende a tutte le famiglie di eventi e non solo alle $\mathcal{F}-A$.

Nel caso in cui si hanno n eventi $A_1, A_2, \dots, A_n$ e $p_i = P(A_i), i=1, \dots, n$, allora

$G = \sum_i S_i (|A_i| - p_i)$ // guadagno totale //

Esempio

Siano dati 3 eventi A, B, C tali che $BC = \emptyset, AC = \emptyset$



I costituenti non sono $2^3 = 8$ ma 5 perché alcune intersezioni sono impossibili, e sono:

$1 = |A \cap B \cap C^c| \rightarrow G_1 = S_1(1-p_1) + S_2(1-p_2) - p_3 S_3$

$1 = |A \cap B^c \cap C^c| \rightarrow G_2 = S_1(1-p_1) - S_2 p_2 - S_3 p_3$

$1 = |A^c \cap B \cap C^c| \rightarrow G_3 = -p_1 S_1 + S_2(1-p_2) - p_3 S_3$

$1 = |A^c \cap B^c \cap C^c| \rightarrow G_4 = -p_1 S_1 - p_2 S_2 + S_3(1-p_3)$

$1 = |A^c \cap B \cap C| \rightarrow G_5 = -p_1 S_1 - p_2 S_2 - p_3 S_3$

$\{G_i\}_{i=1}^5$ sono tutti i possibili guadagni (al variare degli indicatori)

Def Una definizione di probabilità $(p_1, \dots, p_n)$ su $\{A_1, A_2, \dots, A_n\}$ è coerente $\Leftrightarrow \min G_i = \max G_i \leq 0, \forall S_1, \dots, S_n \in \mathbb{R}$.

Valutazione di probabilità

Dati n eventi $\{A_1, \dots, A_n\}$ a cui associo la probabilità $p_1, p_2, \dots, p_n, p_i = P(A_i), i=1, \dots, n$

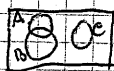
Considero i costituenti $C_1, C_2, \dots, C_m$ che al massimo sono 2^n e ad ognuno di essi associo la probabilità $\lambda_1, \dots, \lambda_m$ si ha:

$$\begin{cases} \sum_{C_i \subseteq A_j} \lambda_i = p_j, j=1, \dots, n \\ \lambda_i \geq 0, i=1, \dots, m \\ \sum_{i=1}^m \lambda_i = 1 \end{cases} \Leftrightarrow A_i = \bigcup_{C_j \subseteq A_i} G_j \Rightarrow p_i = P(A_i) = \sum_{C_j \subseteq A_i} \lambda_j$$

(1)

Def Una valutazione di probabilità è coerente $\Leftrightarrow$ il sistema ha una soluzione.

Ritorniamo l'esempio di prima



$p_1 = P(A), p_2 = P(B), p_3 = P(C)$

Ho 5 costituenti $C_1, \dots, C_5$ a cui associo la probabilità $\lambda_1, \lambda_2, \dots, \lambda_5$

Scrivo A come unione di costituenti $\Rightarrow A = A \cap B \cap C^c \cup A \cap B^c \cap C^c$, sono incompatibili $\Rightarrow$

$p(A) = P(A \cap B \cap C^c \cup A \cap B^c \cap C^c) = P(A \cap B \cap C^c) + P(A \cap B^c \cap C^c) = \lambda_1 + \lambda_2 \Rightarrow$

$\Rightarrow p_1 = \lambda_1 + \lambda_2$, ecco come nasce la prima equazione del sistema.

$$\Rightarrow \begin{cases} p_1 = \lambda_1 + \lambda_2 \\ p_2 = \lambda_1 + \lambda_3 \\ p_3 = \lambda_1 \\ \lambda_i \geq 0, i=1,2,3 \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1 \end{cases}$$

Consideriamo il caso particolare che gli n eventi cost. formano una partizione.

TEO. Siano dati n eventi $\{H_1, H_2, \dots, H_n\}$ che ~~costituiscono~~ formano una partizione.

Consideriamo la ~~valutazione~~ ^{definitiva} di probabilità $\{p_1, p_2, \dots, p_n\}$, $p_i = p(H_i)$, $i=1, \dots, n$. Tale ~~valutazione~~ ^{definitiva} è coerente $\Leftrightarrow \sum_{i=1}^n p_i = 1$.

• Dim. $\Rightarrow$ H_p è coerente $\Leftrightarrow \sum_{i=1}^n p_i = 1$

Per ipotesi $\min G \cdot \max G \leq 0, \forall S_1, S_2, \dots, S_n \in R$, dove $G = \sum_{i=1}^n S_i (|H_i| - p_i)$.

Poiché gli H_i formano una partizione i costituenti sono così formati:

$$C_1 = H_1 H_2^c H_3^c \dots H_n^c$$

$$C_2 = H_1^c H_2 H_3^c H_4^c \dots H_n^c$$

$$C_n = H_1^c \dots H_{n-1}^c H_n$$

$$g_1 = S_1(1-p_1) - \sum_{i=2}^n p_i S_i = S_1 - p_1 S_1 - \sum_{i=2}^n p_i S_i = S_1 - \sum_{i=1}^n p_i S_i$$

$$g_2 = S_2(1-p_2) - \sum_{i \neq 2} p_i S_i = S_2 - \sum_{i=1}^n p_i S_i$$

$$g_n = S_n - \sum_{i=1}^n p_i S_i$$

Per ipotesi $\min G \cdot \max G \leq 0, \forall S_1, \dots, S_n \Rightarrow$ in particolare vale per $S_1 = S_2 = \dots = S_n = 1 \Rightarrow$

$$\Rightarrow g_1 = 1 - \sum_{i=1}^n p_i = g_2 = \dots = g_n$$

$\Rightarrow$ il guadagno non è più aleatorio ma è ben preciso ed $\min G = \max G \Rightarrow$

$$\min G \cdot \max G \leq 0 \Leftrightarrow 1 - \sum_{i=1}^n p_i = 0 \Rightarrow \sum_{i=1}^n p_i = 1 \text{ c.v.d.}$$

• Dim. $\Leftarrow$ $H_p \sum_{i=1}^n p_i = 1$ $T_0(p_1, \dots, p_n)$ è coerente $\Leftrightarrow \min G \cdot \max G \leq 0, \forall S_1, \dots, S_n$

Abbiamo visto che $g_i = S_i - \sum_{i=1}^n p_i S_i, i=1, \dots, n$

Dati $m_1, \dots, m_n \in R$ e $a_1, \dots, a_n \in R: a_i \geq 0 \forall i$ e $\sum_{i=1}^n a_i = 1$ allora

$\bar{m} = a_1 m_1 + a_2 m_2 + \dots + a_n m_n$ si dice combinazione lineare convessa. (vedi pg 9 lemma 1)

Le combinazioni lineari convesse godono della seguente proprietà:

$$\min \{m_1, m_2, \dots, m_n\} \leq \bar{m} \leq \max \{m_1, m_2, \dots, m_n\}$$

Consideriamo una ^{convessa} combinazione lineare dei guadagni $\Rightarrow$

$$\bar{g} = p_1 g_1 + p_2 g_2 + \dots + p_n g_n = p_1 (S_1 - \sum_{i=1}^n p_i S_i) + \dots + p_n (S_n - \sum_{i=1}^n p_i S_i) =$$

$$= \sum_{i=1}^n p_i S_i - \left(\sum_{i=1}^n p_i \right) \left(\sum_{i=1}^n p_i S_i \right) = \sum_{i=1}^n p_i S_i - \sum_{i=1}^n p_i S_i = 0 \Rightarrow \bar{g} = 0$$

$$\text{ma } \min \{g_1, \dots, g_n\} \leq \bar{g} \leq \max \{g_1, \dots, g_n\} \Rightarrow \min G \cdot \max G \leq 0 \quad \forall$$

$$S_1, \dots, S_n \in R \Rightarrow (p_1, \dots, p_n) \text{ è coerente.}$$

c.v.d.

mettendo in evidenza $\sum_{i=1}^n p_i S_i$

⇒ il numero aleatorio $X = 1A - 2B + 3C$ non è in forma normale perché $X \notin \{1, -2, 3\}$ ma $X \in \{-2, -1, 0, 1, 3\}$, infatti consideriamo i costituenti:

$$ABC^c \rightarrow X = 1$$

$C_X = \text{codominio di } X$

$$A^c B C^c \rightarrow X = -2$$

$$A^c B^c C \rightarrow X = 1$$

$$A^c B^c C^c \rightarrow X = 3$$

$$A^c B^c C^c \rightarrow X = 0$$

Allora ogni numero aleatorio semplice che non è scritto in forma normale può essere riscritto in modo che sia in forma normale ⇒

$$X = -1 |ABC^c| - 2 |A^c B C^c| + 1 |A^c B^c C| + 3 |A^c B^c C^c| + 0 |A^c B^c C^c|$$

PREVISIONE DI UN NUMERO ALEATORIO

Consideriamo il numero aleatorio $X = c_1 |A_1| + c_2 |A_2| + \dots + c_n |A_n|$, e $p_i = P(A_i)$, $i=1, \dots, n$.

Si definisce previsione del numero aleatorio X , $P[X] = \sum_{i=1}^n c_i \cdot p_i$, ~~che~~ viene anche detta valore atteso o media. $\parallel X = \sum c_i I_{A_i}, EX = \sum c_i E I_{A_i} = \sum c_i P(A_i) = \parallel$

Se conosco il ^{immagine} ~~codominio~~ di X , $C_X = \{b_1, b_2, \dots, b_m\}$ e $P(X=b_i)$ è la probabilità che $X=b_i$

$$\Rightarrow P(X) = \sum_{i=1}^m b_i \cdot P(X=b_i)$$

Consideriamo la quantità $aX+b$ dove X è un numero aleatorio e $a, b \in \mathbb{R} \Rightarrow$

$$P(aX+b) = a P(X) + b \Rightarrow \text{la previsione è lineare}$$

$$\text{Infatti: } P(aX+b) = P(aX) + P(b) = a P(X) + b$$

Si può vedere come numero aleatorio di un sbo esento che ha $p=1$.

Dati n aleatori $x_1, x_2, \dots, x_n$ sia $X = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ una loro combinazione

lineare $\Rightarrow P(X) = a_1 P(x_1) + a_2 P(x_2) + \dots + a_n P(x_n)$, segue dalla def di P .

Infatti dati due numeri aleatori $X = a_1 |A_1| + \dots + a_n |A_n|$ e $Y = b_1 |B_1| + \dots + b_m |B_m|$.

$\Rightarrow$ si definisce $cX+dY = c a_1 |A_1| + \dots + c a_n |A_n| + d b_1 |B_1| + \dots + d b_m |B_m|$.

Dati X e Y (quelli di prima) sia $p_i' = P(A_i)$, $i=1, \dots, n$ e $p_j'' = P(B_j)$, $j=1, \dots, m$.

$$\Rightarrow P(cX+dY) = \sum_{i=1}^n c a_i p_i' + \sum_{j=1}^m d b_j p_j'' = c \sum_{i=1}^n a_i p_i' + d \sum_{j=1}^m b_j p_j'' =$$

$$= c \cdot P(X) + d \cdot P(Y)$$

Nel caso particolare in cui $X = |A|$ in cui $C_X = \{0, 1\}$ si ha:

$$P(X=0) = P(A^c) = q$$

$$P(X=1) = P(A) = p, \quad p+q=1$$

$$\Rightarrow P(X) = 1 \cdot p + 0 \cdot q = p$$

Cioè in questo caso la previsione si riduce alla probabilità (definita).

Esempio

" A_i " = "Nel lancio di un dado esce la faccia i "

$$\text{Sia } X = 1 \cdot |A_1| + 2 \cdot |A_2| + 3 \cdot |A_3| + 4 \cdot |A_4| + 5 \cdot |A_5| + 6 \cdot |A_6| \cong$$

$$P(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2}$$

(11)

• Sia $X \in \{x_1, x_2, \dots, x_n\} \Rightarrow \min\{x_1, x_2, \dots, x_n\} \leq P(X) \leq \max\{x_1, x_2, \dots, x_n\}$

Infatti se $(p_1, \dots, p_n)$ sono le probabilità $p_i = P(X=x_i)$, $i=1, 2, \dots, n$

Poiché $\sum_{i=1}^n p_i = 1$, (viene preso in considerazione tutto l'eventualità di X) e $P(X) = \sum_{i=1}^n p_i \cdot x_i$

$\Rightarrow P(X)$ è una combinazione lineare convessa $\Rightarrow$ si ha l'attesi.

(Si è supposto che X è in forma normale, se non lo è si può scriverlo in forma normale).

VARIANZA DI UN NUMERO ALEATORIO

Si definisce varianza di un numero aleatorio X , detto anche errore quadratico medio,

$$VAR(X) = P[(X-m)^2], \quad m = P(X)$$

Si dimostra che $VAR(X) = P(X^2) - P^2(X) = P(X^2) - m^2$, infatti:

$$\begin{aligned} VAR(X) &= P[(X-m)^2] = P[X^2 - 2mX + m^2] = P(X^2) - 2mP(X) + m^2 = \\ &= P(X^2) - 2 \cdot P(X) \cdot P(X) + (P(X))^2 = P(X^2) - 2P^2(X) + P^2(X) = P(X^2) - P^2(X) \end{aligned}$$

Qu che $VAR(X) \geq 0$

Infatti sia X un numero aleatorio e sia $X \in \{x_1, x_2, \dots, x_n\}$ e $p_i = P(X=x_i)$.

$$VAR(X) = P((X-m)^2)$$

$(X-m)^2$ è un numero aleatorio semplice e il suo codominio è $\{(x_1-m)^2, (x_2-m)^2, \dots, (x_n-m)^2\}$

Osserviamo che $(X-m)^2 = (x_1-m)^2 \Leftrightarrow X=x_1 \Rightarrow$

$$\Rightarrow P[(X-m)^2 = (x_1-m)^2] = P(X=x_1) = p_1$$

$$\Rightarrow VAR(X) = P((X-m)^2) = \underbrace{p_1}_{\downarrow} \cdot \underbrace{(x_1-m)^2}_{\downarrow} + \underbrace{p_2}_{\downarrow} \cdot \underbrace{(x_2-m)^2}_{\downarrow} + \dots + \underbrace{p_n}_{\downarrow} \cdot \underbrace{(x_n-m)^2}_{\downarrow} \Rightarrow$$

$$\Rightarrow VAR(X) \geq 0$$

Il valore dell'errore $VAR(X)$ è tanto più grande quanto più la distribuzione è dispersa, cioè quanto più i valori x_i sono lontani dal valore atteso $P(X)$, in altre parole la precisione $P(X)$ è tanto più significativa quanto più piccola è la varianza.

Tanto più piccola è la varianza tanto il numero aleatorio è più vicino alla sua precisione.

Esempio.

Siano dati 3 numeri aleatori X, Y, Z :

$$X: \begin{array}{ccc} \frac{1}{2} & & \frac{1}{2} \\ -1 & & 1 \end{array} \rightarrow P(X) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$$

$$Y: \begin{array}{ccccc} \frac{1}{4} & \frac{1}{4} & & \frac{1}{4} & \frac{1}{4} \\ -2 & -1 & 0 & 1 & 2 \end{array} \rightarrow P(Y) = -2 \cdot \frac{1}{4} - 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = 0$$

$$Z: \begin{array}{ccc} \frac{1}{2} & & \frac{1}{2} \\ -2 & & 2 \end{array} \rightarrow P(Z) = -2 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = 0$$

$$VAR(X) = P[(X-0)^2] = (-1-0)^2 \cdot \frac{1}{2} + (1-0)^2 \cdot \frac{1}{2} = 1$$

$$VAR(Y) = (-2-0)^2 \cdot \frac{1}{4} + (-1-0)^2 \cdot \frac{1}{4} + (1-0)^2 \cdot \frac{1}{4} + (2-0)^2 \cdot \frac{1}{4} = 1 + \frac{1}{4} + \frac{1}{4} + 1 = \frac{5}{2}$$

$$VAR(Z) = (-2-0)^2 \cdot \frac{1}{2} + (2-0)^2 \cdot \frac{1}{2} = 2 + 2 = 4$$

$\Rightarrow VAR(X) < VAR(Y) < VAR(Z) \Rightarrow$ ~~quello~~ la $P(X)$ è la più indicativa del proprio numero aleatorio rispetto a $P(Y)$ e $P(Z)$ per Y e Z .

Sia $c \in \mathbb{R} \Rightarrow P(c) = c$, infatti: sia $X = c \cdot 1_{\Omega} \Rightarrow x \in \{c\}$

$$\Rightarrow P(x) = P(c) = c \cdot p(\Omega) = c \cdot 1 = c$$

$$\text{VAR}(c) = 0$$

Sia $X = c \cdot 1_{\Omega} \Rightarrow x \in \{c\} \Rightarrow P(c) = c$

$$\text{VAR}(c) = P[(c - m)^2] = P[(c - c)^2] = P(0) = 0$$

Prossimo che la varianza è non lineare cioè $\text{VAR}(aX + b) \neq a \cdot \text{VAR}(X) + \text{VAR}(b)$

Infatti: $\text{VAR}(aX + b) = P[(aX + b - P(aX + b))^2] = P[(aX + b - aP(X) - b)^2] =$
 $= P[(aX - aP(X))^2] = P[a^2 X^2 + \underbrace{a^2 P^2(X)}_{\text{costante}} - 2a^2 X \cdot \underbrace{P(X)}_{\text{costante}}] =$ per la proprietà di linearità della speranza e ricordando che $P(X)$ è una costante e non un numero aleatorio
 $= a^2 P(X^2) + a^2 P^2(X) - 2a^2 P(X) \cdot P(X) =$
 $= a^2 P(X^2) + a^2 P^2(X) - 2a^2 P^2(X) = a^2 P(X^2) - a^2 P^2(X) =$
 $= a^2 (P(X^2) - P^2(X)) = a^2 \cdot \text{VAR}(X)$

Si dice scarto quadratico medio o deviazione standard del numero aleatorio X la quantità $\sigma_X = \sqrt{\text{VAR}(X)}$

Calcoliamo $P(X^2)$

Sia X un numero aleatorio tale che $X \in \{x_1, x_2, \dots, x_n\} \Rightarrow X^2 \in \{x_1^2, x_2^2, \dots, x_n^2\}$

$$\Rightarrow P(X^2) = x_1^2 \cdot p(X^2 = x_1^2) + x_2^2 \cdot p(X^2 = x_2^2) + \dots + x_n^2 \cdot p(X^2 = x_n^2)$$

$$X^2 = x_1^2 \Leftrightarrow X = x_1 \Rightarrow p(X^2 = x_1^2) = p(X = x_1) \Rightarrow$$

$$P(X^2) = x_1^2 \cdot p(X = x_1) + \dots + x_n^2 \cdot p(X = x_n)$$

Esempio: $X \in \{-2, -1, 1, 2\} \Rightarrow X^2 \in \{1, 4\}$

$$X^2 = 1 \Leftrightarrow (X = 1) \vee (X = -1)$$

$$X^2 = 4 \Leftrightarrow (X = 2) \vee (X = -2)$$

$$\Rightarrow P(X^2 = 1) = p((X = 1) \vee (X = -1)) = p(X = 1) + p(X = -1)$$

$$P(X^2 = 4) = p(X = 2) + p(X = -2)$$

$$P(X^2) = 1 \cdot p(X^2 = 1) + 4 \cdot p(X^2 = 4) = 1 \cdot p(X = 1) + 1 \cdot p(X = -1) + 4 \cdot p(X = 2) + 4 \cdot p(X = -2)$$

$$= 1^2 \cdot p(X = 1) + (-1)^2 \cdot p(X = -1) + 2^2 \cdot p(X = 2) + (-2)^2 \cdot p(X = -2)$$

12-11-02
LUKEB

Siano X e Y due numeri aleatori $\Rightarrow$

$$\text{VAR}(X + Y) = P[(X + Y - m_{X+Y})^2] =$$
 dove $m_X = P(X)$, $m_Y = P(Y)$, $m_{X+Y} = P(X+Y) = m_X + m_Y$
 $= P[(X + Y - m_X - m_Y)^2] = P[(X - m_X) + (Y - m_Y)]^2 =$
 $= P[(X - m_X)^2 + (Y - m_Y)^2 + 2(X - m_X) \cdot (Y - m_Y)] =$
 $= P[(X - m_X)^2] + P[(Y - m_Y)^2] + 2P[(X - m_X) \cdot (Y - m_Y)] =$
 $= \text{VAR}(X) + \text{VAR}(Y) + 2P[(X - m_X) \cdot (Y - m_Y)]$ $\Rightarrow$ da VAR non è lineare

Si definisce covarianza di due numeri aleatori X e Y ,

$$\text{COV}(X, Y) = P[(X - m_X)(Y - m_Y)]$$

Ad differenza di P e VAR , la COV riguarda più numeri aleatori

$$\text{NOTA: } \text{VAR}(X_1 + X_2 + \dots + X_n) = \text{VAR}(X_1) + \dots + \text{VAR}(X_n) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{COV}(X_{i+1}, X_{i+2})$$

dove la sommatoria è estesa a tutte le combinazioni a due a due degli indici (13)
 $1, 2, \dots, m$.

