Contents

Matematica Applicata

Giuseppe Bumma

September 2022

1 Introduzione

Nel XVII secolo scienziati come Fermat, Pascal e Huygens introducono il concetto di probabilità, in maniera molto rozza, con tecniche di calcolo semplici.

Nel XVIII secolo con Laplace, Gauss, de Maivve e Poisson si introduce la variabile causale, la legge dei grandi numeri, teorema del limite centrale, distribuzione normale ecc... Questi concetti però non furono definiti in maniera dettagliata come lo sono oggi.

Nel XX secolo con Kolmogorov (nel 1933) si introduce la teoria della probabilità rigorosa dal punto di vista matematico.

2 Probabilità

2.1 Definizione di Probabilità

2.1.1 Definizione classica

Dato un esperimento con un numero finitio di possibili esiti **equiprobabili**, un evento A associato a questo esperimento ha probabilità

$$P(A) = \frac{n^o \text{ esiti favorevoli}}{n^o \text{ esiti totali}}$$

L'insieme di tutti gli esiti dell'esperimento viene chiamato *spazio campione*, si identifica con il simbolo S o Ω e si rappresenta con un rettangolo.

Valutiamo questa definizione:

- Pro
  - semplicità
- Contro
  - richiesta di esiti equiprobabili (è illogico definire la probabilità usando la definizione di probabilità)
  - richiesta di esiti in numero finito
  - a volte non è possibile contare gli esiti

2.1.2 Definizione frequentista

Si ripete N volte l'esperimento in maniera identica e indipendente, la probabilità dell'evento A si definisce

$$P(A) = \frac{n^o \text{esperimenti con esito A}}{N(n^o \text{ totale di esperimenti})}$$

Valutiamo questa definizione:

- Pro
  - vale per eventi non equiprobabili
  - non è richiesto di contare gli eventi
- Contro
  - gli esiti dell'esperimento devono essere in nuemero finito
  - a volte non è possibile ripetere l'esperimento (molte volte)

2.2 Calcolo combinatorio

2.2.1 Principio fondamentale del calcolo combinatorio o principio di enumerazione

Si supponga di realizzare due esperimenti e si supponga che il primo esperimento presenti n possibili esiti, mentre il secondo esperimento presenti m possibili esiti. Le coppie ordinate che contengono gli esiti del primo e del secondo esperimento saranno $n \times m$.

Esempio

Scatola con 6 palline rosse e 5 palline verdi. Estrazione con reimissione di 3 palline. Qual è la probabilità di ottenere 2 palline rosse e una verde (senza badare all'ordine di estrazione)?

$$A = 2R + 1V \text{ senza ordine}$$

n^o di esiti possibili $= 11 \times 11 \times 11 = 11^3$
n^o di esiti favorevoli $= \underbrace{6 \times 5 \times 6}_{\text{RVR}} + \underbrace{6 \times 6 \times 5}_{\text{RRV}} + \underbrace{5 \times 6 \times 6}_{\text{VRR}} = 3 \times 5 \times 6^2$

$$P(A) = \frac{15 \times 6^2}{11^3}$$

Esmepio di prima senza reimissione

$$B = 2R + 1V \text{ senza ordine e senza reimissioni}$$

n^o di esiti possibili $= 11 \times 10 \times 19$
n^o di esiti favorevoli $= 5 \times 6 \times 5 + 6 \times 5 \times 5 + 6 \times 5 \times 5 = 18 \times 5^2$

$$P(B) = \frac{n^o \text{ esiti favorevoli}}{n^o \text{ esiti possibili}} = \frac{18 \times 5^2}{11 \times 10 \times 9} = \frac{5}{11}$$

2.2.2 Disposizioni semplici di n elementi di classe k

gli n a disposizione senza ripetizioni.

$$D_{n,k} = n \times (n-1) \times (n-2) \times ... \times (n-k+1) = \frac{n!}{(n-k)!}$$

N.B. Le disposizioni semplici sono associate alle estrazioni senza reimissioni.