La covarianza indica la tendenza che ~~hanno~~ hanno due valori ad associarsi.

Se a valori piccoli (o grandi) di X si associano valori piccoli (o grandi) di Y e si aspetta una covarianza positiva. Piccolo o grande è riferito rispetto alla previsione.

valori grandi di X vuol dire maggiori di $P(X) \Rightarrow X - P(X) > 0$; piccoli vuol dire minori di $P(X) \Rightarrow X - P(X) < 0$. Quindi se sono tutte e due grandi ~~la covari~~ o tutte e due piccole la covarianza è ~~positiva~~ positiva, mentre se una si associano valori piccoli e grandi e si aspetta covarianza negativa. Se $\text{cov}(X, Y) = 0 \Rightarrow X$ e Y si dicono incorrelati.

DISEGUAGLIANZA DI MARKOV

H_p Sia X un numero aleatorio non negativo, $X \geq 0$,

$$m = P(X), \quad d \in \mathbb{R}: d \geq m \Rightarrow P(X) < d \Rightarrow m < d$$

Tesi: $P(X \geq d) \leq \frac{m}{d}$ (Se $d \leq m$ sarebbe banale perché $\frac{m}{d} \geq 1 \rightarrow P(X \geq d) \leq 1$ è sempre vero)

Dim.

Consideriamo il codominio di X , $C_X = \{x_1, x_2, \dots, x_n\}$ tale che $x_i \leq x_{i+1}$

Osserviamo che non può essere $d < x_1$ perché vorrebbe dire $P(X) < x_1$ ^{impossibile} ~~perché~~ ^{per ipotesi} $d > P(X)$.

$$\text{Se } d > x_n \Rightarrow P(X \geq d) = 0 \leq \frac{m}{d}$$

Rimane il caso

$$\text{Oppure } x_k < d \leq x_{k+1}$$

Sia $p_i = P(X = x_i)$, $i = 1, \dots, m$.

^{valori di C_X sono tutti non negativi}

$$m = P(X) = x_1 p_1 + x_2 p_2 + \dots + x_{k+1} p_{k+1} + \dots + x_n p_n \geq x_{k+1} p_{k+1} + x_{k+2} p_{k+2} + \dots + x_n p_n \geq$$

$$\geq d p_{k+1} + d p_{k+2} + \dots + d p_n = d (p_{k+1} + p_{k+2} + \dots + p_n) = d \cdot P(X \geq d), \text{ infatti:}$$

$$(d \leq x_{k+1} \leq x_{k+2} \leq \dots \leq x_n)$$

$$(X \geq d) = (X = x_{k+1}) \cup (X = x_{k+2}) \cup \dots \cup (X = x_n), \text{ unione di eventi incompatibili } \Rightarrow$$

$$P(X \geq d) = p_{k+1} + p_{k+2} + \dots + p_n$$

$$\Rightarrow m \geq d \cdot P(X \geq d) \Rightarrow P(X \geq d) \leq \frac{m}{d} \quad \text{c.v.d.}$$

DISEGUAGLIANZA DI CEBICEV

H_p X numero aleatorio, $m = P(X)$, $\sigma = \sqrt{\text{VAR}(X)}$, $k \geq 1$

Tesi: $P(|X - m| \geq k\sigma) \leq \frac{1}{k^2}$ (Se $k=1$ è banale)

Dim.

$$|X - m| \geq k\sigma \Leftrightarrow (X - m)^2 \geq k^2 \sigma^2$$

Considero il numero aleatorio $Y = (X - m)^2$

$$P(Y) = m_Y = P((X - m)^2) = \text{VAR}(X) = \sigma^2 \Rightarrow k^2 \sigma^2 > m_Y \Rightarrow \text{Per Markov}$$

$$P((X - m)^2 \geq k^2 \sigma^2) \leq \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2}$$

$$\Downarrow$$

$$P(|X - m| \geq k\sigma) \leq \frac{1}{k^2} \quad \text{c.v.d.}$$

PRINCIPIO DI INCLUSIONE ED ESCLUSIONE

Siano dati n eventi $A_1, A_2, \dots, A_n$.

Si ha: $P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(A_1^c \cap A_2^c \cap \dots \cap A_n^c)$

Calcoliamo $P(A_1^c \cap A_2^c \cap \dots \cap A_n^c)$

Riesprimiamo che la previsione dell'indicatore coincide con la probabilità dell'evento $\Rightarrow$

$$P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = P(1 - A_1 | 1 - A_2 | \dots | 1 - A_n) = P((1 - A_1)(1 - A_2) \dots (1 - A_n)) =$$

$$= P\left[1 - (|A_1| + |A_2| + \dots + |A_n|) + \sum_{i,j} |A_i| \cdot |A_j| - \sum_{i,j,k} |A_i| \cdot |A_j| \cdot |A_k| + \dots + (-1)^n |A_1| \cdot |A_2| \cdot \dots \cdot |A_n|\right] =$$

$$= P\left[1 - \sum_i |A_i| + \sum_{i,j} |A_i A_j| - \sum_{i,j,k} |A_i A_j A_k| + \dots + (-1)^n |A_1 A_2 \dots A_n|\right] = \text{Per la previsione della previsione}$$

$$= 1 - \sum_i P(A_i) + \sum_{i,j} P(A_i A_j) - \sum_{i,j,k} P(A_i A_j A_k) + \dots + (-1)^n P(A_1 A_2 \dots A_n) =$$

$$= 1 - \sum_i P(A_i) + \sum_{i,j} P(A_i A_j) - \sum_{i,j,k} P(A_i A_j A_k) + \dots + (-1)^n P(A_1 A_2 \dots A_n)$$

$$\Rightarrow P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = 1 - 1 + \sum_i P(A_i) - \sum_{i,j} P(A_i A_j) + \sum_{i,j,k} P(A_i A_j A_k) - \dots + (-1)^{n+1} P(A_1 A_2 \dots A_n) \Rightarrow$$

$$\Rightarrow P(A_1 \cup \dots \cup A_n) = \sum_i P(A_i) - \sum_{i,j} P(A_i A_j) + \sum_{i,j,k} P(A_i A_j A_k) - \dots + (-1)^{n+1} P(A_1 A_2 \dots A_n)$$

Per $n=2$ si ha il Teorema delle probabilità totali.

Es. Gioco delle concordanze (cappelli pg 13 Chap 1)

Il problema si può porre considerando n palline numerate da 1 a n che si estraggono senza restituzione da un'urna. Si dice concordanza l'evento "alla i -esima estrazione esce la pallina con il numero i , $i=1, \dots, n$ ".

La probabilità di avere 1 concordanza $P(A_1) = \frac{(n-1)!}{n!}$

$$\begin{matrix} n & & n & 2 & n & \text{è } P(A_1 A_2) = \frac{(n-2)!}{n!} & \text{(Per due elementi precisi)} \end{matrix}$$

$$\begin{matrix} n & & n & 3 & n & \text{è } P(A_1 A_2 A_3) = \frac{(n-3)!}{n!} & \text{(Per tre elementi ben precisi)} \end{matrix}$$

$$\begin{matrix} n & & n & h & \text{concordanze è } P(A_{i_1} A_{i_2} \dots A_{i_h}) = \frac{(n-h)!}{n!} & \text{(Per } h \text{ elementi ben precisi)} \end{matrix}$$

Le probabilità sono indipendenti dagli indici, cioè la probabilità è la stessa a prescindere da quali sono gli h elementi $\Rightarrow$ poiché h elementi su un totale di n si possono scegliere in $\binom{n}{h}$ modi $\Rightarrow$

$$S_h = \text{probabilità di avere } h \text{ concordanze a prescindere da quali sono gli } h \text{ elementi che considero} = \binom{n}{h} \cdot P(A_{i_1} A_{i_2} \dots A_{i_h}) = \binom{n}{h} \cdot \frac{(n-h)!}{n!} =$$

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

Adesso, se poniamo per convenzione $S_0 = 1$, per quello visto nella dimostrazione

del principio di inclusione ed esclusione si ha:

$$P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = \sum_{h=0}^n (-1)^h S_h = \sum_{h=0}^n (-1)^h \frac{1}{h!} \approx e^{-1} \quad \text{Sviluppo in serie di McLaurin nel punto } x=-1$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{h=1}^n (-1)^{h+1} S_h = \sum_{h=1}^n (-1)^{h+1} \frac{1}{h!} \approx 1 - e^{-1}$$

Quindi la probabilità di avere nessuna concordanza è: $P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) \approx \frac{1}{e} \approx 0,3679$

• $VAR(X+Y) = P[(X+Y) - (m_x+m_y)]^2 = VAR(X) + VAR(Y) + 2P[(X-m_x)(Y-m_y)]$ 16-11-02 (15) VERERBI

• $COV(X,Y) = P[(X-m_x)(Y-m_y)]$

• $VAR(aX+b) = a^2 VAR(X)$

• $P(|A|) = p$, $p = P(A)$

• $VAR(|A|) = P[|A|^2] - P^2(|A|) = P[|A|^2] - p^2$, vediamo quanto vale $P[|A|^2]$;
 $p = P(A)$, $|A|^2 \in \{0, 1\}$

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

$\Rightarrow P(|A|^2) = 0 \cdot q + 1 \cdot p = p \Rightarrow$

$VAR(|A|) = p - p^2 = p(1-p) = p \cdot q$

• $COV(X,Y) = P[(X-m_x)(Y-m_y)] =$, $m_x = P(X)$, $m_y = P(Y)$

$= P[X \cdot Y - m_y \cdot X - m_x \cdot Y + m_x \cdot m_y] =$, per la linearità di P , m_x e m_y sono costanti,

$= P(X \cdot Y) - m_y \cdot P(X) - m_x \cdot P(Y) + m_x \cdot m_y =$

$= P(X \cdot Y) - m_y \cdot m_x - m_x \cdot m_y + m_x \cdot m_y = P(X \cdot Y) - m_y \cdot m_x =$

$= P(X \cdot Y) - P(Y) \cdot P(X) = COV(X,Y)$

• Esempio (di cosa indica la covarianza).

Consideriamo due urne A e B contenenti entrambe una pallina nera e una pallina bianca. Estraggo una pallina da A e la metto in B, e considero i numeri aleatori:

$X =$ numero di palline bianche in A, $\Rightarrow X \in \{0, 1\}$

$Y =$ numero di palline bianche in B, $\Rightarrow Y \in \{1, 2\}$

Calcoliamo la $COV(X,Y) = P(X \cdot Y) - P(X) \cdot P(Y)$

$p(X=0) = \frac{1}{2} \Leftrightarrow p(Y=2) = \frac{1}{2}$

$p(X=1) = \frac{1}{2} \Leftrightarrow p(Y=1) = \frac{1}{2}$, allora

$P(X) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$; $P(Y) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}$

Chi è $X \cdot Y$? $X \cdot Y$ è il numero aleatorio dei possibili prodotti $X \cdot Y$. Ad esempio non può essere $X \cdot Y = 0 \cdot 1$ perché se $X=0 \Rightarrow$ necessariamente $Y=2 \Rightarrow$

$X \cdot Y \in \{0 \cdot 2 = 0, 1 \cdot 1 = 1\} = \{0, 1\}$

$p(X \cdot Y = 0) = p(X=0, Y=2) = p(X=0) = \frac{1}{2}$

~~$p(X \cdot Y = 1) = p(X=1, Y=1) = p(X=1) = \frac{1}{2}$~~

$\Rightarrow P(X \cdot Y) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$

$\Rightarrow COV(X,Y) = P(X \cdot Y) - P(X) \cdot P(Y) = \frac{1}{2} - \frac{1}{2} \cdot \frac{3}{2} = \frac{1}{2} - \frac{3}{4} = -\frac{1}{4} < 0$, come ci si aspettava

perché a valori piccoli di X (cioè $X=0 < P(X)=\frac{1}{2}$) si associano valori grandi di Y (cioè $Y=2 > P(Y)=\frac{3}{2}$, perché a $X=0$ corrisponde $Y=2$).

• $COV(X,X) = VAR(X)$, infatti:

$COV(X,X) = P(X \cdot X) - P(X) \cdot P(X) = P(X^2) - P^2(X) = VAR(X)$

• Siano dati i numeri casuali $X = \sum_{i=1}^n a_i x_i$, $Y = \sum_{j=1}^m b_j y_j \Rightarrow$

$$\text{COV}(X, Y) = \text{COV}\left(\sum_{i=1}^n a_i x_i, \sum_{j=1}^m b_j y_j\right) =$$

$$= \mathbb{P}\left[\left(\sum_{i=1}^n a_i x_i - \mathbb{P}\left(\sum_{i=1}^n a_i x_i\right)\right) \cdot \left(\sum_{j=1}^m b_j y_j - \mathbb{P}\left(\sum_{j=1}^m b_j y_j\right)\right)\right] =$$

Op. gli $x_i \in \{x_1, \dots, x_n\}$
si possono vedere come numeri
casuali associati all'esperimento
 $\Rightarrow X$ è combinazione lineare
di numeri casuali $\Rightarrow$ per la linearità
di $\mathbb{P}$ segue:

per OSS

$$\uparrow = \mathbb{P}\left[\left(\sum_{i=1}^n a_i x_i - \sum_{i=1}^n a_i \mathbb{P}(x_i)\right) \cdot \left(\sum_{j=1}^m b_j y_j - \sum_{j=1}^m b_j \mathbb{P}(y_j)\right)\right] =$$

$$= \mathbb{P}\left[\sum_{i=1}^n \sum_{j=1}^m a_i b_j x_i y_j - \sum_{i=1}^n \sum_{j=1}^m a_i b_j x_i \mathbb{P}(y_j) - \sum_{i=1}^n \sum_{j=1}^m a_i b_j y_j \mathbb{P}(x_i) + \sum_i \sum_j a_i b_j \mathbb{P}(x_i) \mathbb{P}(y_j)\right] =$$

$$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathbb{P}(x_i y_j) - \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathbb{P}(x_i) \mathbb{P}(y_j) - \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathbb{P}(y_j) \mathbb{P}(x_i) + \sum_i \sum_j a_i b_j \mathbb{P}(x_i) \mathbb{P}(y_j)$$

$$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathbb{P}(x_i y_j) - \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathbb{P}(x_i) \mathbb{P}(y_j) =$$

$$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j (\mathbb{P}(x_i y_j) - \mathbb{P}(x_i) \mathbb{P}(y_j)) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{COV}(x_i, y_j) \quad (1)$$

• Esempio

$$\text{COV}(3X + 2Y, X + 2T) = 3 \text{COV}(X, X) + 6 \text{COV}(X, T) + 2 \text{COV}(Y, X) + 4 \text{COV}(Y, T)$$

$$\bullet \text{COV}(X, X) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{COV}(x_i, x_j) =$$

$$= \sum_{i=1}^n a_i^2 \text{VAR}(x_i) + 2 \sum_{i < j} a_i a_j \text{COV}(x_i, x_j)$$

$\downarrow$
 $i=j$

$\downarrow$
se $i \neq j$, $a_i a_j$ formano una matrice simmetrica $\Rightarrow$
con $i < j$ è conto soltanto una volta, da qui nasce il 2.

• Si definisce indice di correlazione o covarianza normalizzata il valore

$$\rho_{X,Y} = \frac{\text{COV}(X, Y)}{\sigma_X \cdot \sigma_Y}, \text{ dove } \sigma_X = \sqrt{\text{VAR}(X)}, \sigma_Y = \sqrt{\text{VAR}(Y)}$$

Osserviamo che $\rho_{X,Y} = \text{COV}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) = \frac{1}{\sigma_X} \frac{1}{\sigma_Y} \cdot \text{COV}(X, Y)$, per quello visto in (1)

Si dimostra che valgono le seguenti proprietà:

1) $-1 \leq \rho_{X,Y} \leq 1$

2) $\rho_{X,Y} = \pm 1 \Leftrightarrow Y = aX + b$ (il senso $\Rightarrow$ è vero con probabilità 1, cioè non è detto che sia vero ma è molto probabile, e' è l'implicazione $\Rightarrow$ può essere che in qualche caso non si verifichi)

Allo stesso modo dire che un evento ha probabilità 0 non vuol dire che sia impossibile ma che è molto improbabile.

Proviamo la 1)

$$\text{VAR}\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = \frac{1}{\sigma_X^2} \text{VAR}(X) + \frac{1}{\sigma_Y^2} \text{VAR}(Y) + \frac{2}{\sigma_X \sigma_Y} \text{COV}(X, Y) \geq 0, \text{ perché } \text{VAR}(X) \geq 0$$

$$\Rightarrow 1 + 1 + 2 \rho_{X,Y} \geq 0 \Leftrightarrow \rho_{X,Y} \geq -1$$

$$\text{VAR}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = \frac{1}{\sigma_X^2} \text{VAR}(X) + \frac{1}{\sigma_Y^2} \text{VAR}(Y) - \frac{2}{\sigma_X \sigma_Y} \text{COV}(X, Y) \geq 0$$

$$\Rightarrow 2 - 2 \rho_{X,Y} \geq 0 \Leftrightarrow \rho_{X,Y} \leq 1 \text{ c.v.d.}$$

Proviamo la 2) $\rho_{X,Y} = \pm 1 \Leftrightarrow Y = aX + b$

Hp $Y = aX + b$, Toci $\rho_{X,Y} = \pm 1$

$$\rho_{x,y} = \frac{\text{COV}(x,y)}{\sigma_x \sigma_y} = \frac{\text{COV}(x, ax+b)}{\sigma_x \sigma_y} = \frac{a \text{COV}(x,x) + \text{COV}(x,b)}{\sigma_x \sigma_y} \quad (12)$$

$$\sigma_y = \sqrt{\text{VAR}(Y)}, \quad \text{VAR}(Y) = \text{VAR}(aX+b) = a^2 \text{VAR}(X) \Rightarrow \sigma_y = \sqrt{a^2 \text{VAR}(X)} = |a| \cdot \sigma_x$$

$$\Rightarrow \rho_{x,y} = \frac{a \text{COV}(x,x) + \text{COV}(x,b)}{|a| \sigma_x^2} = \frac{a \text{VAR}(x)}{|a| \sigma_x^2} = \frac{a}{|a|} = \begin{cases} 1 & \text{se } a > 0 \\ -1 & \text{se } a < 0 \end{cases}$$

On $\text{COV}(x,c) = 0$, $c = \text{costante}$ perché:

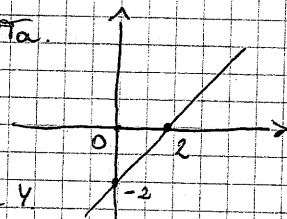
$$\text{COV}(x,c) = P(c \cdot x) - P(c) \cdot P(x) = c \cdot P(x) - c \cdot P(x) = 0$$

La covarianza tra un numero aleatorio e una costante è nulla.

Il coefficiente di correlazione non lo dimostriamo, ma sappiamo che si verifica con probabilità 1, cioè ~~non~~ è molto probabile che accenga ma non è detto che accenga sempre.

Il coefficiente di correlazione indica la dispersione di x e y attorno ad una retta ^{c.v.d.} ~~non~~ indica il coefficiente angolare. Se $\rho_{x,y} \approx \pm 1$ i punti stanno molto vicini o stanno sulla retta. Se $\rho_{x,y} \approx 0$ i punti sono più distanti dalla retta.

Il segno negativo di $\rho_{x,y}$ indica che la retta è inclinata in senso opposto.



Se $\rho_{x,y} \approx 1$ è lecito supporre che ci sia un legame lineare tra x e y .

Nell'esempio delle due urne entrambe con una pallina bianca e con una pallina nera si ha: $\text{COV}(X,Y) = -1$, infatti c'è un legame lineare:

$$x + y = 2, \text{ quindi } y = -x + 2$$

per questo -1 e non 1, non c'è bisogno di calcolarlo.

FUNZIONE DI RIPARTIZIONE.

Sia X un numero aleatorio, $X \in \{x_1, x_2, \dots, x_n\}$ con $x_1 < x_2 < \dots < x_n$. Sia $p_i = P(X=x_i)$, $i=1,2,\dots,n$.

$\Rightarrow$ si dice funzione di ripartizione $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$.

Osserviamo che $F: \mathbb{R} \rightarrow [0,1]$, infatti:

$$F(x) = \begin{cases} 0 & \text{se } x < x_1 \\ p_1 & \text{se } x_1 \leq x < x_2 \\ p_1 + p_2 & \text{se } x_2 \leq x < x_3 \\ \vdots & \vdots \\ p_1 + p_2 + \dots + p_{n-1} & \text{se } x_{n-1} \leq x < x_n \\ 1 = p_1 + p_2 + \dots + p_n & \text{se } x \geq x_n \end{cases} \quad \begin{aligned} & \rightarrow (X \leq x \Leftrightarrow (X=x_1 \cup X=x_2) \Rightarrow \\ & P(X \leq x) = P(X=x_1 \cup X=x_2) = p_1 + p_2 \end{aligned}$$

Un numero aleatorio X si dice discreto se il suo codominio C_X è finito o numerabile.

Una distribuzione di probabilità di un numero aleatorio discreto è tale che:

$$1) P(X=x) > 0 \quad \forall x \in C_X$$

$$2) P(X=x) = 0 \quad \forall x \notin C_X$$

$$x_n \in C_X \Rightarrow \sum_{n=1}^{\infty} P(X=x_n) = 1 \quad \text{dal punto di vista assiomatico}$$

$$\sum_{n=1}^{\infty} P(X=x_n) \leq 1 \quad \text{dal punto di vista soggettivo}$$

Per un numero aleatorio discreto X si definisce la previsione $P(x) = \sum_{i=1}^{\infty} x_i p_i$ in questo modo se si verificano queste due proprietà:

$$1) \sum_{i=1}^{\infty} p_i = 1$$

$$2) \sum_{i=1}^{\infty} |x_i| p_i < \infty, \text{ cioè la serie deve essere assolutamente convergente.}$$

Quest'ultima proprietà fa sì che per la serie (o previsione in questo caso) valga la proprietà associativa e commutativa e ciò permette insieme alla 1) di evitare che, modificando l'ordine o associando i termini della serie, si possa ottenere per $P(x)$ qualunque valore. (Teorema di Riemann-Dini).

26-11-07
LUNEDÌ

Eventi condizionati

Un evento condizionato $A|B$, da leggersi "A dato B", oppure "A condizionato ad B" è un ente a tre valori:

1) VERO, quando essendo vero B anche A risulta vero.

2) FALSO, quando essendo vero B anche A risulta falso.

3) Indeterminato quando B risulta falso e quindi non interessa sapere se A è vero o falso.

Teorema delle probabilità composte

Si può interpretare un evento condizionato come un numero aleatorio X a tre valori: ai tre casi VERO, FALSO, indeterminato corrispondano rispettivamente i valori 1, 0 e

$p = P(E|H)$, cioè il numero aleatorio X si può rappresentare nella forma:

$$X = (E|H) + p \cdot (1 - |H|)$$

Ricordiamo che la previsione $P(x)$ di un numero aleatorio semplice X rappresenta la somma da versare per ricevere un importo pari al valore che assumerà $X \Rightarrow$

p deve coincidere con la previsione $P(x)$ del numero aleatorio $\Rightarrow$

$$p = P(x) = P((E|H) + p(1 - |H|)) = P(E|H) + p \cdot P(1 - |H|) =$$

$$= P(E|H) + p + p \cdot p(H) \Leftrightarrow p = P(E|H) + p \cdot p(H)$$

$$\Leftrightarrow \underline{p(E|H) = P(E|H) \cdot p(H)} \quad \text{la dim. non è obbligatoria.}$$

o o o, $P(A|B) = P(A|B) \cdot P(B)$

$$\Rightarrow P(A|B) P(B) = P(B|A) \cdot P(A)$$

$$P(B|A) = P(B|A) \cdot P(A)$$

* $A \text{ e } A^c \Rightarrow A \cup A^c = \Omega \Rightarrow P(A) + P(A^c) = 1$

Se $B \neq \emptyset \Rightarrow (A|B)^c = (A^c|B) \Rightarrow P(A|B) + P(A^c|B) = 1$

$$\Rightarrow P(A|B) + P(A^c|B) = 1$$

* Dati n eventi $A_1, A_2, \dots, A_n : A_i \cap A_j = \emptyset, i \neq j$ e $B \neq \emptyset \Rightarrow$

$$\Rightarrow P(A_1 \cup A_2 \cup \dots \cup A_n | B) = P(A_1|B) + P(A_2|B) + \dots + P(A_n|B)$$

lo dimostriamo nel caso $n=2$

$$P(A_1 \cup A_2 | B) = \frac{P[(A_1 \cup A_2) \cap B]}{P(B)} = \frac{P[(A_1 \cap B) \cup (A_2 \cap B)]}{P(B)} =$$

Osserviamo che $A_1 \cap B$ e $A_2 \cap B$ sono incompatibili perché A_1 e A_2 per ipotesi $\Rightarrow$

$$= \frac{P(A \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} = P(A_1|B) + P(A_2|B) \quad \text{e.v.d.}$$

(19)

NOTA. Nel caso in cui $A \cap B \neq \emptyset \Rightarrow P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C) =$
 $= P(A|C) + P(B|C) - P(A \cap B|C)$, infatti $A \cap B|C = A \cap B$.

Se $P(A|B) > P(A) \Rightarrow A$ si dice correlato positivamente a B .

Se $P(A|B) = P(A) \Rightarrow A$ si dice stocasticamente indipendente da B .

Se $P(A|B) < P(A) \Rightarrow A$ si dice stocasticamente correlato negativamente a B .

Infatti se $P(A|B) > P(A)$ e calcoliamo la covarianza tra $|A|$ e $|B|$ si ha:

$$\begin{aligned} \text{COV}(|A|, |B|) &= P(|A| \cdot |B|) - P(|A|) \cdot P(|B|) = P(A \cap B) - P(A) \cdot P(B) = \\ &= P(A \cap B) - P(A) \cdot P(B) = P(A|B) \cdot P(B) - P(A) \cdot P(B) = P(B) [P(A|B) - P(A)] > 0, \end{aligned}$$

da qui il nome correlato positivamente.

• Due eventi A e B si dicono logicamente indipendenti, quando conosciuto il valore di verità di uno (l'indicatore) di uno l'altro non è univocamente determinato, mentre si dicono stocasticamente indipendenti se l'esito di A è indipendente da B (il relativo alla probab. $P(A)$).

• Consideriamo gli eventi A e B . In pratica, $1(A \Rightarrow B)$, $1(A \Rightarrow \neg B)$, $1(\neg A \Rightarrow B)$, $1(\neg A \Rightarrow \neg B)$.

$A =$ "Un individuo ha una certa malattia"

$B =$ "l'individuo ha un'età compresa tra a e b "

quindi $A \cap B$, $A \cap B^c$, $A^c \cap B$, $A^c \cap B^c$

Questi due eventi sono logicamente indipendenti perché può accadere $A \cap B$, $A^c \cap B$, $A \cap B^c$, $A^c \cap B^c$ e possono essere stocasticamente indipendenti se una particolare malattia colpisce un individuo indipendentemente dall'età che ha oppure non essere stocasticamente indipendenti se la malattia colpisce solo in una ben determinata fascia di età, e cioè i due eventi possono essere sia logicamente indip. che stocasticamente indip. oppure essere logicamente ind. e non stocasticamente indipendenti.

• Se due eventi sono stocasticamente indipendenti (o indipendenti) $\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

infatti $P(A|B) = P(A)$ perché sono "indep."

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

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

$$\Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$$

• Osserviamo che se due eventi sono indipendenti $\Rightarrow$

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

$$P(A^c|B^c) = P(A^c)$$

$$P(A^c|B) = P(A^c)$$

$$P(B|A^c) = P(B)$$

$$P(B^c|A^c) = P(B^c)$$

$$P(A|B^c) = P(A)$$

• Se $A_1, \dots, A_n$ sono stocasticamente indipendenti $\Rightarrow \forall s \geq 1, s \leq n, \forall A_1, \dots, A_s$ si ha:

$$P(A_1 \cap A_2 \cap \dots \cap A_s) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_s)$$

ed inoltre $\forall h \geq 1$ e $\forall s \geq 1$ e $\forall A_1, \dots, A_s$ si ha:

$$P(A_1^c \cap \dots \cap A_h^c \cap A_{h+1} \cap \dots \cap A_s) = P(A_1^c) \cdot \dots \cdot P(A_h^c) \cdot P(A_{h+1}) \cdot \dots \cdot P(A_s)$$

(basta osservare che $P(E_1^c \cap E_2 \cap \dots \cap E_s) = P(E_1^c | E_2 \cap E_3 \cap \dots \cap E_s) \cdot P(E_2 \cap E_3 \cap \dots \cap E_s)$ per il Teorema delle probab. $P(A)$ composte.

SCAMBIABILITÀ

Dati n eventi $A_1, \dots, A_n$ essi si dicono scambiabili se $\forall h \geq 0, \forall n \geq 1$ si ha:

$$P(A_1 \cap A_2 \cap \dots \cap A_h \cap A_{h+1}^c \cap \dots \cap A_n^c) = p(h, n),$$

diversificarsi di h eventi.

cio' significa che se gli eventi sono scambiabili la probabilità dell'intersezione di h di essi non dipende da quali eventi considero ma da quanti ne considero veri e falsi, cioè da h e da n .

Esempio. Se A, B e C sono tre eventi scambiabili $\Rightarrow$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2^c \cap A_3) = P(A_1 \cap A_2 \cap A_3^c) = P(A_1^c \cap A_2 \cap A_3)$$

Teorema

Indipendenza + equiprobabilità $\Rightarrow$ Scambiabilità

Dim. Dati n eventi $A_1, \dots, A_n$: $p = P(A_1) = \dots = P(A_n)$ ed $q = 1 - p = P(A_i^c) \Rightarrow$

$$P(A_1 \cap \dots \cap A_h \cap A_{h+1}^c \cap \dots \cap A_n^c) = P(A_1) \cdot \dots \cdot P(A_h) \cdot P(A_{h+1}^c) \cdot \dots \cdot P(A_n^c) = p^h \cdot q^{n-h}$$

scambiabilità $\Rightarrow$ equiprobabilità Dim. Scegli $n=h+1$ verrà la tesi perché dipende da h e n .

DISTRIBUZIONE BINOMIALE

Dati n eventi $A_1, \dots, A_n$ indipendenti ed equiprobabili con $p = P(A_i), i=1, \dots, n$ e $q = P(A_i^c) = 1 - p \quad \forall i=1, \dots, n$ si dice distribuzione binomiale dipendente dai parametri n, p .

quella del numero aleatorio $X = |A_1| + |A_2| + \dots + |A_n| =$ numero di eventi che si verificano

$$X \in \{0, 1, \dots, n\}$$

$$P(X=0) = P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = P(A_1^c) \cdot P(A_2^c) \cdot \dots \cdot P(A_n^c) = q^n$$

$$P(X=1) = P[(A_1 \cap A_2^c \cap \dots \cap A_n^c) \cup (A_1^c \cap A_2 \cap A_3^c \cap \dots \cap A_n^c) \cup \dots \cup (A_1^c \cap \dots \cap A_{n-1}^c \cap A_n)] =$$

$$= P(A_1 \cap A_2^c \cap \dots \cap A_n^c) + P(A_1^c \cap A_2 \cap A_3^c \cap \dots \cap A_n^c) + \dots + P(A_1^c \cap \dots \cap A_{n-1}^c \cap A_n) = n \cdot p \cdot q^{n-1}$$

$$P(X=h) = P(\cup_{i_1, \dots, i_h} A_{i_1} \cap \dots \cap A_{i_h} \cap A_{i_{h+1}}^c \cap \dots \cap A_n^c) = \binom{n}{h} p^h \cdot q^{n-h}$$

Si dice che X ha distribuzione binomiale di parametri n e p e si indica $X \sim B_{n,p}$

Esempio

* Estrazione con restituzione da un'urna di composizione nota

N = palline, c = palline bianche, $N-c$ = nere

n = numero di estrazioni con restituzione.

A_i = "l' i -esima pallina estratta è bianca"

Consideriamo il numero aleatorio $X = |A_1| + |A_2| + \dots + |A_n| =$ numero di palline bianche uscite in n estrazioni.

$$P(A_i) = \frac{c}{N} \quad \forall i=1, \dots, n$$

$$P(X=h) = \binom{n}{h} \left(\frac{c}{N}\right)^h \cdot \left(\frac{N-c}{N}\right)^{n-h} \Rightarrow X \sim B_{n, \frac{c}{N}}$$

• Supponiamo che $X \sim B_{n,p}$, $X = \sum_{i=1}^n |A_i|$

$$\Rightarrow P(X) = ?, \text{VAR}(X) = ?$$

Calcoliamo

(21)

$$P(X) = \sum_{i=1}^m 1 \cdot p_i = \sum_{j=1}^m p = p \cdot m = \frac{m}{N} \cdot m$$

$$VAR(X) = VAR(A_1 + A_2 + \dots + A_m) = \sum_{i=1}^m VAR(A_i) + 2 \sum_{i < j} COV(A_i, A_j) =$$

Toss da se A e B sono due eventi stocasticamente indipendenti $\Rightarrow COV(A, B) = 0$

infatti: $COV(A, B) = P(A \cdot B) - P(A) \cdot P(B) = P(A \cap B) - P(A) \cdot P(B) =$

$$= P(AB) - P(A) \cdot P(B) = P(AB) - P(AB) = 0 \quad \checkmark$$

eventi
indipendenti

$$= m \cdot p \cdot q + 0 = m \cdot p \cdot q$$

30.11.08
VENERDI

• Esempio

Si ha una pallina urna con una pallina bianca e una nera. Si effettuano 3 estrazioni con restituzione, mettendo nell'urna un'altra pallina dello stesso colore di quella estratta. Consideriamo l'evento $E_i = \text{"i-esima pallina estratta è bianca"}$

Calcoliamo le probab. P.Ta $P(E_2)$ e $P(E_3)$

$$E_2 = E_2 \cap \Omega = E_2 \cap (E_1 \cup E_1^c) = E_2 E_1 \cup E_2 E_1^c = \text{unione di eventi incompatibili} \Rightarrow$$

$$P(E_2) = P(E_2 E_1) + P(E_2 E_1^c) = \text{per i P.Ta. delle probab. P.Ta. composte}$$

$$= P(E_2 | E_1) P(E_1) + P(E_2 | E_1^c) P(E_1^c) = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$$

Teorema delle probab. P.Ta. composte generalizzato

$$P(E_1 \cap E_2 \cap \dots \cap E_m) = P(E_m | E_1, E_2, \dots, E_{m-1}) \cdot P(E_{m-1} | E_1, E_2, \dots, E_{m-2}) \cdot \dots \cdot P(E_2 | E_1) \cdot P(E_1)$$

$$E_3 = E_3 \cap \Omega = E_3 \cap (E_1 E_2 \cup E_1 E_2^c \cup E_1^c E_2 \cup E_1^c E_2^c) =$$

Intersezione dei costituenti di E_1 e $E_2 = \Omega$

$$= E_1 E_2 E_3 \cup E_1 E_2^c E_3 \cup E_1^c E_2 E_3 \cup E_1^c E_2^c E_3 = \text{unione di eventi incompatibili}$$

$$\Rightarrow P(E_3) = P(E_1 E_2 E_3) + P(E_1 E_2^c E_3) + P(E_1^c E_2 E_3) + P(E_1^c E_2^c E_3) = \text{per Teo. della prob. composte gener.}$$

$$= P(E_3 | E_1 E_2) \cdot P(E_2 | E_1) \cdot P(E_1) + P(E_3 | E_2^c E_1) \cdot P(E_2^c | E_1) \cdot P(E_1) +$$

$$+ P(E_3 | E_2 E_1^c) \cdot P(E_2 | E_1^c) \cdot P(E_1^c) + P(E_3 | E_2^c E_1^c) \cdot P(E_2^c | E_1^c) \cdot P(E_1^c) =$$

$$= \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4} + \frac{1}{12} + \frac{1}{12} + \frac{1}{12} = \frac{6}{12} = \frac{1}{2}$$

$$\Rightarrow P(E_3) = \frac{1}{2}$$

Proviamo che il numero aleatorio $X = |E_1| + |E_2|$ ha distribuzione uniforme cioè X

ha assumere con le stesse probab. P.Ta. i valori del suo codominio $C_X = \{0, 1, 2\}$

$$P(X=0) = P(E_1^c E_2^c) = P(E_2^c | E_1^c) \cdot P(E_1^c) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$P(X=1) = P(E_1 E_2^c \cup E_1^c E_2) = \text{poiché sono incompatibili} = P(E_1 E_2^c) + P(E_1^c E_2) =$$

$$= P(E_2^c | E_1) \cdot P(E_1) + P(E_2 | E_1^c) \cdot P(E_1^c) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3}$$

$$p(X=2) = p(E_1 E_2) = p(E_2|E_1) \cdot p(E_1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$$

$\Rightarrow X$ ha distribuzione uniforme sull'insieme $\{0, 1, 2\} \ni x$.