2.2.3 Disposizioni con ripetizione di n elementi di classe k

Dati n elementi distinti, si dicono disposizioni con ripetizione di n elementi di classe k tutti gli allineamenti di k elementi (anche ripetuti) presi dall'insieme degli elementi dati.

$$D^R_{n,k} = n^k$$

N.B. k può essere qualsiasi. N.B.2 Le disposizioni con ripetizione sono associate all'estrazione con reimissione.

2.2.4 Permutazioni semplici di n elementi

Dati n elementi distinti, s dicono permutazioni semplici degli n elementi tutti gli allineamenti degli n elementi.

$$P_n = D_{n,n} = n!$$

2.2.5 Permutazioni con ripetizione di n elementi

Dati n elementi <u>non necessariamente distinti</u>, si dicono permutazioni con ripetizione degli n elementi tutti gli allineamenti di questi n elementi. $n = k_1 + k_2 + k_3 + ... + k_j$ $k_n = n^o$ di elementi non distinti di tipo h

$$P^R_n = k_1 + k_2 + ... + k_j = \frac{n!}{k_1!k_2!...k_j!}$$

Anagrammi della parola ”MATEMATICA”

Abbiamo $n = 10$ tra cui 3A, 2M, 2T, 1E, 1I, 1C.

$$P_n^R = \frac{n!}{k_1!k_2!...k_j!} = \frac{10!}{3!2!2!1!1!1!} = \frac{10!}{24}$$

2.2.6 Combinazioni semplici di n elementi semplici di classe k, $(k \leq n)$

Dati n elementi distinti, si dicono combinazioni semplici di n elementi di classe k $(k \leq n)$ tutti i gruppi di k elementi che si possono formare senza ripetizioni a partire da n dati.

$$C_{n,k} = \frac{D_{n,k}}{P_k} = \binom{n}{k}$$

Combinazioni semplici associate all'estrazione simultanea.

Esempio del paradosso dei compleanni

Immaginiamo di avere n persone nate in anni non bisestili in una stanza $(1 \leq n \leq 365)$. Qual è la probabilità che tutte le persone abbiano date di compleanno diverse? (Evento C).
Ipotesi: tutte le date sono equiprobabili.
$C =$ ”tutte le persone hanno date di compleanno diverse”.

Possiamo comportarci come se avessimo un caso di estrazione con reimissione; gli esiti totali saranno quindi 365^n oppure $D^R_{365,n}$.

n^o esiti favorevoli $= 365 \times 364 \times 363 \times ... \times (365 - n + 1)$ oppure $D_{365,n} = \dfrac{365!}{(365-n)!}$

$$P(C) = \frac{365 \times 364 \times 363 \times ... \times (365 - n + 1)}{365^n}$$

Qual è invece la probabilità che non tutte le persone abbiano date di compleanno diverse?
Definiamo E^C come l' **evento complementare** di un evento E, cioè l'insieme di tutti gli esiti di S che non stanno in E, definito come $P(E^C) = 1 - P(C)$

$$P(C^C) = 1 - P(C) = 1 - \frac{365 \times 364 \times 363 \times ... \times (365 - n + 1)}{365^n}$$

Facendo dei calcoli vengono i seguenti risultati:

- se $n = 20 \Longrightarrow P(C^C) = 41\%$
- se $n = 40 \Longrightarrow P(C^C) = 89\%$
- se $n = 60 \Longrightarrow P(C^C) = 99,4\%$

Già con 60 persone (molto meno rispetto al massimo di 365) si ha una probabilità del quasi 100%, per questo si verifica un paradosso.

2.3 Problemi di astrazioni

Immaginiamo di avere un box con 5 palline verdi (V) e 3 palline blu (B).