• Esempio

Sia n eventi $E_1, \dots, E_n$ indipendenti ed equiprobabili, con $p = p(E_i), i=1, \dots, n$ ed
 $S_n = \sum_{i=1}^n |E_i|$ = numero di eventi che si verificano. Si chiede di determinare la distribuzione del numero aleatorio $X = |E_1| \mid (S_n = z)$, $z \in \{0, 1, \dots, n\}$ e l'immagine di X .

$$p(X=0) = p(|E_1|=0 \mid S_n=z) = p(E_1^c \mid S_n=z) = \frac{p(E_1^c \cap (S_n=z))}{p(S_n=z)}$$

$$= \frac{\binom{n-1}{z} p^z q^{n-z}}{\binom{n}{z} p^z q^{n-z}} = \frac{(n-1)!}{z! (n-z-1)!} \cdot \frac{(n-z)! z!}{n!} = \frac{n-z}{n} = 1 - \frac{z}{n}$$

$$p(E_i \mid S_n=z) = \frac{p(E_i \cap S_n=z)}{p(S_n=z)} = \frac{\text{se ne devono verificare } z \text{ su } n \text{ ma } E_i \text{ si è verificato sicuramente} \Rightarrow}{\binom{n}{z} p^z q^{n-z}} = \frac{\binom{n-1}{z-1} p^{z-1} q^{n-z+1}}{\binom{n}{z} p^z q^{n-z}}$$

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

• Esempio

Sia $m < n$ e sia $S_m = \sum_{i=1}^m |E_i|$

Calcolare la distribuzione di S_m condizionato a $S_n=z$, cioè di $S_m \mid (S_n=z)$

$S_m \in \{0, 1, \dots, m\}$

$p(S_m=h \mid S_n=z), h=0, 1, \dots, m$

$$= \frac{p(S_m=h, S_n=z)}{p(S_n=z)}$$

$$\text{Sia } S'_m = \sum_{i=1}^m |E'_i|$$

Osserviamo che $(S_m=h, S_n=z) = (S_m=h, S'_m=z-h)$, con la differenza che gli eventi del secondo membro sono incompatibili indipendenti e $S_m \sim B(m, p)$ ed

$$S'_m \sim B(n-m, p) \Rightarrow$$

$$= \frac{\binom{m}{h} p^h q^{m-h} \cdot \binom{n-m}{z-h} p^{z-h} q^{n-m-z+h}}{\binom{n}{z} p^z q^{n-z}} = \frac{\binom{m}{h} \binom{n-m}{z-h} p^z q^{n-z}}{\binom{n}{z} p^z q^{n-z}} = \frac{\binom{m}{h} \binom{n-m}{z-h}}{\binom{n}{z}}$$

Un gioco si dice equo quando la previsione del guadagno è nulla

• Esempio

$$\text{equo} \Leftrightarrow EG = 0$$

Un giocatore paga un euro per giocare al seguente gioco: si lanciano tre dadi e vince

1 euro se la faccia 1 appare una volta;

2 euro se la faccia 1 appare due volte;

8 euro se la faccia 1 appare 3 volte e 0 euro altrimenti.

Stabilire se il gioco è giusto.

X = "numero di volte in cui appare la faccia 1", $x \in \{0, 1, 2, 3\}$.

$$\text{guadagno } G = (1-1) \cdot |X=1| + (2-1) \cdot |X=2| + (8-1) \cdot |X=3| + (0-1) \cdot |X=0|$$

$$p(G) = p(|X=2| + X \cdot |X=3| - 1 \cdot |X=0|) = p(X=2) + 7 \cdot p(X=3) - p(X=0) =$$

$$= \binom{3}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right) + 7 \cdot \left(\frac{1}{6}\right)^3 - \left(\frac{5}{6}\right)^3 \neq 0$$

⇒ il gioco non è equo. Quando dovrebbe ricevere il giocatore in caso di $x=3$, affinché il gioco sia equo?

$$\downarrow \quad \binom{3}{x} \frac{5}{6^3} + (x-1) \frac{1}{6^3} - \frac{5^3}{6^3} = 0 \Leftrightarrow 15 + x - 1 - 5^3 = 0 \Leftrightarrow x = 111 \text{ euro.}$$

Il gioco risulta equo se il giocatore riceve 111 euro quando escono Tre 1.

Exp

3-12-07

LUNEDÌ

Sia data un'urna con N palline di cui c bianche e $N-c$ nere.

Si effettuano n estrazioni senza restituzione. Consideriamo l'esempio:

$E_i =$ "l' i -esima pallina estratta è bianca"

Sia $X = |E_1| + |E_2| + \dots + |E_n|$ il numero di palline bianche estratte.

• Dimostriamo per induzione $p(E_i) = \frac{c}{N} \quad \forall i=1, \dots, n$

Base dell'induzione $p(E_1) = \frac{c}{N}$ vero

Supponiamo che sia vero $p(E_{i-1}) = \frac{c}{N}$. Dimostriamo che $p(E_i) = \frac{c}{N}$

Introduciamo il numero di palline bianche estratte nelle $i-1$ estrazioni precedenti:

$$p(E_i) = p(E_i \cap \Omega) = p(E_i \cap \bigcup_{k=0}^{i-1} (X_{i-1} = k)) = p\left(\bigcup_{k=0}^{i-1} E_i \cap (X_{i-1} = k)\right) = \sum_k p(E_i \cap (X_{i-1} = k)) =$$

$$= \sum_k p(E_i | X_{i-1} = k) \cdot p(X_{i-1} = k) = \sum_k \frac{c-k}{N-(i-1)} \cdot p(X_{i-1} = k) =$$

$$= \frac{c}{N-i+1} \underbrace{\sum_k p(X_{i-1} = k)}_{\substack{\text{1 per definizione} \\ \text{Tutti i possibili casi}}} = \frac{1}{N-i+1} \underbrace{\sum_k k \cdot p(X_{i-1} = k)}_{\substack{\text{1 per definizione} \\ \text{Tutti i possibili casi}}} =$$

$$= \frac{c - p(X_{i-1})}{N-i+1} =$$

$$X_{i-1} = |E_1| + \dots + |E_{i-1}|, \quad p = \frac{c}{N} = p(E_h), \quad h = 1, \dots, i-1$$

$$p(X_{i-1}) = p(|E_1| + |E_2| + \dots + |E_{i-1}|) = (i-1)p = (i-1) \frac{c}{N}$$

$$= \frac{c - (i-1) \frac{c}{N}}{N-i+1} = \frac{Nc - (i-1)c}{N(N-i+1)} = \frac{c}{N} \quad \text{c.v.d.} \Rightarrow \text{sono equiprobabili.}$$

$$p(E_2 | E_1) = \frac{c-1}{N-1} \neq \frac{c}{N} = p(E_1) = p(E_2)$$

$$p(E_2 | E_1) = \frac{p(E_2, E_1)}{p(E_1)} = \frac{\frac{c-1}{N-1} \cdot \frac{c}{N}}{\frac{c}{N}} \neq \frac{c}{N} \Rightarrow \text{non sono indipendenti.}$$

Si dimostra che sono scambiabili

Sia $X = |E_1| + |E_2| + \dots + |E_n|$

$$p(X=h) = p\left(\bigcup_{i_1, \dots, i_h} E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_h}^c\right) = \binom{n}{h} p(E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_h}^c) =$$

$$p(E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_h}^c) = p(E_{i_h}^c | E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_{h-1}}^c) \dots p(E_{i_2}^c | E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_{h-1}}^c) \cdot p(E_{i_1}^c | E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_{h-1}}^c) \cdot p(E_{i_1} \dots E_{i_h} E_{i_1}^c \dots E_{i_h}^c) =$$

$$\begin{aligned}
 &= \frac{(N-z) - (n-1-h)}{N-(n-1)} \cdot \frac{(N-z) - (n-1-h-1)}{N-(n-2)} \cdots \frac{N-z}{N-h} \cdot \frac{z-(h-1)}{N-(h-1)} \cdots \frac{z-1}{N-1} \cdot \frac{z}{N} \\
 &= \text{(da qui si capisce che gli elementi sono scambiabili perché questo prodotto non dipende dagli elementi della h-esima estrazione, ma dal fatto che sono h. Se consideriamo veri altri h esenti il prodotto non cambia)} \\
 &= \frac{(N-m)!}{N!} \cdot \frac{z!}{(z-h)!} \cdot \frac{(N-z)!}{[(N-z)-(n-h)]!} = \frac{(N-m)!}{N!} \cdot \frac{z!}{(z-h)!} \cdot \frac{(N-z)!}{[(N-z)-(n-h)]!} \cdot \frac{h!}{h!} \cdot \frac{m!}{m!} \cdot \frac{(n-h)!}{(n-h)!} \\
 &= \frac{\binom{z}{h} \binom{N-z}{n-h}}{\binom{N}{m} \binom{m}{h}} \\
 \Rightarrow P(X=h) &= \binom{m}{h} \frac{\binom{z}{h} \binom{N-z}{n-h}}{\binom{N}{m} \binom{m}{h}} = \frac{\binom{z}{h} \binom{N-z}{n-h}}{\binom{N}{m}} \quad (1)
 \end{aligned}$$

Def $\Rightarrow$ Si dice che X ha distribuzione ipergeometrica di parametri N, m, z , $X \sim H(N, m, z)$.

DSS Cerchiamo di spiegare la (1) in termini di casi favorevoli su casi possibili.

Infatti, effettuare n estrazioni senza restituzione equivale ad estrarre n palline in blocco $\Rightarrow$

$\binom{N}{n}$ = numero di modi in cui si possono scegliere n palline su N = casi possibili.

$\binom{z}{h}$ = numero di modi in cui si possono scegliere h palline bianche su z

$\binom{N-z}{n-h}$ = numero di modi in cui si possono scegliere $n-h$ palline nere su $N-z$

$\Rightarrow \binom{z}{h} \binom{N-z}{n-h}$ = casi favorevoli.

• Calcoliamo $P(X)$ e $VAR(X)$ nel caso in cui $X = |E_1| + \dots + |E_m|$ ha distribuzione ipergeometrica di parametri N, m, z .

$$P(X) = P(|E_1| + \dots + |E_m|) = \sum_{i=1}^m P(|E_i|) = \sum_{i=1}^m p(E_i) = \sum_{i=1}^m \frac{z}{N} = m \cdot \frac{z}{N} = mp$$

$$q = (1-p)$$

$$VAR(X) = VAR(|E_1| + \dots + |E_m|) = \sum_{i=1}^m VAR(|E_i|) + 2 \sum_{i < j} COV(|E_i|, |E_j|) = mpq + 2 \sum_{i < j} COV(|E_i|, |E_j|)$$

$$COV(|E_i|, |E_j|) = P(|E_i|, |E_j|) - P(|E_i|) \cdot P(|E_j|) = P(E_i E_j) - P(E_i) \cdot P(E_j) = \text{poiché sono scambiabili}$$

$$= P(E_1 E_2) - P(E_1) \cdot P(E_2) = P(E_2 | E_1) \cdot P(E_1) - P(E_1) \cdot P(E_2) =$$

$$= \frac{z-1}{N-1} p - p^2$$

$$= mpq + 2 \cdot \binom{m}{2} p \left(\frac{z-1}{N-1} - p \right) = mpq + \frac{2 \cdot m(m-1)}{2} p \left(\frac{z-1}{N-1} - p \right) = mpq + m(m-1)p \left(\frac{N-z-1}{N(N-1)} \right)$$

$$= mpq + m(m-1)p \frac{(z-N)}{N(N-1)} = mpq + m(m-1)p \frac{N}{N(N-1)} =$$

$$= mpq - m(m-1)p \frac{q}{N-1} = mpq \left(1 - \frac{m-1}{N-1} \right) = mpq \frac{N-1-m+1}{N-1} = mpq \frac{(N-m)}{(N-1)}$$

Dopo il 3-12-07 e prima del 14-12-07

Estrazione con restituzione da un'urna con composizione incognita.

Sia data un'urna contenente N palline di cui alcune bianche ed altre nere. Si effettuano n estrazioni con restituzione.

Consideriamo gli eventi: E_i = "l' i -esima pallina estratta è bianca".

Ci chiediamo: gli eventi E_i sono equiprobabili? Sono indipendenti?

Consideriamo il numero aleatorio $X = |E_1| + \dots + |E_N|$

Consideriamo gli eventi $H_c = \text{"l'urna contiene } c \text{ palline bianche"} \quad c=0,1,\dots,N$

Osserviamo che gli H_c formano una partizione $\Rightarrow$

$E_i = E_i \cap \Omega = E_i \cap (H_0 \cup H_1 \cup \dots \cup H_N) = E_i \cap H_0 \cup E_i \cap H_1 \cup \dots \cup E_i \cap H_N$, sono incompatibili.

$$\Rightarrow p(E_i) = p(E_i \cap H_0 \cup E_i \cap H_1 \cup \dots \cup E_i \cap H_N) = \sum_{c=0}^N p(E_i \cap H_c) = \sum_{c=0}^N p(E_i | H_c) \cdot p(H_c) = \sum_{c=0}^N \frac{c}{N} \cdot p(H_c)$$

$$\Rightarrow p(E_i) = \sum_{c=0}^N \frac{c}{N} \cdot p(H_c) \quad (1)$$

$\Rightarrow$ gli E_i sono equiprobabili perché la (1) non dipende da i .

Mentre gli E_i non sono indipendenti perché ogni volta che estraggo una pallina so qualcosa in più sulla composizione dell'urna (ad esempio se estraggo una pallina bianca sicuramente H_0 è impossibile).

$p(\cdot | H_c)$ è una probabilità

Invece gli E_i verificano l'ipotesi di indipendenza subordinata, cioè gli E_i subordinatamente all'ipotesi $H_c = \text{"l'urna contiene } c \text{ palline bianche"}$, sono indipendenti $\Rightarrow$

$$p(E_i E_j | H_c) = p(E_i | H_c) \cdot p(E_j | H_c) = \frac{c}{N} \cdot \frac{c}{N} = \frac{c^2}{N^2}$$

Calcoliamo la probabilità: $p(X=h) =$ in n estrazioni si estraggono h palline bianche.

$$p(X=h) = p((X=h) \cap \Omega) = p((X=h) \cap (H_0 \cup \dots \cup H_N)) = p((X=h) \cap H_0 \cup (X=h) \cap H_1 \cup \dots \cup (X=h) \cap H_N) = \sum_{c=0}^N p(X=h, H_c) = \sum_{c=0}^N p(X=h | H_c) \cdot p(H_c) = \sum_{c=0}^N \binom{n}{h} \left(\frac{c}{N}\right)^h \left(\frac{N-c}{N}\right)^{n-h} p(H_c)$$

dato $H_c \Leftrightarrow$ conosco la composizione dell'urna

$\Rightarrow$ è una distribuzione binomiale: $p = p(H_c) = \frac{c}{N}$

quattro di binomiali

prob. che una pallina data H_c .

Gli E_i sono scambiabili?

$$p(E_{i_1} E_{i_2} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c) = p(E_{i_1} E_{i_2} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c \cap \Omega) =$$

$$= p[(E_{i_1} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c) \cap (H_0 \cup H_1 \cup \dots \cup H_N)] =$$

$$p[(E_{i_1} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c \cap H_0) \cup (E_{i_1} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c \cap H_1) \cup \dots \cup (E_{i_1} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c \cap H_N)] =$$

$$= \sum_{c=0}^N p(E_{i_1} \dots E_{i_k} E_{i_{k+1}}^c \dots E_{i_n}^c | H_c) \cdot p(H_c) =$$

$$= \sum_{c=0}^N p(E_{i_1} | H_c) \cdot \dots \cdot p(E_{i_k} | H_c) \cdot p(E_{i_{k+1}}^c | H_c) \cdot \dots \cdot p(E_{i_n}^c | H_c) \cdot p(H_c) =$$

$$= \sum_{c=0}^N \left(\frac{c}{N}\right)^k \left(\frac{N-c}{N}\right)^{n-k} p(H_c) \Rightarrow \text{sono scambiabili.}$$

Estrazioni senza restituzione da un'urna di composizione incognita

Sia data un'urna contenente N palline di cui alcune bianche e altre nere. Si effettuano n estrazioni ^{senza} restituzione.

Consideriamo gli eventi $E_i = \text{"l'esima pallina estratta è bianca"}$. Come prima ci chiediamo se sono equiprobabili, indipendenti, e scambiabili.

Consideriamo il numero aleatorio $X = |E_1| + \dots + |E_n|$ e l'evento

$H_c = \text{"l'urna contiene } c \text{ palline bianche"}$, $c=0,1,\dots,N$ che costituiscono una partizione.



$$p(E_i) = p(E_i \cap \Omega) = p(E_i \cap (H_1 \cup H_2 \cup \dots \cup H_N)) = \sum_{c=0}^N p(E_i \cap H_c) = \sum_{c=0}^N \underbrace{p(E_i | H_c)}_1 p(H_c) = \sum_{c=0}^N \frac{1}{N} p(H_c) = \frac{1}{N}$$

$\Rightarrow$ Gli E_i sono equiprobabili perché

$p(E_i)$ non dipende da i .

dimostrato il 3-12-08 ne inducono che $p(E_i | H_c) = \frac{1}{N}$

Siamo nel caso di un'unica composizione nota senza restituzione e abbiamo dimostrato il 3-12-08 ne inducono che $p(E_i | H_c) = \frac{1}{N}$

Ovviamente non sono indipendenti, né stocasticamente né subordinatamente.

~~per la prima volta~~

$$p(X=R) = \sum_{c=0}^N p(X=R | H_c) \cdot p(H_c) = \sum_{c=0}^N \frac{\binom{r}{h} \binom{N-r}{n-h}}{\binom{N}{n}} \cdot p(H_c)$$

(*)

per il ragionamento di prima

Siamo nel caso della distribuzione ipergeometrica.

"misto" di ipergeometriche = somma di ipergeometriche per la probabilità di essere composta è una.

~~Non sono scambiabili perché~~

Scambiabilità:

$$p(E_{i_1} \dots E_{i_r} \cdot E_{j_1}^c \dots E_{j_m}^c) = \sum_{c=0}^N p(E_{i_1} \dots E_{i_r} \cdot E_{j_1}^c \dots E_{j_m}^c | H_c) \cdot p(H_c) =$$

$$= \sum_{c=0}^N p(X=r | H_c) \cdot p(H_c) = \sum_{c=0}^N \frac{\binom{r}{h} \binom{N-r}{n-h}}{\binom{N}{n}} \cdot p(H_c)$$

?

? ? probabilità totale diviso il numero di casi possibili $\binom{n}{r}$ infatti in (*) si sottointende che ?

$X=r$ può verificarsi in $\binom{n}{r}$ modi. ?

Def

Un numero aleatorio X si dice assolutamente continuo se:

1) $p(X=x) = 0 \quad \forall x \in \mathbb{R}$ ($\mathbb{R}$ ha cardinalità del continuo $\Rightarrow$ gli $x \in \mathbb{R}$ sono troppi e non si possono attribuire probabilità positive, se no la somma sarebbe > 1)

2) $f(x) \geq 0$ integrabile secondo Riemann. Tale che $\forall C \subseteq \mathbb{R}$, con C misurabile secondo P.J, si ha: $p(X \in C) = \int_C f(x) dx$

$f(x)$ si dice densità di probabilità e di conseguenza si ha: $\int_{\mathbb{R}} f(x) dx = 1$

$f(x)$ per essere una densità di probabilità deve essere:

1) $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

2) $\int_{\mathbb{R}} f(x) dx = 1$

Ricordando che la funzione di ripartizione $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$

(vedi pg 17)

si ha che: $p(X \leq x) = p(X \in]-\infty, x]) = \int_{-\infty}^x f(x) dx$

$$\Rightarrow F(x) = \int_{-\infty}^x f(x) dx \Rightarrow \underline{f(x) = F'(x)} \quad \leftarrow \text{NOTA}$$

DISTRIBUZIONE UNIFORME

numero aleatorio assolutamente continuo.

Il simbolo $X \sim U[a,b]$ si legge X ha distribuzione uniforme in $[a,b]$.