1. estrazione con reimissione, 3 estrazioni di una pallina per volta.
   $E_1 = 2V + 1B$ (senza ordine)

$$\text{Esiti totali} = 8 \times 7 \times 6 = D_{8,3} = \frac{8!}{5!}$$

$$\text{Esiti favorevoli} = \text{VVB+VBV+BBV} = (5 \times 4 \times 4) + (5 \times 3 \times 4) + (3 \times 5 \times 4) = 180 = 3^2 \times 4 \times 5$$

$$P(E_1) = \frac{3^2 \times 4 \times 5}{8 \times 7 \times 6} = \frac{15}{28}$$

2. estrazione simultanea di 3 palline.
   $E_2 = 2V + 1B$ (senza ordine)

$$\text{Esiti totali} = C_{8,3} = \binom{8}{3} = \frac{8!}{5! \times 3!} = \frac{8 \times 7 \times 6}{3!} = 56$$

$$\text{Esiti favorevoli} = C_{5,2} \times C_{3,1} = \binom{5}{2} \times \binom{3}{1} = \frac{5!}{3! \times 2!} \times \frac{3!}{2! \times 1!} = 30$$

$$P(E_2) = \frac{30}{56} = \frac{15}{28}$$

   N.B. Questi due casi danno lo stesso risultato perché se non considero l'ordine è indifferente usare un'estrazione simultanea (caso 2) o di una pallina per volta (caso 1).

3. estrazione con reimissione di 3 palline, una alla volta.
   $E_3 = 2V + 1B$ (senza ordine)

$$\text{Esiti totali} = D^R_{8,3} = 8^3$$

$$\text{Esiti favorevoli} = (5 \times 5 \times 3) + (5 \times 3 \times 5) + (3 \times 5 \times 5) = 3^2 \times 5^2$$

$$P(E_3) = \frac{3^2 \times 5^2}{8^3}$$

2.3.1 Notazioni e definizioni preliminari

- **Unione di eventi:** $A, B \subset S, \quad A \cup B =$ evento di S che contiene tuti gli esiti di A e/o B.
- **Intersezione:** $A, B \subset S, \quad A \cap B =$ evento di S che contiene tutti gli esiti sia in A che in B.
- **Insieme vuoto:** $\emptyset =$ evento che non contiene esiti.
- **Eventi disgiunti/mutuamente esclusivi:** $A, B \subset S, \quad A \cap B = \emptyset$
- **Evento complementare:** $A \subset S, \quad A^C = S - A$
  N.B. $S^C = \emptyset$

$$A, B \subset S, \quad A^C \cap B^C = (A \cup B)^C \qquad A^C \cup B^C = (A \cap B)^C$$

$$A_1, A_2, ..., A_m \subset S, \quad \bigcap_{k=1}^{m} A_k = A_1 \cap A_2 \cap ... \cap A_m \qquad \bigcup_{k=1}^{m} A_k = A_1 \cup A_2 \cup ... \cup A_m$$

2.4 Assiomi di Kolmogorov

2.4.1 Definizione assiomatica di probabilità

Dato u esperimento che prevede più esiti possibili e a cui è associato uno spazio campione S, e dato un evento $E \subset S$, si definisce probabilità di E $P(E) \in \mathbb{R}$, il numero tale che

$A_1)$ $0 \leq P(E) \leq 1$

$A_2)$ P(S) $= 1$

$A_3)$ Dati $E_1, E_2, ..., E_m \subset S$ mutuamente esclusivi, allora

$$P\Big(\bigcup_{k=1}^{m} E_k\Big) = \sum_{k=1}^{m} P(E_k)$$

Proprietà

$P_1)$ $E \subset S$

$$P(E^C) = 1 - P(E)$$

$P_{1bis})$

$$P(\emptyset) = 0$$

$P_2)$ $A, B \subset S$

$$A \subset B \Rightarrow P(A) \leq P(B)$$

$P_3)$ $A, B \subset S$

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

$P_4)$ $A, B, C \subset S$

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

2.5 Definizione classica di probabilità (come conseguenza degli assiomi di Kolmogorov) o spazio degli esiti equiprobabili