In questo caso $f(x) = \begin{cases} K & \text{se } x \in [a,b] \\ 0 & \text{se } x \notin [a,b] \end{cases}$, con $K > 0 \Rightarrow$

$$\Rightarrow \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$\text{ma} \int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^a 0 dx + \int_a^b k dx + \int_b^{+\infty} 0 dx = 1 \Leftrightarrow \int_a^b k dx = 1$$

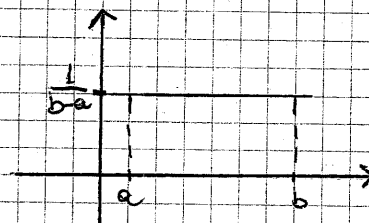
(2)

$$\Leftrightarrow k \cdot (b-a) = 1 \Rightarrow k = \frac{1}{b-a}$$

$$\Rightarrow \text{dato } [c, d] \subseteq [a, b], p(x \in [c, d]) = p(c \leq x \leq d) = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$$

Si dice distribuzione uniforme perché a intervalli di ampiezza uguali corrispondono probabilità uguali (ovviamente per avere la stessa probabilità gli intervalli devono essere contenuti in (a, b))

la probabilità che $x \in [c, d]$ è uguale all'area sotto della curva in $[c, d]$.



Se $X \sim U(I)$, $c \in \mathbb{R}$

$$\Rightarrow p(x \in c) = \frac{\text{mis}(C \cap I)}{\text{mis}(I)}$$

• Se X è assolutamente continuo e $f(x)$ è la densità, si definisce previsione del numero aleatorio X :

$$P(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$

• la quantità $P(x^c)$ si dice momento d'ordine c. la previsione di X è il momento di ordine 1. $P(x^c) = \int_{-\infty}^{+\infty} x^c f(x) dx$

Se indichiamo con $m = P(X)$ si ha:

$$\text{VAR}(X) = P[(X-m)^2] = \int_{-\infty}^{+\infty} (x-m)^2 f(x) dx \quad \text{oppure}$$

$$\text{VAR}(X) = P(X^2) - P^2(X) = \int_{-\infty}^{+\infty} x^2 f(x) dx - \left[\int_{-\infty}^{+\infty} x f(x) dx \right]^2 \quad (1)$$

$\Rightarrow$ se $X \sim U[a, b]$ si ha: $EX = \frac{a+b}{2}$

$$P(x) = \int_{-\infty}^{+\infty} x f(x) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2} \quad \text{notiamo che è il punto medio dell'intervallo.}$$

$$P(x^2) = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b = \frac{b^3 - a^3}{3(b-a)}$$

$$\Rightarrow \text{VAR}(X) = \frac{b^3 - a^3}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{b^2 + ab + a^2}{3} - \frac{b^2 + 2ab + a^2}{4} = \frac{4b^2 + 4ab + 4a^2 - 3b^2 - 6ab - 3a^2}{12} = \frac{b^2 + a^2 - 2ab}{12} = \frac{(a-b)^2}{12} \quad \text{Var} X = \frac{(b-a)^2}{12}$$

• Esempio

14-12-0X

VENERDI

Sia X un numero aleatorio assolutamente continuo tale che

$X \sim U[0, 2]$ e sia Y un altro numero aleatorio ass. continuo tale che $Y = -1 + \pi \cdot X$

Calcolare $p(x, y)$, $\text{cov}(x, y)$ e $p(x, y > 0)$.

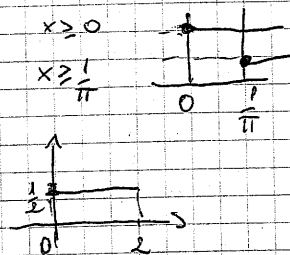
Osserviamo che ovviamente c'è dipendenza tra X e $Y \Rightarrow p(x, y) \neq 1$ Il segno dipende dal coefficiente angolare della retta e in questo caso $\pi > 0 \Rightarrow p(x, y) = 1$.

$$\begin{aligned} \text{COV}(X, Y) &= \text{COV}(X, -1 + \pi X) = \text{COV}(X, -1) + \text{COV}(X, \pi X) = 0 + \pi \text{COV}(X, X) = \pi \text{VAR}(X) = \\ &= \pi \cdot \frac{(2-0)^2}{12} = \pi \cdot \frac{4}{12} = \frac{\pi}{3} \end{aligned}$$

$$P(X \leq 0 \vee X \geq \frac{1}{\pi}) = P(X(-1 + \pi X) \geq 0) = , \quad X(-1 + \pi X) \geq 0 \quad \begin{matrix} x \geq 0 \\ x \geq \frac{1}{\pi} \end{matrix}$$

$$= P(X \leq 0 \vee X \geq \frac{1}{\pi}) = P(X \leq 0) + P(X \geq \frac{1}{\pi}) = ,$$

$$\text{Ricordiamo che } X \sim U[0, 2] \Rightarrow f(x) = \frac{1}{b-a} = \frac{1}{2}$$



$$= 0 + P(X \geq \frac{1}{\pi}) = \int_{\frac{1}{\pi}}^2 \frac{1}{2} dx = \frac{1}{2} (2 - \frac{1}{\pi})$$

• Esempio

Siano date due urne U_1 e U_2 . U_1 contiene 10 palline numerate da 1 a 10 e U_2 contiene 9 palline non numerate. Si estraggono 3 palline da U_1 e vengono messe in U_2 . Dopo si effettuano 5 estrazioni da U_2 :

E_1 = "la pallina 1 viene immessa in U_2 "

E_2 = "la pallina 1 viene estratta da U_2 "

$$P(E_1) = \frac{\binom{9}{2}}{\binom{10}{3}} \leftarrow \begin{matrix} \text{3 termini con una pallina bianca fissata} \\ \text{tutte le 3 palline da } U_1 \end{matrix}$$

Calcolare $P(E_2) = P(E_1 | E_2^c)$

$$P(E_2) = P(E_2 \cap (E_1 \cup E_1^c)) = P(E_2 E_1 \cup E_2 E_1^c) = P(E_2 E_1) + P(E_2 E_1^c) = P(E_2 | E_1) P(E_1) +$$

$$= \frac{\binom{11}{4}}{\binom{12}{5}} \cdot \frac{\binom{9}{2}}{\binom{10}{3}} = \frac{330}{495} \cdot \frac{36}{120} = \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10}$$

$$P(E_1 | E_2^c) = \frac{P(E_1 E_2^c)}{P(E_2^c)} = \frac{P(E_2^c | E_1) P(E_1)}{1 - P(E_2)} = \frac{[1 - P(E_2 | E_1)] P(E_1)}{1 - \frac{3}{10}} = \frac{[1 - \frac{\binom{11}{4}}{\binom{12}{5}}] \cdot \frac{\binom{9}{2}}{\binom{10}{3}}}{1 - \frac{3}{10}} =$$

$$= \frac{(1 - \frac{330}{495}) \cdot \frac{36}{120}}{\frac{7}{10}} = \frac{\frac{165}{495} \cdot \frac{36}{120}}{\frac{7}{10}} = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{7}{10}} = \frac{1}{10} = \frac{1}{5}$$

• Esempio

Un lotto contiene 2 pezzi difettosi e 28 buoni. Vengono presi 4 pezzi. Qual è la probabilità di estrarre due pezzi difettosi consecutivamente?

Fissati i due difettosi i casi possibili sono: (d, d, b, b); (b, d, d, b); (b, b, d, d).

I due buoni si possono scegliere ogni volta in $\binom{28}{2}$ modi $\Rightarrow$

$$p = \frac{3 \cdot \binom{28}{2}}{\binom{30}{4}} = 0,0413$$

• Esempio

Data un'urna contenente 20 palline nere e 30 bianche si effettuano 5 estrazioni senza restituzione

E_i = "pallina bianca alla i -esima estrazione"

$$\text{Oss che: } P(E_1^c E_2^c E_3^c E_4^c E_5 \vee E_1^c E_2^c E_3^c E_5^c E_4^c) = P(E_1^c E_2^c E_3^c E_4^c) =$$

$$= P(E_4 | E_1^c E_2^c E_3^c) \cdot P(E_3^c | E_1^c E_2^c) \cdot P(E_2^c | E_1^c) \cdot P(E_1^c) = \frac{30}{47} \cdot \frac{18}{48} \cdot \frac{19}{49} \cdot \frac{20}{50} \approx 0,034$$

Calcolare la probabilità che le prime due palline siano dello stesso colore sapendo

le delle prime 5, 3 sono bianche.

(29)

X = numero palline bianche estratte.

$$p(E_1 E_2 \cup E_1^c E_2^c | X=3) = \frac{p[(E_1 E_2 \cup E_1^c E_2^c) \cap X=3]}{p(X=3)}$$

Se su 5 palline 3 sono bianche l'evento $E_1 E_2$ può verificarsi in 3 modi mentre l'evento

$E_1^c E_2^c$ può verificarsi in 1 solo modo $\Rightarrow$ 2 palline dello stesso colore si hanno in 4 modi.

3 palline bianche in 5 estrazioni si possono estrarre in $\frac{\binom{30}{3} \binom{20}{2}}{\binom{50}{5}}$ modi

che va diviso per $\binom{5}{3}$ = numero di modo di ordinare le palline

$$\Rightarrow = \frac{4 \cdot \frac{\binom{30}{3} \binom{20}{2}}{\binom{50}{5}} \cdot \binom{5}{3}}{\binom{30}{3} \binom{20}{2}} = 4 \cdot \frac{1}{\binom{5}{3}} = \frac{4}{10} = \frac{2}{5}$$

Cioè le 5 palline le posso estrarre in 20 modi di ordine diversi:

Se si a nostro favore sono i primi tre e l'ultimo.

$$\Rightarrow \frac{4}{10} = \frac{2}{5}$$

b b b n n
b b n b n
b b n n b
b n b n b
b n b m b
b n n b b
n b b b n
n b b n b
n b n b b
n n b b b

• Esempio

Sia X un numero casuale ass. continuo avente distribuzione di probabilità $P.T.a$

$f(x) = c \cdot x^3$ nell'intervallo $[1, 2]$ e $f(x) = 0$ altrove.

Calecolare c :

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \Leftrightarrow \int_1^2 c x^3 dx = 1 \Leftrightarrow c \left[\frac{x^4}{4} \right]_1^2 = 1 \Leftrightarrow c \left(4 - \frac{1}{4} \right) = 1 \Leftrightarrow c = \frac{4}{15}$$

Sia $Y = -2 + 3X$ un altro numero casuale ass. continuo.

Determinare la distribuzione di probabilità $P.T.a$ ~~g(y)~~ $g(y)$.

Ricordiamo che: $g(y) = G'(y)$

$$G(y) = P(Y \leq y) = P(-2 + 3X \leq y) = P(X \leq \frac{y+2}{3}) = *$$

$f(x) \neq 0$ solo in $[1, 2]$ quindi se $(\frac{y+2}{3} < 1 \Leftrightarrow y < 1) \Rightarrow G(y) = 0 \Rightarrow g(y) = 0$

se $1 \leq \frac{y+2}{3} \leq 2 \Leftrightarrow 1 \leq y \leq 4 \Rightarrow$

$$* = \int_1^{\frac{y+2}{3}} c x^3 dx = c \left[\frac{(\frac{y+2}{3})^4}{4} - \frac{1}{4} \right] = G(y) \text{ se } 1 \leq y \leq 4$$

$$\Rightarrow g(y) = \frac{4}{15} \cdot \frac{4}{4} \left(\frac{y+2}{3} \right)^3 \cdot \frac{1}{3} = \frac{4}{15} \left(\frac{y+2}{3} \right)^3 \cdot \frac{1}{3} \text{ se } 1 \leq y \leq 4$$

Per $y > 4 \Rightarrow X = \frac{y+2}{3} > 2 \Rightarrow P(Y \leq y) = 1 = G(y) \Rightarrow G'(y) = g(y) = 0$

• Esempio

$f(x) = \frac{2}{x^3}$ per $x \in [a, \frac{3}{4}]$ e $f(x) = 0$ altrove, $f(x)$ = funzione di densità di X

Calcolare i valori possibili di a :

$$\int_a^{\frac{3}{4}} \frac{2}{x^3} dx = 1 \Leftrightarrow 2 \int_a^{\frac{3}{4}} x^{-2} dx = 2 \left[\frac{x^{-1}}{-1} \right]_a^{\frac{3}{4}} = 2 \left[-\frac{1}{x} \right]_a^{\frac{3}{4}} = 2 \left[-\frac{1}{\frac{3}{4}} + \frac{1}{a} \right] = 1$$

$$\Leftrightarrow 2 \left[-\frac{16}{3} + \frac{1}{a} \right] = 1 \Leftrightarrow -\frac{32}{3} + \frac{2}{a} = 1 \Leftrightarrow \frac{2}{a} = \frac{19}{3} \Leftrightarrow a^2 = \frac{3}{19} \Leftrightarrow a = \pm \sqrt{\frac{3}{19}}$$

$$\Rightarrow a = \sqrt{\frac{3}{19}}$$

Infine calcolare:

$$P\left(X \leq \frac{x}{10} \mid X \geq \frac{13}{20}\right) = \frac{P\left(\frac{13}{20} \leq X \leq \frac{x}{10}\right)}{P\left(X \geq \frac{13}{20}\right)} = \frac{\int_{\frac{13}{20}}^{\frac{x}{10}} \frac{2}{x^3} dx}{\int_{\frac{13}{20}}^{\frac{3}{4}} \frac{2}{x^3} dx}$$

$$\text{Se } \frac{13}{20} < a \Rightarrow P\left(X \leq \frac{x}{10} \mid X \geq \frac{13}{20}\right) = \frac{\int_{\frac{13}{20}}^{\frac{x}{10}} \frac{2}{x^3} dx}{1}$$

$$f(x) = 0 \text{ per } x < a$$

17-12-07
LUNEDÌ

DISTRIBUZIONE NORMALE

Sia X un numero aleatorio continuo che può assumere tutti i valori reali, cioè $X(\Omega) = \mathbb{R}$. Si dice che $X \sim N_{0,1}$, X ha distribuzione normale standard rispetto ai parametri 0 e 1 se la funzione densità di probabilità $f(x)$ è:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

Il suo grafico prende il nome di Gaussiana ed è:

e osserviamo che X è asintoto orizzontale e

l'asse y è asse di simmetria

Indichiamo con $\Phi_{0,1}(x)$ la funzione di ripartizione ed è

$$\Phi_{0,1}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Poiché $f(x)$ è simmetrica rispetto all'asse $y \Rightarrow x > 0$ si ha:

$$\Phi_{0,1}(-x) = 1 - \Phi_{0,1}(x), \text{ ricordando che } \int_{-\infty}^{+\infty} f(x) dx = 1$$

Osserviamo che $P(-3 \leq X \leq 3) \approx 0,9974$ cioè quasi tutta l'area è concentrata nell'intervallo $[-3, 3]$, ciò vuol dire che un numero aleatorio con questa distribuzione ha una probabilità molto piccola "1 - 0,9974 = 0,0026" di assumere valori esterni all'intervallo $[-3, 3]$.

$$P(-x \leq X \leq x) = 2\Phi_{0,1}(x) - 1, \text{ infatti}$$

$$P(-x \leq X \leq x) = \Phi_{0,1}(x) - \Phi_{0,1}(-x) = \Phi_{0,1}(x) - 1 + \Phi_{0,1}(x) = 2\Phi_{0,1}(x) - 1$$

valore per qualsiasi
tipo di funzione di
ripartizione, non solo per Φ

Osserviamo che i parametri 0 e 1 sono la previsione e lo scarto quadratico medio del numero aleatorio X . Infatti:

$$P(X) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0, \text{ perché nel caso di simmetria la precisione coincide con il punto di mezzo medio, in tal caso } 0 \quad (1)$$

$$VAR(X) = \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx = \int_{-\infty}^{+\infty} x \cdot x f(x) dx = \text{integrando per parti, si ottiene } = 1$$

$\Rightarrow \sqrt{VAR(X)} = \sigma = 1$ sostituendo $f(x)$ con la sua espressione e

Se numero aleatorio X è continuo la distribuzione normale di parametri m e σ , $X \sim N_{m,\sigma}$ se la funzione $f(x)$ della densità di probabilità è:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2}$$

Se $m=0$ e $\sigma=1$ otteniamo la normale Standard.

Il grafico di $f(x)$ ha sempre forma a campana ed è simmetrico rispetto alla retta $x=m$.

La funzione di ripartizione è:

$$\Phi_{m,\sigma}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2} dx$$

I suoi valori si possono ottenere attraverso la funzione di ripartizione della normale Standard.

Anche in questo caso si ha $P(X) = m$ e $\sqrt{VAR(X)} = \sigma$.

Dato un numero aleatorio continuo X con $P(X) = m$ e $\sigma = \sqrt{VAR(X)}$, il numero aleatorio $Y = \frac{X-m}{\sigma}$ si dice numero aleatorio ridotto o standardizzato.

Si ha che $P(Y) = 0$ e $VAR(Y) = 1$, infatti:

$$P(Y) = P\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma} (P(X) - m) = \frac{m}{\sigma} - \frac{m}{\sigma} = 0$$

$$Y = \frac{X - E(X)}{\sqrt{VAR(X)}}$$

$$VAR(Y) = VAR\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma^2} VAR(X-m) = \frac{1}{\sigma^2} \cdot \sigma^2 = 1$$

$\Rightarrow Y \sim N_{0,1}$, da qui il nome ridotto o standardizzato.

• Sia $X \sim N_{m,\sigma}$ e $Y = aX + b$. Si prova che $Y \sim N_{m',\sigma'}$.

Cioè anche Y ha distribuzione normale ma con parametri diversi da quelli di X .

(funzione di ripartizione di Y) $= G(y) = P(Y \leq y)$

$g(y) = \text{densità di probabilità}$.

Si ha che $G'(y) = g(y)$.

funzione di ripartizione di X

$$G(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = \Phi_{m,\sigma}\left(\frac{y-b}{a}\right)$$

$$g(y) = G'(y) = \frac{1}{a} \cdot \Phi'_{m,\sigma}\left(\frac{y-b}{a}\right), \text{ (derivata di funzioni composte)}$$

Già sappiamo che $X \sim N_{m,\sigma}$ e $f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2}$

$$\Phi'_{m,\sigma}(x) = f(x) = \text{densità di probabilità} \quad \Downarrow$$

$$\Rightarrow g(x) = \frac{1}{a} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{y-b-m}{\sigma} \right)^2} = \frac{1}{\sqrt{2\pi} \cdot a \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{y-(b+am)}{a\sigma} \right)^2}$$

$$\Rightarrow Y \sim N_{am+b, a\sigma} \text{ e in generale se } Y = ax+b, a \in \mathbb{R} \Rightarrow Y \sim N_{am+b, |a|\sigma}$$

• Se $X \sim N_{m,\sigma}$, $m = P(X)$, $\sigma = \sqrt{VAR(X)}$ si ha:

$$\Phi_{m,\sigma}(x) = P(X \leq x) = P(X-m \leq x-m) = P\left(\frac{X-m}{\sigma} \leq \frac{x-m}{\sigma}\right) = \left(\text{ricordando che } \frac{X-m}{\sigma} \sim N_{0,1} \right) = \Phi_{0,1}\left(\frac{x-m}{\sigma}\right) \quad \left[\Phi_{m,\sigma}(x) = \Phi_{0,1}\left(\frac{x-m}{\sigma}\right) \right]$$

Esercizio Si genera a caso un numero nell'insieme $[1,5]$. Sia $I_1 = [1,4]$, $I_2 = [3,5]$

Calcolare la probabilità che:

a) Il numero sta in $I_1 \cap I_2$

b) Il numero sta in $I_1 \setminus I_2$

c) Il numero sta in $I_2 \setminus I_1$

Sia X : il numero Poiché è generato a caso, vuol dire che ha distribuzione uniforme in tutto l'intervallo $\Rightarrow X \sim U[1,5]$ $b-a=4$

$$f(x) = \begin{cases} \frac{1}{4} & \text{se } x \in [1,5] \\ 0 & \text{se } x \notin [1,5] \end{cases}$$

$$a) X \in I_1 \cap I_2 \Leftrightarrow X \in [3,4], P(X \in [3,4]) = \int_3^4 \frac{1}{4} dx = \frac{1}{4} \left[\frac{x^2}{2} \right]_3^4 = \frac{1}{4} \left(\frac{16}{2} - \frac{9}{2} \right) = \frac{1}{4} \cdot \frac{7}{2} = \frac{7}{8}$$

$$b) X \in I_1 \setminus I_2 \Leftrightarrow X \in [1,3], P(X \in [1,3]) = \frac{1}{4} \cdot 2 = \frac{1}{2}$$

$$c) X \in I_2 \setminus I_1 \Leftrightarrow X \in [4,5], P(X \in [4,5]) = \frac{1}{4} \cdot 1 = \frac{1}{4}$$

• Esercizio In un gioco giocatore e banco lancia entrambi un dado. Il giocatore vince solo se il suo numero è strettamente più grande di quello del banco. Qual'è la probabilità di: $A = \text{"Il giocatore vince"}$

Il casi possibili sono 36.