Dato un esperimento che presenta un numero finito N di esiti equiprobabili $S = \{e_1, e_2, ..., e_N\}$ con $N \in \mathbb{N}$ e $N < +\infty$

$$P(e_1) = P(e_2) = ... = P(e_N) = P, \quad 0 < P < 1$$

N.B. non ha senso porre la condizione $0 \leq P \leq 1$

$e_1, e_2, ..., e_N$ sono disgiunti in quanto esiti

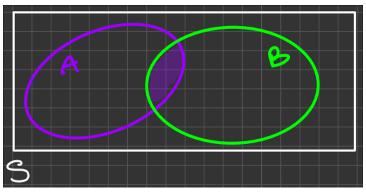
$S = \bigcup_{k=1}^{N} e_k$

$$1 \underset{A_2}{=} P(S) = P\Big(\bigcup_{k=1}^{N} e_k\Big) \underset{A_3}{=} \sum_{k=1}^{N} P(e_k) \underbrace{=}_{\text{equiprob}} \sum_{k=1}^{N} N \cdot P \Longrightarrow P = \frac{1}{N}$$

2.6 Probabilità condizionata

Dati due eventi $A, B \subset S$ con $P(B) \neq 0$ si definisce probabilità di A condizionata da B

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



N.B. Nella probabilità condizionata il mio spazio campione si restringe da S a B.

2.6.1 Eventi indipendenti

Dati due eventi $A, B \subset S$ essi si dicono indipendenti se

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

Esempio

Immaginiamo di avere 52 carte da poker (un mazzo intero).
A = 'estrazione di un asso' (4 assi per mazzo);
B = 'estrazione di una carta di quadri';
A e B sono indipendenti?

$$P(A) = \frac{4}{52} = \frac{1}{13} \qquad\qquad P(B) = \frac{13}{52} = \frac{1}{4}$$

$P(A|B) = P(\text{estrarre asso di quadri}) = \dfrac{1}{52}$ (c'è un solo asso di quadri per mazzo).
Verifichiamo se $P(A|B) = P(A) \cdot P(B)$

$$P(A) \cdot P(B) = \frac{1}{13} \cdot \frac{1}{4} = \frac{1}{52} \Longrightarrow A, B \text{ sono indipendenti}$$

Osservazione

Dati $A, B \subset S$ indipendenti con $P(B) \neq 0$

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

Se A e B sono indipendenti $P(A|B) = P(A)$, ovvero B non condiziona A.

Teorema

Dati 2 eventi $A, B \subset S$ indipendenti, allora anche Ae B^C sono indipendenti tra loro, e lo stesso vale per A^Ce B e A^C e B^C.

2.6.2 Indipendenza tra 3 eventi

Dati $A, B, C \subset S$, essi si dicono indipendenti se

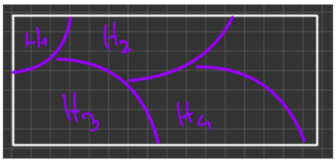
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

$$P(A \cap B) = P(A) \cdot P(B) \qquad P(A \cap C) = P(A) \cdot P(C) \qquad P(B \cap C) = P(B) \cdot P(C)$$

In generale N eventi sono indipendenti se tutte le loro intersezioni tra $N, N-1, N-2, ..., 2$ soddisfano condizioni simili alle precedenti.

2.6.3 Partizione di uno spazio campione

Dato uno spazio campione S, si dice partizione di S una suddivisione in eventi $\{H_1, H_2, ..., H_m\}(m \in \mathbb{N}/\{0, 1\})$ tali che $H_i \cap H_k = \emptyset$ se $i \neq k$ e $\bigcup_{k-1}^{m} H_k = S$



Gli elementi della partizione $H_1, H_2, ..., H_m$ sono detti **ipotesi**.

2.7 Formula (o teorema) delle probabilità totali

Dato uno spazio campione$S$ e una sua partizione $\{H_1, H_2, ..., H_m\}$ e dato un evento $A \subset S$

$$P(A) = \sum_{k=1}^{m} P(A|H_k) \cdot P(H_k)$$

2.8 Teorema di Bayes (o teorema della probabilità posteriore)

Dato uno spazio campione S e una sua partizione $\{H_1, H_2, ..., H_m\}$ e considerato un evento $E \subset S$ con $P(E) \neq \emptyset$

$$P(H_j|E) = \frac{P(E|H_j) \cdot P(H_j)}{\sum_{k=1}^{m} P(E|H_k) \cdot P(H_k)}$$

3 Variabili casuali discrete

Dato un esperimento che può avere esiti diversi, si associa a ciascun esito il valore di una variabile detta **casuale o aleatoria**.