$\forall n \in [1,6] \cap \mathbb{N}$ ci sono $6-n$ casi favorevoli

$$\Rightarrow \text{Casi favorevoli} = \sum_{n=1}^6 (6-n) = 36 - \sum_{n=1}^6 n = 36 - \frac{6 \cdot 7}{2} = 36 - 21 = 15$$

$$\Rightarrow P(A) = \frac{15}{36} = \frac{5}{12}$$

Esercizio. Due sedi A e B sono collegate con linee di Trasmissione indipendenti (il funzionamento o meno di una non influenza quello di un'altra $\Rightarrow A_i$ indipendenti). Ogni linea è operativa con probabilità dell'80%.

$A_i = \text{"l'i-esima linea è operativa"}$, $i=1, \dots, n$, $P(A_i) = \frac{80}{100}$

Esprimere in funzione di n la probabilità che le due sedi possano comunicare.

Possiamo comunicare se almeno una è operativa $\Rightarrow$ bisogna calcolare $P\left(\bigcup_{i=1}^n A_i\right)$.

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = 1 - P(A_1^c) \cdot P(A_2^c) \cdot \dots \cdot P(A_n^c) = 1 - \left(\frac{20}{100}\right)^n$$

sono indipendenti

Quanto deve valere n affinché questa probabilità sia superiore al 90%?

$$1 - \left(\frac{20}{100}\right)^n > \frac{90}{100} \Rightarrow \frac{1}{5^n} < 1 - \frac{9}{10} = \frac{1}{10} \Rightarrow 5^n > 10 \Rightarrow n \geq 2.$$

Qual'è la probabilità che ci sia almeno una linea interrotta?

$$P(A_1^c \cup A_2^c \cup \dots \cup A_n^c) = 1 - P(A_1 A_2 \dots A_n) = 1 - P(A_1) \cdot P(A_2) \dots P(A_n) = 1 - \left(\frac{80}{100}\right)^n$$

Qual'è la probabilità che ci sia almeno una linea interrotta sapendo che la comunicazione è attiva.

H = "la comunicazione è attiva", B = "almeno una linea è interrotta".

$$P(B|H) = ?$$

$$B = (A_1^c \cup A_2^c \cup \dots \cup A_n^c), \quad H = A_1 A_2 \dots A_n$$

$$\begin{aligned} P(B|H) &= P(A_1^c \cup A_2^c \cup \dots \cup A_n^c | A_1 A_2 \dots A_n) = 1 - P(A_1 A_2 \dots A_n | A_1 A_2 \dots A_n) = \\ &= 1 - \frac{P(A_1 A_2 \dots A_n \cap A_1 A_2 \dots A_n)}{P(A_1 A_2 \dots A_n)} = 1 - \frac{P(A_1 A_2 \dots A_n)}{P(A_1 A_2 \dots A_n)} = 1 - \frac{1 - \left(\frac{20}{100}\right)^n}{1 - \left(\frac{20}{100}\right)^n} \end{aligned}$$

7-01-08

Vettori aleatori discreti.

Siano dati n numeri aleatori $X_1, X_2, \dots, X_n$. Si dice vettore aleatorio discreto l'ensemble $X = (X_1, X_2, \dots, X_n)$. X si dice discreto se il suo codominio C_X è discreto, cioè numerabile o finito.

Una distribuzione di probabilità su X è discreta se:

- 1) $P(X=x) = P(X_1=x_1, X_2=x_2, \dots, X_n=x_n) > 0 \quad \forall x = (x_1, x_2, \dots, x_n) \in C_X$.
- 2) $P(X=x) = 0 \quad \forall x \notin C_X$.

• $P(X_1=x_1, X_2=x_2, \dots, X_n=x_n)$ si dice distribuzione di probabilità congiunta e dà la probabilità che i numeri aleatori assumano certi valori.

Def. La distribuzione di un qualunque sottovettore si chiama distribuzione marginale. La distribuzione congiunta determina univocamente la distribuzione marginale, ma il viceversa non è vero.

• Determinare la distribuzione marginale di un dato evento.

Siano $Y = (X_{i_1}, \dots, X_{i_m})$ e $T = (X_{j_1}, \dots, X_{j_{n-m}})$ due sottovettori di un vettore aleat. X .

Determinare la dist. marginale $\Leftrightarrow$ determinare $\forall y \in C_Y, P(Y=y), y = (x_{i_1}, \dots, x_{i_m})$.

$$\Rightarrow P(Y=y) = P(X_{i_1}=x_{i_1}, \dots, X_{i_m}=x_{i_m}) = P((X_{i_1}=x_{i_1}, \dots, X_{i_m}=x_{i_m}) \cap \Omega) =$$

dove Ω = l'unione di tutti i possibili valori che possono essere assunti da T

$$= P((X_{i_1}=x_{i_1}, \dots, X_{i_m}=x_{i_m}) \cap \left(\bigcup_{x_{j_1}=x_{j_1}, \dots, x_{j_{n-m}}=x_{j_{n-m}}} (X_{j_1}=x_{j_1}, \dots, X_{j_{n-m}}=x_{j_{n-m}}) \right)) =$$

$$= P\left(\bigcup_{x_{j_1}=x_{j_1}, \dots, x_{j_{n-m}}=x_{j_{n-m}}} (X_{i_1}=x_{i_1}, \dots, X_{i_m}=x_{i_m}, X_{j_1}=x_{j_1}, \dots, X_{j_{n-m}}=x_{j_{n-m}}) \right) = \text{gli eventi sono incompatibili.}$$

$$= \sum_{x_{j_1}=x_{j_1}, \dots, x_{j_{n-m}}=x_{j_{n-m}}} P(X_{i_1}=x_{i_1}, \dots, X_{i_m}=x_{i_m}, X_{j_1}=x_{j_1}, \dots, X_{j_{n-m}}=x_{j_{n-m}})$$

"distribuzione congiunta in X ".

Esempio

Sia $(X, Y) \in \{(0,0), (0,1), (0,2), (1,1), (1,2)\}$. Il vettore aleatorio (X, Y) può assumere con la stessa probabilità ogni elemento del suo codominio, $p((X, Y) = (x, y)) = \frac{1}{5}, \forall (x, y) \in C_{(X, Y)}$.
 $\Rightarrow$ in questo caso la distribuzione congiunta è uniforme.

$$X \in \{0, 1\} \text{ e } Y \in \{0, 1, 2\}$$

Determiniamo le distribuzioni marginali

Con lo stesso ragionamento di prima:

$$p(X=0) = p(X=0 \cap \Omega) = p(X=0 \cap (Y=0 \cup Y=1 \cup Y=2)) =$$

$$p((X=0, Y=0) \cup (X=0, Y=1) \cup (X=0, Y=2)) =, \text{ eventi incompatibili,}$$

$$= p(X=0, Y=0) + p(X=0, Y=1) + p(X=0, Y=2) = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5}$$

questa è
la congiunta

questa è
la congiunta

questa è
la congiunta

$$p(X=1) = p(X=1, Y=0) + p(X=1, Y=1) + p(X=1, Y=2) = 0 + \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

$$p(Y=0) = \frac{1}{5}$$

$$p(Y=1) = \frac{2}{5}$$

$$p(Y=2) = \frac{2}{5}$$

$$\bigcup_{Y=0,1,2} p(Y=Y) = 1$$

- Due numeri aleatori X e Y si dicono indipendenti quando la distribuzione congiunta è uguale al prodotto delle distribuzioni marginali. Cioè:

Sia dato il vettore (X, Y)

la dist. congiunta $p(X_i, Y_j) = p(X=X_i, Y=Y_j)$

Sia $p_i^X = p(X=X_i)$ dist. marginale di X

$p_j^Y = p(Y=Y_j)$ dist. marginale di Y

X e Y sono indipendenti se $p_i^X \cdot p_j^Y = p(X_i, Y_j), \forall (X_i, Y_j)$

COVARIANZA

$$\text{COR}(X, Y) = p(X, Y) - p(X) \cdot p(Y)$$

Se X e Y sono indipendenti $\Rightarrow$ sono incoerenti, cioè $\text{COR}(X, Y) = 0$

Esempio

Sia $(X, Y) \in \{(-1,0), (0,0), (1,0), (-1,1), (0,1), (1,1)\}$ con distribuzione di probabilità congiunta $\Rightarrow$

$$\text{COR}(X, Y) = p(X, Y) - p(X) \cdot p(Y)$$

$X \in \{-1, 0, 1\}, Y \in \{0, 1\} \Rightarrow$ le distribuzioni marginali sono:

$$p(X=-1) = \frac{1}{4}$$

$$p(Y=0) = \frac{2}{4} = \frac{1}{2}$$

$$p(X=0) = \frac{5}{10} = \frac{1}{2}$$

$$p(Y=1) = \frac{1}{4}$$

$$p(X=1) = \frac{1}{4}$$

$X, Y \in \{-1, 0, 1\}$ e la sua distribuzione marginale:

$$P(X, Y = -1) = \frac{1}{16}$$

$$P(X, Y = 0) = \frac{7}{8}$$

$$P(X, Y = 1) = \frac{1}{16}$$

$$\Rightarrow P(X) = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0$$

$$P(Y) = 0 \cdot \frac{2}{10} + 1 \cdot \frac{13}{10} = \frac{13}{10}$$

$$P(X, Y) = -1 \cdot \frac{1}{16} + 0 \cdot \frac{7}{8} + 1 \cdot \frac{1}{16} = 0$$

$\Rightarrow \text{COV}(X, Y) = 0 \Rightarrow X$ e Y sono incorrelati, ma vediamo che ~~non~~ sono indipendenti.

$$P(-1, 0) \Rightarrow \frac{1}{4} \cdot \frac{2}{10} \neq \frac{3}{16} \text{ infatti, il prodotto delle marginali non è uguale alla congiunta.}$$

• Calcolare la probabilità di un evento condizionato:

$$P(\min(X, Y) \geq 0 \mid \max(X, Y) < 1) =$$

$$= \frac{P(\min(X, Y) \geq 0, \max(X, Y) < 1)}{P(\max(X, Y) < 1)} = \frac{P(0, 0)}{P(-1, 0) + P(0, 0)} = \frac{\frac{3}{10} \cdot \left(\frac{3}{16} + \frac{3}{10}\right)}{\frac{3}{10} \cdot \frac{39}{16}} = \frac{8}{13}$$

VECTORI ALEATORI ASSOLUTAMENTE CONTINUI

(X, Y) è abs. continuo se:

$$1) P(X=x, Y=y) = 0 \quad \forall (x, y) \in \mathbb{R}^2$$

$$2) \exists f(x, y) \geq 0 \text{ integrabile secondo Riemann: } \forall C \in \mathbb{R}^2, \text{ misurabile secondo}$$

$$\text{Peano-Jordan si ha: } P((X, Y) \in C) = \int_C f(x, y) dx dy$$

La funzione $f(x, y)$ si chiama funzione di densità congiunta e a ~~possere delle~~ ~~verificare~~

$$\text{Se } 1) f(x, y) \geq 0 \quad \forall (x, y) \in \mathbb{R}^2$$

$$2) \int_{\mathbb{R}^2} f(x, y) dx dy = 1 \quad \parallel \Rightarrow f(x, y) \text{ è una densità}$$

Def. Le funzioni di densità marginali:

$$f_1(x) = \int_{\mathbb{R}} f(x, y) dy = \int_{-\infty}^{+\infty} f(x, y) dy$$

$$f_2(y) = \int_{\mathbb{R}} f(x, y) dx = \int_{-\infty}^{+\infty} f(x, y) dx$$

Diciamo che due vettori aleatori X e Y si dicono indipendenti se la densità congiunta è uguale al prodotto delle marginali $\Leftrightarrow f(x, y) = f_1(x) \cdot f_2(y), \forall (x, y) \in \mathbb{R}^2$

Se X e Y sono indipendenti $\Rightarrow$ sono incorrelati, ma non vale il viceversa.

Dim $\Rightarrow$

$$\text{COV}(X, Y) = P(X, Y) - P(X) \cdot P(Y)$$

$$P(X) = \int_{-\infty}^{+\infty} x \cdot f_1(x) dx \quad ; \quad P(Y) = \int_{-\infty}^{+\infty} y \cdot f_2(y) dy$$

$$P(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f(x, y) dx dy$$



Per ipotesi x e y sono indipendenti $\Rightarrow$

$$P(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_1(x) \cdot f_2(y) dx dy =$$

$$= \int_{-\infty}^{+\infty} x \cdot f_1(x) dx \cdot \int_{-\infty}^{+\infty} y \cdot f_2(y) dy = P(x) \cdot P(y)$$

$\Rightarrow P(x, y) - P(x) \cdot P(y) = 0 = \text{cov}(x, y) \Rightarrow x$ e y sono incorrelati.

18-01-08
VENERDI

DISTRIBUZIONE GEOMETRICA

X = numero di prove fino ad ottenere il 1° successo

A_i = "successo alla i -esima prova"

p =
Data una successione di eventi $A_1, \dots, A_n$, indipendenti ed equiprobabili, con probabilità p , si dice distribuzione geometrica (dipendente dal parametro p), quella del numero X di prove fino ad ottenere il 1° successo.

A_i = "successo alla i -esima prova"

$$X = 1 \cdot |A_1| + 2 \cdot |A_2| + \dots + n \cdot |A_n| + \dots$$

$$p = P(A_i) \quad , \quad X \in \{1, 2, \dots, n, \dots\}$$

$$P(X=1) = P(A_1) = p \quad ; \quad q = P(A_i^c) = 1-p$$

$$P(X=2) = P(A_1^c \cdot A_2) = P(A_1^c) \cdot P(A_2) = q \cdot p$$

$$P(X=n) = P(A_1^c \cdot A_2^c \cdot \dots \cdot A_{n-1}^c \cdot A_n) = P(A_1^c) \cdot P(A_2^c) \cdot \dots \cdot P(A_{n-1}^c) \cdot P(A_n) = q^{n-1} \cdot p$$

Proprietà di assenza di memoria.

$$P(X > n+m | X > n) = P(X > m)$$

Cioè le prime n prove non influenzano il risultato delle prove successive.

Ad esempio nel gioco del lotto il fatto che un numero non viene estratto da molto tempo non aumenta la sua probabilità di uscire nelle prossime estrazioni.

TEO. Un numero aleatorio discreto ha distribuzione geometrica se e solo se vale la proprietà di assenza di memoria.

Dim.

H_0 vale la proprietà di assenza di memoria

Tesi: $X \sim G(p)$, allora c'è un libro pag 48 ma forse non bisogna farla

H_p $X \sim G(p)$

Tesi vale la prop. di assenza di memoria.

Vedi anche [pg 55 Tila]

$$P(X > n) = P(X = n+1 \vee \dots \vee X = t \vee \dots) = \sum_{t=n}^{+\infty} q^t \cdot p = 1 - P(X \leq n) =$$

$$= 1 - P(X=1 \vee \dots \vee X=n) = 1 - \sum_{t=1}^n q^{t-1} \cdot p = 1 - p \sum_{t=1}^n q^{t-1} = 1 - p \frac{1-q^n}{1-q} = 1 - \frac{p}{p} (1-q^n) =$$

$$= 1 - 1 + q^n = q^n$$

$$\text{Ovvero } \sum_{t=n}^{+\infty} q^t \cdot p = q^n$$

$P(X > n) \Leftrightarrow$ nelle prime n prove non si ha avuto successo

$$\Rightarrow P(X > n) = P(A_1^c \cdot \dots \cdot A_n^c) = q^n$$

$$P(X > n+m | X > n) = \frac{P(X > n+m, X > n)}{P(X > n)} = \frac{P(X > n+m)}{P(X > n)} = \frac{q^{n+m}}{q^n} = q^m$$

ovvero che $X > n+m$ implica $X > n$
 $\Rightarrow (X > n+m) \wedge (X > n) = X > n+m$

$$= \frac{P(X > n+m)}{P(X > n)} = \frac{q^{n+m}}{q^n} = q^m$$

$\Rightarrow P(X > n+m | X > n) = P(X > m) \Rightarrow$ vale la proprietà di assenza di memoria e.v.d.

Dato $X \sim G(p)$, calcoliamo $P(X)$ e $VAR(X)$.

$$P(X) = \sum_{n=1}^{\infty} n \cdot q^{n-1} \cdot p = p \sum_{n=1}^{\infty} n \cdot q^{n-1} = p \sum_{n=1}^{\infty} \frac{d}{dq} (q^n) = p \frac{d}{dq} \left(\sum_{n=1}^{\infty} q^n \right) = p \frac{d}{dq} \left(q \sum_{n=1}^{\infty} q^{n-1} \right) =$$

$$= p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \frac{(1-q) - q(-1)}{(1-q)^2} = p \frac{1}{(1-q)^2} = \frac{p \cdot 1}{p^2} = \frac{1}{p}$$

$$VAR(X) = P(X^2) - P^2(X)$$

$$P(X^2) = \sum_{n=1}^{\infty} n^2 \cdot q^{n-1} \cdot p = \sum_{n=1}^{\infty} n(n-1+1) q^{n-1} \cdot p = \sum_{n=1}^{\infty} n(n-1) q^{n-1} \cdot p + \sum_{n=1}^{\infty} n \cdot q^{n-1} \cdot p =$$

$$= p q \sum_{n=1}^{\infty} \frac{d^2}{dq^2} (q^n) + P(X) = p q \frac{d^2}{dq^2} \left(q \cdot \sum_{n=1}^{\infty} q^{n-1} \right) + \frac{1}{p} = p q \frac{d^2}{dq^2} \left(\frac{q}{1-q} \right) + \frac{1}{p} =$$

$$= p q \frac{d}{dq} \left(\frac{1}{(1-q)^2} \right) + \frac{1}{p} = p \cdot q \cdot \frac{2(1-q)}{(1-q)^4} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}$$

$$VAR(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q + p - 1}{p^2} = \frac{2q + q}{p^2} = \frac{q}{p^2}$$

DISTRIBUZIONE DI PASCAL

Sia X = numero di prove fino ad ottenere k successi.

DSS. da distribuzione geometrica è un caso particolare della distribuzione di Pascal che si ottiene per $k=1$.

A_i = "successo alla i -esima prova". Gli A_i sono equiprobabili p e indipendenti $\Rightarrow$
 $\Rightarrow p = p(A_i)$ e $q = 1-p = p(A_i^c)$

$$X \in \{k, k+1, \dots, n, \dots\}$$

$P(X=k) =$, si sono effettuate k prove tutte di successo, $\Rightarrow$
 $= P(A_1 \cdot A_2 \cdot \dots \cdot A_k) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_k) = p^k$

$P(X=k+1) = k \cdot q \cdot p^k$

perché k successi ~~tra~~ su $k+1$ si possono scegliere in $k+1$ modi ma dobbiamo escludere il caso in cui l'insuccesso si ha alla $k+1$ -esima prova $\Rightarrow$ rimangono k modi.