Esempio

Nel caso del lancio di una moneta ho che il mio spazio $S = \{T, C\}$. Posso associare a questo esperimento valori diversi per testa e croce, quindi variabili casuali diverse, ad esempio $X \in \{-1, 1\}$.

Invece nel caso del lancio di un dado cubico possiamo indicare come insieme delle possibili facce del dado $X \in \{1, 2, 3, 4, 5, 6\}$

Ogni valore della variabile casuale è associato a un esito e quindi anche alla probabilità di verificarsi dell'esito.
Se $X \in \{q_1, q_2, ..., q_m$ con $m \in \mathbb{N}/\{0\}$

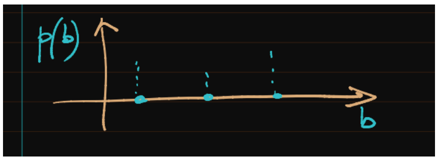
$$P(X = P_k) = P(e_k)$$

con e_k esito k-simo associato al valore q_k.

3.1 Funzione di massa di probabilità

Se X è una variabile casuale discreta, ovvero assume valori discreti: $X \in \{a_1, .., a_m\}$, si può definire la **funzione di massa di probabilità**

$$\varphi(y) = \Phi(x = y) = \begin{cases} \varphi(e_k) \text{ se } y = a_k \quad k = 1, ..., m \\ 0 \text{ se } y \neq a_k \quad k = 1, ..., m \end{cases}$$



Proprietà

1. $0 \leq \varphi(b) \leq 1 \quad \forall b \in \mathbb{R}$
2. $\sum_{k=1}^{n} \varphi(a_k) = P(S) = 1$

3.2 Funzione di ripartizione o distribuzione di probabilità

$F : \mathbb{R} \longrightarrow [0, 1]$

$$F(b) = P(X \leq b) \quad \forall b \in \mathbb{R}$$

Proprietà

1. $0 \leq F(b) \leq 1$
2. $\lim_{b \to +\infty} F(b) = P(X \leq +\infty = 1)$
3. $\lim_{b \to -\infty} F(b) = P(X \leq -\infty) = 0$

4. $b, c \in \mathbb{R}$ tali che $b < c$

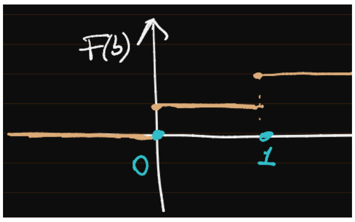
$$F(b) = P(X \leq b) \underbrace{\leq}_{p_2} P(X \leq c)$$
$$= F(c) \implies F(b) \leq F(c) \Leftarrow b < c$$

N.B. $(X \leq b) \subset (X \leq c) \implies P(X \leq b) \leq P(X \leq c)$
F è una funzione decrescente

5. Se X è una variabile casuale discreta F è una funzione 'a gradini'.
Esempio 1
lancio di una moneta equilibrata
$S = \{T, C\}, X = \{0.1\} \begin{cases} X = 0 \text{ se esce } T \\ X = 1 \text{ se esce } C \end{cases}$

$$\varphi(b) = P(X = b) = \begin{cases} \frac{1}{2} & \text{se } b = 0 \text{ oppure } b = 1 \\ 0 & \text{altrove} \end{cases}$$

$$F(b) = P(X \leq b) = \begin{cases} 0 & \text{se } b < 0 \\ \frac{1}{2} & \text{se } b = 0 \\ = P(X = 0) & \text{se } 0 < b < 1 \\ 1 & \text{se } b = 1 \\ 1 & \text{se } b > 1 \end{cases}$$



Esempio 2
$X \in \{-1, 0, 1\}$
$p(-1) = \frac{1}{6}, p(0) = \frac{1}{3}, p(1) = \frac{1}{2}$

$$F(b) = P(X \leq b) = \begin{cases} 0 & b < -1 \\ P(X = -1) = \frac{1}{6} & b = -1 \\ P(X = -1) = \frac{1}{6} & -1 < b < 0 \\ P(X = -1 \cup X + 0) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2} & b = 0 \\ P(X = -1 \cup X = 0) = \frac{1}{2} & 0 < b < 1 \\ P(X = -1 \cup X = 0 \cup X = 1) = 1 & b = 1 \\ 1 & b > 1 \end{cases}$$