$P(X=n) = \binom{n-1}{k-1} p^k \cdot q^{n-k} = \binom{n-1}{n-k} p^k \cdot q^{n-k}$, per lo stesso motivo $\uparrow$
 $\Rightarrow$ come mettere $n-k$ insuccessi in $n-1$ posti.

come mettere $k-1$ successi in $n-1$ posti.
 $\left[\begin{matrix} \text{ultimo successo} \\ \text{è posto all'ultima} \\ \text{prova} \end{matrix} \right] \binom{n-1}{n-k} = \frac{(n-1)!}{(n-k)! (k-1)!} \Rightarrow \binom{n-1}{k-1} = \binom{n-1}{n-k}$

In generale $\binom{n}{k} = \binom{n}{n-k}$ = formula di Vandermonde.

Introduco i numeri aleatori $X_1, X_2, \dots, X_k$:

X_1 = numero di prove per ottenere il primo successo

X_2 = numero di prove per ottenere il secondo successo a partire dal primo

X_k = numero di prove per ottenere il k -esimo successo a partire dal $(k-1)$ -esimo.

Ma allora $X = X_1 + X_2 + \dots + X_K$

Osserviamo che abbiamo ciascun X_i a distribuzione geometrica di parametro p ,
 $X_i \sim G(p)$ ed inoltre $X_1, X_2, \dots, X_K$ sono indipendenti perché riguardano eventi indipendenti.

$$\Rightarrow P(X) = P(X_1 + X_2 + \dots + X_K) = P(X_1) \cdot P(X_2) \cdot \dots \cdot P(X_K) = \frac{1}{p} \cdot \frac{1}{p} \cdot \dots \cdot \frac{1}{p} = \frac{1}{p^K}$$

$$VAR(X) = VAR(X_1 + \dots + X_K) = \sum_{i=1}^K VAR(X_i) = \sum_{i=1}^K \frac{1-p}{p^2} = \frac{K \cdot (1-p)}{p^2}$$

$$E(X) = \frac{K}{p}, \quad VAR(X) = \frac{K \cdot (1-p)}{p^2}$$

DISTRIBUZIONE DI POISSON

La distribuzione di Poisson nasce per calcolare il numero di morti dovuti a colpi di cannone nell'esercito prussiano.

La distribuzione di Poisson, dipendente da un parametro $\lambda > 0$, è quella di un numero aleatorio $X \in \{0, 1, \dots, n\}$ ^{che può assumere valori del dominio} con probabilità:

$$P(X=n) = \frac{\lambda^n}{n!} \cdot e^{-\lambda}, \quad X \sim P(\lambda)$$

~~Sotto~~ sotto le condizioni che $\lambda > 0$:

- 1) La probabilità di avere un arrivo in un intervallo di ampiezza h è circa $\lambda \cdot h$;
- 2) la probabilità di avere 2 o più arrivi in un intervallo di ampiezza h è trascurabile rispetto alla probabilità di avere un solo arrivo.
- 3) Gli arrivi in intervalli disgiunti sono indipendenti.

Questo tipo di distribuzione di probabilità si utilizza per calcolare il numero di telefonate ad un centralino, gli arrivi di un cliente ad uno sportello, ecc. in un intervallo di tempo.

Sia I un intervallo di tempo $[0, t]$. Vogliamo calcolare $P(X=k)$ in $[0, t]$, cioè la probabilità che nell'intervallo di tempo $[0, t]$ ci siano k arrivi. Dividiamo l'intervallo I in n sottointervalli di ampiezza $\frac{t}{n}$. Possiamo supporre che $n > k$ perché per o quanto meno in un intervallo ci sia un solo arrivo.

Per le condizioni 1), 2) e 3) si ha:

$$P(X=k) \stackrel{\text{binomiale}}{\approx} \binom{n}{k} \cdot (\lambda \cdot \frac{t}{n})^k \cdot (1 - \lambda \cdot \frac{t}{n})^{n-k} = \binom{n}{k} \left(\frac{\lambda \cdot t}{n}\right)^k \cdot \left(1 - \frac{\lambda \cdot t}{n}\right)^{n-k}$$

Se rimpiccioliamo gli intervalli, cioè $n \rightarrow \infty$

$$\Rightarrow P(X=k) = \lim_{n \rightarrow \infty} \binom{n}{k} \left(\frac{\lambda \cdot t}{n}\right)^k \left(1 - \frac{\lambda \cdot t}{n}\right)^{n-k} =$$

$$= \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2) \dots (n-k+1)}{k!} \cdot \frac{(\lambda \cdot t)^k}{n^k} \cdot \frac{\left(1 - \frac{\lambda \cdot t}{n}\right)^n}{\left(1 - \frac{\lambda \cdot t}{n}\right)^k} = \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2) \dots (n-k+1)}{k!} \cdot \frac{(\lambda \cdot t)^k}{n^k} \cdot \left(1 - \frac{\lambda \cdot t}{n}\right)^n$$

$$= \frac{(\lambda \cdot t)^k}{k!} \cdot \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2) \dots (n-k+1)}{n^k} \cdot \left(1 - \frac{\lambda \cdot t}{n}\right)^n = \frac{(\lambda \cdot t)^k}{k!} \cdot e^{-\lambda \cdot t}$$

che è la distribuzione di Poisson con $\lambda = \lambda \cdot t$

Calcoliamo la ~~per~~ $P(X)$ quando $X \sim P(\lambda)$

$$P(X) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

(39)

• Dati i numeri aleatori X_1 e X_2 tali che $X_1 \sim P(\lambda_1)$ e $X_2 \sim P(\lambda_2)$ con X_1 e X_2 indipendenti, proviamo che il numero aleatorio $Y = X_1 + X_2$ ha distribuzione $Y \sim P(\lambda_1 + \lambda_2)$.

$$Y \in \{0, 1, \dots, n\}$$

$$\begin{aligned} P(Y=n) &= P(X_1 + X_2 = n) = \sum_{m=0}^n P(X_1=m, X_2=n-m) = \sum_{m=0}^n P(X_1=m) \cdot P(X_2=n-m) \\ &= \sum_{m=0}^n \frac{\lambda_1^m}{m!} e^{-\lambda_1} \cdot \frac{\lambda_2^{n-m}}{(n-m)!} e^{-\lambda_2} = e^{-\lambda_1 - \lambda_2} \sum_{m=0}^n \frac{\lambda_1^m \cdot \lambda_2^{n-m}}{m! (n-m)!} \\ &= e^{-\lambda_1 - \lambda_2} \sum_{m=0}^n \frac{n!}{m! (n-m)!} \frac{\lambda_1^m \lambda_2^{n-m}}{n!} = e^{-\lambda_1 - \lambda_2} \sum_{m=0}^n \binom{n}{m} \frac{\lambda_1^m \lambda_2^{n-m}}{n!} \\ &= \frac{e^{-\lambda_1 - \lambda_2}}{n!} \sum_{m=0}^n \binom{n}{m} \lambda_1^m \lambda_2^{n-m} = \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n \Rightarrow Y \sim P(\lambda_1 + \lambda_2) \end{aligned}$$

"binomio di Newton"

• Supponiamo di avere $X_1 \sim P(\lambda_1)$ e $X_2 \sim P(\lambda_2)$. Qual'è la distribuzione di $X_1 / (X_1 + X_2 = n)$.

$$X_1 \in \{0, 1, 2, \dots, n\} \text{ perché } X_2 \geq 0, X_1 \leq n - X_2$$

Dobbiamo calcolare $P(X_1=h | (X_1 + X_2 = n))$, dove $h \in \{0, 1, \dots, n\}$

$$\begin{aligned} P(X_1=h | X_1 + X_2 = n) &= \frac{P(X_1=h, X_1 + X_2 = n)}{P(X_1 + X_2 = n)} = \frac{P(X_1=h, X_2=n-h)}{P(X_1 + X_2 = n)} \\ &= \frac{P(X_1=h) \cdot P(X_2=n-h)}{P(X_1 + X_2 = n)} = \frac{\frac{\lambda_1^h}{h!} e^{-\lambda_1} \cdot \frac{\lambda_2^{n-h}}{(n-h)!} e^{-\lambda_2}}{\frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1 + \lambda_2)}} \\ &= \frac{\lambda_1^h \lambda_2^{n-h}}{(\lambda_1 + \lambda_2)^{n-h-h}} \cdot \frac{n!}{(n-h)! h!} = \binom{n}{h} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^h \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-h} = \binom{n}{h} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^h \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{n-h} \end{aligned}$$

$\Rightarrow X_1 / (X_1 + X_2 = n)$ ha distribuzione binomiale, $X_1 / (X_1 + X_2 = n) \sim B(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$

DISTRIBUZIONE UNIFORME IN $\mathbb{R}^2$

21-01-08
LUXEN

Un vettore aleatorio (x, y) si dice che ha distribuzione uniforme in $C \subseteq \mathbb{R}^2$ se la sua funzione di densità è fatta così:

$$f(x, y) = \begin{cases} k & \text{se } (x, y) \in C \\ 0 & \text{se } (x, y) \notin C \end{cases}$$

La funzione di densità deve verificare le due seguenti proprietà:

- 1) $f(x, y) \geq 0$
- 2) $\int_{\mathbb{R}^2} f(x, y) dx dy = 1$

Calcoliamo k affinché f sia una densità:

$$\int_{\mathbb{R}^2} f(x, y) dx dy = 1$$

$$\text{ma: } \int_{\mathbb{R}^2} f(x,y) dx dy = \int_C k dx dy = k \int_C dx dy = k \cdot \text{mis}(C)$$

$$\Rightarrow k \cdot \text{mis}(C) = 1 \Rightarrow k = \frac{1}{\text{mis}(C)}$$

• Esempio

Sia (X,Y) un vettore aleatorio con distribuzione uniforme in $C = x^2 + y^2 \leq 1$, cerchio unitario, con centro nell'origine. Verificare che x e y sono ^{non} indipendenti.

$$\text{mis}(C) = \pi \cdot 1^2 = \pi$$

$$\Rightarrow f(x,y) = \begin{cases} \frac{1}{\pi} & \text{se } (x,y) \in C \\ 0 & \text{se } (x,y) \notin C \end{cases}$$

Calcoliamo le funzioni densità marginali:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y) dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{1}{\pi} \cdot 2\sqrt{1-x^2}, \quad x \in [-1, 1]$$

$$f_1(x) = 0 \quad \text{se } x \notin [-1, 1]$$

$$f_2(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \frac{2}{\pi} \sqrt{1-y^2}, \quad y \in [-1, 1]$$

$$\Rightarrow f(x,y) \neq f_1(x) \cdot f_2(y) \Rightarrow x \text{ e } y \text{ non sono indipendenti.}$$

Vediamo che sono incorrelati (confermando che incorrelati $\nRightarrow$ indipendenza).

$$\text{COV}(X,Y) = P(X,Y) - P(X) \cdot P(Y)$$

$$P(X) = \int_{-1}^1 x \cdot \frac{2\sqrt{1-x^2}}{\pi} dx = 0, \quad \text{perché è l'integrale di una funzione dispari in un intervallo simmetrico e analogamente si ha}$$

$$P(Y) = \int_{-1}^1 y \cdot \frac{2\sqrt{1-y^2}}{\pi} dy = 0$$

$$P(X,Y) = \iint_{\mathbb{R}^2} x \cdot y \cdot \frac{1}{\pi} dx dy = \frac{1}{\pi} \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} x \cdot y dx dy = 0 \Rightarrow \text{COV}(X,Y) = 0 \Rightarrow \text{sono incorrelati.}$$

DISTRIBUZIONE ESPONENZIALE

Sia X un numero aleatorio continuo che indica il tempo di durata fino al guasto di un certo dispositivo. Diremo che X ha distribuzione esponenziale e lo indichiamo con $X \sim \text{EXP}(\lambda)$, $\lambda > 0$ se la sua densità è fatta così:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Calcoliamo la probabilità che la macchina si guasti tra 0 e $X \Rightarrow$

$$\Leftrightarrow P(X \leq x) = \int_0^x f(x) dx = \int_0^x \lambda \cdot e^{-\lambda x} dx = [-e^{-\lambda x}]_0^x = 1 - e^{-\lambda x} = F(x) = \text{funzione di ripartizione}$$

$$P(X > x) = 1 - F(x) = 1 - 1 + e^{-\lambda x} = e^{-\lambda x} = G(x) = \text{funzione di sopravvivenza.}$$

$$P(x) = \int_0^{+\infty} x \cdot \lambda \cdot e^{-\lambda x} dx = [x \cdot (-e^{-\lambda x})]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = 0 - \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^{+\infty} = \frac{1}{\lambda}$$

per parti

$$\begin{aligned} f(x) &= x & g'(x) &= \lambda \cdot e^{-\lambda x} \\ f'(x) &= 1 & g(x) &= -e^{-\lambda x} \end{aligned}$$

$VAR(X) = P(X^2) - P^2(X)$

$P(X^2) = \int_0^{+\infty} x^2 \cdot \lambda \cdot e^{-\lambda x} dx = [x^2 \cdot (-e^{-\lambda x})]_0^{+\infty} + 2 \int_0^{+\infty} x \cdot e^{-\lambda x} dx = 0 + \frac{2}{\lambda} \int_0^{+\infty} x \cdot \lambda \cdot e^{-\lambda x} dx =$

per parti: $f(x) = x^2 \quad g'(x) = \lambda \cdot e^{-\lambda x}$
 $f'(x) = 2x \quad g(x) = -e^{-\lambda x}$

$= \frac{2}{\lambda} P(X) = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2} \Rightarrow VAR(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

OSS Nel discreto la proprietà di assenza di memoria vale $\Leftrightarrow$ il numero ha distribuzione geometrica. ~~Tuttavia~~ Notiamo che:

Anche in questo caso è valida la proprietà di assenza di memoria. $\downarrow$

$P(X > x + x_0 | X > x) = P(X > x_0)$

$H_0: X \sim ESP(\lambda) \quad , \quad Tesi: P(X > x + x_0 | X > x) = P(X > x_0)$

3 dispositivi considerati non sono soggetti ad usura

Dim.

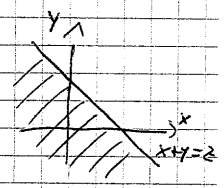
$P(X > x + x_0 | X > x) = \frac{P(X > x + x_0, X > x)}{P(X > x)} = \frac{P(X > x_0 + x)}{P(X > x)} = \frac{S(x + x_0)}{S(x)} = \frac{e^{-\lambda(x+x_0)}}{e^{-\lambda x}} = e^{-\lambda x_0} = S(x_0) = P(X > x_0) \quad c.r.d.$

INTEGRALE DI CONVOLUZIONE

Sia dato il vettore (x, y) con densità $f(x, y)$ e siano x e y indipendenti.

Determiniamo la distribuzione di $Z = x + y$.

$G(z) =$ funzione di ripartizione $= P(Z \leq z) = P(x + y \leq z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f(x, y) dx dy =$



Cambio variabile, ponendo $(x+y) = t \quad , \quad y = t-x$

$= \int_{-\infty}^{+\infty} \int_{-\infty}^z f(x, t-x) dx dt = \int_{-\infty}^{+\infty} \int_{-\infty}^z f_1(x) \cdot f_2(t-x) dx dt =$

$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f_1(x) \cdot f_2(t-x) dx \right] dt \Rightarrow G'(z) =$ funz. di densità Z $= \int_{-\infty}^{+\infty} f_1(x) \cdot f_2(t-x) dx$
poiché x e y sono indipendenti
integrale di convoluzione

Una distribuzione di probabilità ha una struttura del tipo:

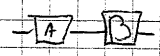
$F(x) = \int_{-\infty}^x f(x) dx$

$\Rightarrow$ la distribuzione di z è l'integrale di convoluzione, cioè:

$g(z) = \int_{-\infty}^{+\infty} f_1(x) \cdot f_2(z-x) dx = \int_{-\infty}^{+\infty} f_1(t-y) \cdot f_2(y) dy$
variabile proprietà commutativa

Applicazioni

● Dispositivi in serie



Tutto il dispositivo si guasta quando se ne guasta uno. Siano:

$X =$ Tempo di durata fino al guasto di A, $X \sim ESP(\lambda_1)$

$Y =$ Tempo di durata fino al guasto di B, $Y \sim ESP(\lambda_2)$

$Z =$ Tempo di durata di tutto il dispositivo $\Rightarrow Z = \min(x, y)$

u
u
u
u
u
u
u
u
u
u

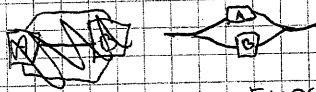
Q22) Determiniamo la sua distribuzione

$$G(z) = P(Z \leq z) = P(\min(X, Y) \leq z) = 1 - P(\min(X, Y) > z) = 1 - P(X > z, Y > z) = 1 - P(X > z) \cdot P(Y > z) = 1 - S(x) S(y) = 1 - e^{-\lambda_1 z} \cdot e^{-\lambda_2 z} = 1 - e^{-(\lambda_1 + \lambda_2)z}$$

perché x e y sono indip.

$$g(z) = G'(z) = (\lambda_1 + \lambda_2) \cdot e^{-(\lambda_1 + \lambda_2)z}, z \geq 0 \Rightarrow Z \sim \text{ESP}(\lambda_1 + \lambda_2)$$

Dispositivi in parallelo.



Siano X e Y gli stemi di prima, $X \sim \text{EXP}(\lambda_1)$ e $Y \sim \text{EXP}(\lambda_2)$ e X e Y indipendenti.

Z = Tempo di durata di tutto il dispositivo.

Poiché il dispositivo è collegato in parallelo, si guasta quando si guastano tutti e due

$$\Rightarrow Z = \max(X, Y)$$

Determiniamo la distribuzione di Z .

$$G(z) = P(Z \leq z) = P(\max(X, Y) \leq z) = P(X \leq z, Y \leq z) = P(X \leq z) \cdot P(Y \leq z) = (1 - e^{-\lambda_1 z}) \cdot (1 - e^{-\lambda_2 z})$$

$$g(z) = G'(z) = \lambda_1 e^{-\lambda_1 z} (1 - e^{-\lambda_2 z}) + (1 - e^{-\lambda_1 z}) \cdot \lambda_2 e^{-\lambda_2 z} = \lambda_1 \cdot e^{-\lambda_1 z} + \lambda_2 \cdot e^{-\lambda_2 z} - (\lambda_1 + \lambda_2) \cdot e^{-(\lambda_1 + \lambda_2)z}$$

Dispositivi in parallelo sequenziale.

In questo tipo di collegamento uno inizia a funzionare quando si guasta l'altro

$$\Rightarrow Z = X + Y$$

Per calcolare l'integrale la distribuzione di Z usiamo il Teo di convoluzione, infatti ne sono verificate le ipotesi.

$$g(z) = \int_{-\infty}^{+\infty} f_1(x) \cdot f_2(z-x) dx$$

$$f_1(x) = \begin{cases} \lambda_1 \cdot e^{-\lambda_1 x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$f_2(z-x) = \begin{cases} \lambda_2 \cdot e^{-\lambda_2(z-x)} & \text{se } z-x \geq 0 \\ 0 & \text{se } z-x < 0 \Leftrightarrow z \leq x \end{cases}$$

$$\Rightarrow g(z) = \int_0^z \lambda_1 \cdot e^{-\lambda_1 x} \cdot \lambda_2 \cdot e^{-\lambda_2(z-x)} dx = \lambda_1 \cdot \lambda_2 \cdot e^{-\lambda_2 z} \cdot \int_0^z e^{-(\lambda_1 + \lambda_2)x} dx = \frac{-\lambda_1 \lambda_2 \cdot e^{-\lambda_2 z}}{(\lambda_1 + \lambda_2)} \int_0^z -(\lambda_1 + \lambda_2) \cdot e^{-(\lambda_1 + \lambda_2)x} dx = \frac{-\lambda_1 \lambda_2 \cdot e^{-\lambda_2 z}}{(\lambda_1 + \lambda_2)} (e^{-(\lambda_1 + \lambda_2)z} - 1) = \frac{\lambda_1 \lambda_2 \cdot e^{-\lambda_2 z}}{(\lambda_1 + \lambda_2)} (1 - e^{-(\lambda_1 + \lambda_2)z})$$

da funzione di ripartizione è $g(z) = G'(z)$, si devono fare i calcoli.

TEOREMA DI BAYES (È molto utilizzato in Statistica)

Viene utilizzato per calcolare i difetti prodotti da dei macchinari.

Supponiamo che sia data una partizione $\{H_1, H_2, \dots, H_n\}$ di Ω . Ad esempio Ω può rappresentare un'urna ^{che contiene} (con palline di colore diverso di cui non si conosce il numero) $\Rightarrow$ gli eventi $H_1, \dots, H_n$ rappresentano le varie ipotesi che si possono

fare sulla composizione dell'urna ^{supponiamo che} e ~~ognuna~~ di queste ipotesi si può verificare con probabilità $p_1 = P(H_1), \dots, p_n = P(H_n)$. $(p_1, \dots, p_n)$ viene detta distribuzione iniziale o distribuzione a priori o probabilità iniziale (queste p_i vengono assegnate inizialmente cioè quando non si sa nulla sulla composizione dell'urna). Dopo effettuare un esperimento E e indichiamo con $E = \text{"risultato dell'esperimento"}$.

Nell'esempio dell'urna l'esperimento consiste nell'estrarre una pallina dalle palline, non tutte, e vedere di che colore sono. Lo scopo del Teorema di Bayes è quello di riguardare le probabilità iniziali alla luce del risultato dell'esperimento E cioè calcolare $P(H_c | E)$, $c=1, \dots, n$, che vengono dette probabilità finali.

Usiamo il Teo. delle probabilità composte

$$P(H_c | E) = \frac{P(E | H_c)}{P(E)} = \frac{P(E | H_c) \cdot P(H_c)}{P(E)}$$

$$P(E) = P(E \cap \Omega) = P(E \cap (H_1 \cup H_2 \cup \dots \cup H_n)) = \sum_{c=1}^n P(E \cap H_c) = \sum_{c=1}^n P(E | H_c) \cdot P(H_c)$$

$$\Rightarrow P(H_c | E) = \frac{P(E | H_c) \cdot P(H_c)}{\sum_{c=1}^n P(E | H_c) \cdot P(H_c)}$$

Def

$P(E | H_c)$ viene detta verosimiglianza perché tanto più è alta $P(E | H_c)$ tanto più è corretta l'ipotesi H_c .

Dopo aver calcolato le probabilità finali si avrà hanno e così:

- 1) Alcune ipotesi si rafforzano di più, cioè assumono maggiore considerazione rispetto a quanto ne avevano prima dell'esperimento
- 2) Altre ipotesi invece diventano più probabili.

L'ipotesi che si rafforza di più è data da quella H_c tale che (1)

$$\frac{P(H_c | E)}{P(H_c)} = \max_c \frac{P(H_c | E)}{P(H_c)} = \max_c \left(\frac{1}{P(E)} \cdot \frac{P(E | H_c) \cdot P(H_c)}{P(H_c)} \right) = \max_c \frac{P(E | H_c)}{P(E)} = \max_c P(E | H_c)$$

= $\max_c P(E | H_c)$, cioè quella che si rafforza di più è quella più verosimile.

Quella che diventa più probabile invece è quella H_c tale che (2)

$$P(H_c | E) = \max_c P(H_c | E) = \max_c c \cdot P(E | H_c) \cdot P(H_c) = \max_c P(E | H_c) \cdot P(H_c)$$

Nota: se $P(H_c) = \text{costante } k$, allora la (1) e (2) coincidono.

Sia $X \sim \text{Exp}(\lambda)$

$$\Rightarrow f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

22-01-08
MARTEDI

Ricordiamo la proprietà di assenza di memoria $P(x > x_1 + x_2 | x > x_1) = P(x > x_2)$

Dimostriamo che se x è un numero aleatorio continuo, $x \geq 0$ e vale la proprietà di assenza di memoria $\Rightarrow X \sim \text{Exp}(\lambda)$

si ha che:

$$P(X > x+x_0 | X > x) = \frac{P(X > x+x_0)}{P(X > x)} = P(X > x_0) \Leftrightarrow$$

[Vedi libro pg 100]

$$\Leftrightarrow S(x+x_0) = S(x_0) \cdot S(x)$$

Osserviamo che x_0 è fisso e x è variabile.

Osserviamo che $\frac{S'(x)}{S(x)} \leq 0$ perché $S(x) \geq 0$ essendo una probabilità e

$S'(x) < 0$ perché $S(x)$ è non crescente $\Rightarrow \frac{S(x)}{S'(x)} = -d(x)$, con $d(x) \geq 0$

Osserviamo ancora che

$$S'(x+x_0) = S(x_0) \cdot S'(x) = \frac{S'(x+x_0)}{S'(x+x_0)} = \frac{S(x_0)}{S(x_0)} \cdot \frac{S'(x)}{S'(x)} = \frac{S'(x)}{S'(x)} = -d(x)$$

$$\Rightarrow -d(x+x_0) = -d(x) \Rightarrow -d(x) \text{ è costante} \Rightarrow -d(x) = -d \Rightarrow$$

$$\frac{S'(x)}{S(x)} = -d \Leftrightarrow \ln[S(x)] = -dx \quad \begin{matrix} \text{integrando} \\ \text{eq. diff.} \end{matrix} \Leftrightarrow \ln(S(x)) = -dx + c \Leftrightarrow$$

$$S(x) = c_1 \cdot e^{-dx}$$

$$\text{Quanto vale } c_1? \text{ Oss. che } S(0) = 1 \Rightarrow 1 = c_1 \cdot e^0 \Rightarrow c_1 = 1$$

$$\Rightarrow S(x) = 1 \cdot e^{-dx} = e^{-dx}$$

$$\Rightarrow X \sim \text{EXP}(d)$$

$$\text{Adda se } X \sim \text{EXP}(d) \Rightarrow P(X > x) = e^{-dx} \Leftrightarrow F(x) = P(X \leq x) = 1 - e^{-dx}; f(x) = F'(x) = d \cdot e^{-dx}$$

$$\bullet P(X > x+x_0 | X > x) \leq P(X > x_0)$$

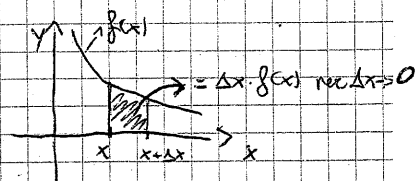
Nel caso $\oplus$ si ha la proprietà di assenza di memoria $\Leftrightarrow X \sim \text{EXP}(\lambda)$

Nel caso $\ominus$ il dispositivo ringiovanisce

Nel caso $\oplus$ c'è usura, il dispositivo invecchia.

FUNZIONE DI RISCHIO

$$P(X \leq x+\Delta x | X > x) = \frac{P(x < X \leq x+\Delta x)}{P(X > x)} \sim \frac{\Delta x \cdot f(x)}{S(x)}$$



Si dice funzione di rischio $h(x) = \frac{f(x)}{S(x)}$; $X \geq 0$

Osserviamo che se: $h(x)$ è crescente $\Rightarrow$ si ha invecchiamento perché $P(X \leq x+\Delta x | X > x)$ è crescente
 $h(x)$ è decrescente $\Rightarrow$ si ha ringiovanimento.

si spiega ricordando che $P(X \leq x+\Delta x | X > x) \approx \Delta x \cdot h(x)$.

Notiamo che c'è una corrispondenza biunivoca tra densità di probabilità e funzione di rischio. Infatti se è data $f(x) \Rightarrow h(x)$ è univocamente determinata perché:

$$h(x) = \frac{f(x)}{\int_x^{+\infty} f(x) dx} = \frac{f(x)}{S(x)}$$

e viceversa se è data $h(x)$ è univocamente determinata $f(x)$ perché:

$$\text{per def. } h(x) = \frac{f(x)}{S(x)} = \frac{-S'(x)}{S(x)} \text{ perché } f(x) = F'(x) \quad \left| \int \right| \Rightarrow f(x) = -S'(x) \quad \Downarrow$$

$$\frac{S'(x)}{S(x)} = -h(x) \Leftrightarrow \ln[S(x)] = -\int_0^x h(x) dx + c \Leftrightarrow$$

$$\Leftrightarrow S(x) = e^{-\int_0^x h(x) dx + c} = e^c \cdot e^{-\int_0^x h(x) dx} = c_1 \cdot e^{-\int_0^x h(x) dx}$$

$$\text{Poiché } x \geq 0 \Rightarrow S(0) = 1 \Rightarrow c_1 = 1 \Rightarrow S(x) = e^{-\int_0^x h(x) dx}$$

$$\left[f(x) = -S'(x) = h(x) \cdot e^{-\int_0^x h(x) dx} \right]$$

La funzione di rischio della funzione esponenziale è costante, poiché vale la proprietà di assenza di memoria. Infatti: se $X \sim \text{EXP}(1) \Rightarrow$

$$\Rightarrow h(x) = \frac{f(x)}{S(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda \Rightarrow h(x) = \lambda \Rightarrow \text{è costante.}$$

Proprietà della funzione di rischio:

$$1) h(x) = \frac{f(x)}{S(x)} \geq 0 \quad (f(x) \text{ è una densità, } S(x) = P(X \geq x) > 0)$$

$$2) \int_0^{+\infty} h(x) dx = +\infty \Leftrightarrow \lim_{x \rightarrow +\infty} \int_0^x h(x) dx = +\infty$$

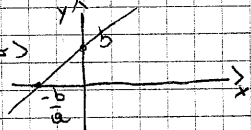
$$\text{Infatti: } \lim_{x \rightarrow +\infty} S(x) = \lim_{x \rightarrow +\infty} e^{-\int_0^x h(x) dx} = 0 \quad \text{ma } S(x) = e^{-\int_0^x h(x) dx} \Rightarrow$$

$$\Rightarrow \lim_{x \rightarrow +\infty} e^{-\int_0^x h(x) dx} = 0 \Leftrightarrow \lim_{x \rightarrow +\infty} \int_0^x h(x) dx = +\infty$$

DSS

Sia $h(x) = ax + b$ una funzione lineare. Può essere una funzione di rischio?

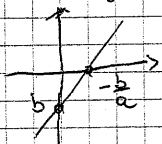
1) Caso $a > 0, b > 0 \Rightarrow$



Supporto $X \geq 0 \Rightarrow$

Se $x > 0 \Rightarrow h(x) > 0 \Rightarrow$ il primo integrale diverge $\Rightarrow h(x)$ può essere una funzione di rischio. Questo è l'unico caso di funzione lineare che è di rischio. Infatti, si ha:

2) caso $a \geq 0, b \leq 0$



non è una funzione di rischio perché $h(x) > 0$ non è sempre vero

3) caso $a < 0, b > 0$ si ha che se $x > -\frac{b}{a} \Rightarrow h(x) < 0 \Rightarrow$ non è una funzione di rischio

4) caso $a < 0, b \leq 0 \Rightarrow \forall x > 0, h(x) < 0 \Rightarrow$ non è una funzione di rischio

DISTRIBUZIONE MULTINOMIALE

Consideriamo n numeri aleatori $X_1, \dots, X_n$ e ciascuno di essi ha m possibili valori $\Rightarrow$

$$X_1 \in \{x_{1,1}, x_{1,2}, \dots, x_{1,m}\}$$

$$X_2 \in \{x_{2,1}, x_{2,2}, \dots, x_{2,m}\}$$

$$\vdots$$

$$X_n \in \{x_{n,1}, \dots, x_{n,m}\}$$

X_{ij} = evento i alla prova j

Supponiamo che siano eventi dello stesso tipo che riguardano ~~prove~~ diverse (cioè il numero aleatorio X_i si riferisce alla prova X_i) $\Rightarrow x_{1,1}, x_{1,2}, \dots, x_{n,1}$ indicano lo stesso evento prova per prova

Supponiamo inoltre che si ottiene X_{ij} verificarsi di E_{ij}

Ad esempio si lancia un dado e l'esito E_{ij} = "all' i-esimo lancio di un dado esce la faccia j".

Osserviamo che $P(X_i = X_{ij}) = p_j \quad \forall j = 1, \dots, m$ e non dipende da i , perché la probabilità che esce i è sempre la stessa sia al primo lancio o al secondo.

$\Rightarrow$ il risultato della i -esima prova è indipendente dal risultato della j -esima prova

$\Rightarrow X_1, \dots, X_n$ sono indipendenti

$\Rightarrow$ gli esiti E_{ij} e E_{ik} sono indipendenti se $i \neq k$ cioè le prove sono diverse.

Consideriamo ora i numeri aleatori $Y_1, Y_2, \dots, Y_m$ dove

$$Y_1 = \sum_{i=1}^n |X_i = X_{1i}| = \text{quante volte } X_{1i} \text{ è ottenuto il risultato di Tipo 1.}$$

$$Y_m = \sum_{i=1}^n |X_i = X_{mi}| = \text{quante volte in } n \text{ prove si è ottenuto il risultato di Tipo } m.$$

Ci chiediamo:

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_m = y_m) \text{ dove ovviamente } y_1 + y_2 + \dots + y_m = n$$

$$= \frac{n!}{y_1! y_2! \dots y_m!} \cdot p_1^{y_1} \cdot p_2^{y_2} \dots p_m^{y_m}$$

Coefficiente multinomiale = numero di modi in cui si possono disporre m oggetti in m scatole in modo che la scatola m_i ne contenga y_i e assai via.

Ora che $Y_1 \in \{0, 1, \dots, n\}$, $Y_2 \in \{0, 1, \dots, n\}$, $Y_m \in \{0, 1, \dots, n\}$,

Vogliamo ora calcolare $P(Y_1 = h)$, $h \in \{0, 1, \dots, n\}$

$$P(Y_1 = h) = \sum_{\substack{y_2, \dots, y_m: \\ y_2 + \dots + y_m = n-h}} \frac{n!}{h! y_2! \dots y_m!} \cdot p_1^h \cdot p_2^{y_2} \dots p_m^{y_m} \quad \left| \text{Questa è la distribuzione marginale} \right.$$

$$= \frac{n!}{h! \cdot 0!} \sum_{\substack{y_2, \dots, y_m: \\ y_2 + \dots + y_m = n-h}} \frac{1}{y_2! \dots y_m!} p_2^{y_2} \dots p_m^{y_m} =$$

$$= \frac{n!}{h! (n-h)!} p_1^h \sum_{\substack{y_2, \dots, y_m: \\ y_2 + \dots + y_m = n-h}} \frac{(n-h)!}{y_2! \dots y_m!} p_2^{y_2} \dots p_m^{y_m} = \quad \text{Sviluppo di un polinomio}$$

$$= \binom{n}{h} p_1^h \cdot (p_2 + \dots + p_m)^{n-h} = \binom{n}{h} p_1^h (1-p_1)^{n-h}$$

$$\Rightarrow Y_1 \sim B(n, p_1) \text{ e analogamente } Y_2 \sim B(n, p_2)$$

$$Y_m \sim B(n, p_m)$$

TEOREMA DEL LIMITE CENTRALE

Sia data una successione di numeri aleatori $X_1, \dots, X_n$ con esiti la stessa distribuzione ed indipendenti e siano $\mu = E(X_i)$, $\sigma^2 = \text{VAR}(X_i)$ ed

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} \quad \text{e allora} \quad P(S_n) = \frac{1}{n} \cdot n \cdot \mu = \mu \quad \text{e}$$

$$\text{medie aritmetica dei primi } n \text{ di } X_i$$

$$\text{VAR}(S_n) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}, \quad \text{cov} = 0, \quad \text{allora} \quad \downarrow$$

A questo punto consideriamo

$$Y_n = \frac{S_n - n}{\sqrt{n}} = \text{numeri aleatori ridotto o standardizzato corrispondente}$$

(42)

a S_n per i quali si ha: $P(Y_n) = 0$ e $VAR(Y_n) = 1$

Sotto queste ipotesi si dimostra che per $n \rightarrow +\infty$ la distribuzione di Y_n

Tende ad una distribuzione normale di punti 0, 1

↓ Non bisogna mai scrivere $Y_n \sim N_{0,1}$

OSS. Ci sono altre versioni di questo Teorema con ipotesi più forti o complicate

#

* = good

6, 7, 8 De Finetti

11

13

18

20 Scambiabilità

22 Exp erenti condizionati, Gioco equo

23 Exp $p(E_i)$ 2 palline bianche, $N-2$ nere

23-24 Hypergeometric *

25-26 come 23, ma con restituzione, 26 X ass continuo, 30 Distrib normale

30 Assenza Memoria *, Vettori continui

~~37~~ 37 Distrib. Pascal *

38 Distrib. Poisson *

43 Assenza di Memoria, per Exp(2) *

44 Funzione di rischio

46 Distrib binom

42 Bayes verosimiglianze

45 Dist multinomiale

Esercizi

pg 28

pg 32

belle di contingenza a due entrate.

Supponiamo di classificare n individui o oggetti secondo due

criteri A e B , che esistano r classificazioni $A_1, A_2, \dots, A_r$ in A

ed s classificazioni $B_1, B_2, \dots, B_s$ in B e che il numero di individui appartenenti ad A_i e B_j è

N_{ij} . Abbiamo una Tabella di contingenza $r \times s$ con frequenza di cella N_{ij} .

	B_1	B_2
A_1	N_{11}	
A_r		N_{rs}

Vogliamo verificare l'ipotesi H_0 che le classificazioni A e B sono indipendenti cioè:

$$\text{che } p(B_j | A_i) = p(B_j) \text{ e } p(A_i | B_j) = p(A_i) \Leftrightarrow p(A_i, B_j) = p(A_i) p(B_j)$$

Se indichiamo con $p_i = p(A_i)$, $i=1, \dots, r$ e $p_j = p(B_j)$, $j=1, \dots, s$

$$\Rightarrow H_0: p_{ij} = p_i p_j$$

La verosimiglianza è $L = \prod_{i,j} p_{ij}^{n_{ij}}$, ^{numero di individui in ij} e ha il suo massimo in $\hat{p}_{ij} = \frac{n_{ij}}{n}$

--- sul libro pag 455.

Cioè per applicare il Test del chi-quadrato dobbiamo parametrizzare l'ipotesi:

Nel caso $r=2$ e $s=2 \Rightarrow \Theta = \{(\theta_1, \theta_2, \theta_3) = (p_{11}, p_{12}, p_{21}) : p_{11} \geq 0, p_{12} \geq 0, p_{21} \geq 0$

$p_{11} + p_{12} + p_{21} \leq 1\}$. Parametizziamo così:

$$\Theta = \{(\theta_1', \theta_2', \theta_3') : \theta_1' = p_{11} - p_1 p_1, \theta_2' = p_{12} \text{ e } \theta_3' = p_{21}\}$$

L'ipotesi diventa $H_0: \theta_1' = 0$, θ_2' e θ_3' non specificati $\Rightarrow H_0$ è nella forma richiesta dal Teorema.