4 Variabili casuali continue

Una variabile casuale si dice *continua* se ad essa è associata una funzione (detta di densità di probabilità) $f : \mathbb{R} \longrightarrow \mathbb{R}^+ \cup \{0\}$ tale che $\forall \mathcal{B} \in \mathbb{R}$

$$P(X \in \mathcal{B}) = \int_{\mathcal{B}} f(s)ds$$

Proprietà

1. $f(\varkappa) \geq 0 \qquad \forall \varkappa \in \mathbb{R}$
2. $$\underbrace{P(X \in \mathbb{R}}_{\int_{\mathbb{R}} f(s)ds} = P(X \leq +\infty) = 1 \Longrightarrow \int_{-\infty}^{+\infty} f(s)ds = 1$$

4.1 Funzione di ripartizione o distribuzione di probabilità

$$F : \mathbb{R} \longrightarrow [0,1] \text{ tale che } F(b) = P(X \leq b) \qquad \forall b\ in \mathbb{R}$$

Proprietà

1. identica al caso delle variabili casuali discrete
2. identica al caso delle variabili casuali discrete
3. identica al caso delle variabili casuali discrete
4. identica al caso delle variabili casuali discrete
5. $$F(b) = P(X \leq b) = \int_{-\infty}^{b} f(s)ds \Longrightarrow f(b) = \frac{dF(b)}{db}$$

   N.B. F è una funzione continua, quindi non può essere a '2 gradini'. Inoltre $f(b) \geq 0$ perché $F(b)$ è non decrescente.

Domanda 1: perché non si può associare la funzione di densità di probabilità alle variabili casuali discrete? Nel caso di variabili casuali discrete la funzione di ripartizione è discontinua (a 2 gradini) e quindi non derivabile.

Domanda 2: È possibile definire la funzione di massa di probabilità per una variabile casuale continua?

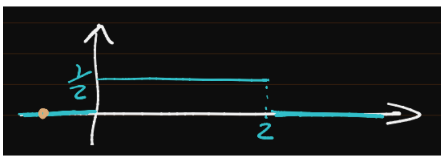
$$\varphi(b) = P(X = b) \underbrace{=}_{\text{se } X \text{ v.c. continua}} \int_{X=b} f(s)ds = \int_{b}^{b} f(s)ds = 0 \qquad \forall b \in \mathbb{R}$$

Quindi nel caso continuo non è significativa.
Nel caso di variabile casuale continua non è significativo domandarsi quanto vale P(X=q). È invece naturale domandarsi qual è la probabilità che X appartenga a un dato intervallo.

Esempio

$X \in [0,2]$ variabile casuale uniforme.

$$f(b) = \begin{cases} \frac{1}{2} & 0 \leq b \leq 2 \\ \\ 0 & \text{altrove} \end{cases}$$



$$P(X=1)=0, \quad P(0.99 \leq X \leq 1.01)=\int_{0.99}^{1.01} \frac{1}{2} ds=\frac{1}{2}(1.01-0.99)=10^{-2}$$

$$F(b)=P(X \leq b)=\begin{cases} \displaystyle\int_{-\infty}^{b} 0 \ ds=0 & b<0 \\ \displaystyle\int_{-\infty}^{0} 0 \ ds+\int_{0}^{b} \frac{1}{2} \ ds=\frac{b}{2} & 0 \leq b \leq 2 \\ \displaystyle\int_{-\infty}^{0} 0 \ ds+\int_{0}^{2} \frac{1}{2} \ ds \int_{2}^{b} 0 \ ds=1 & b>2 \end{cases}$$

N.B. F(funzione di ripartizione di probabilità) è utile nel caso di X variabile casuale continua per velocizzare il calcolo delle probabilità sugli intervalli

$$P(a<X \leq b)=F(b)-F(a)$$

5 Coppie di variabili casuali o variabili casuali doppie discrete

$$(X, Y) \rightarrow \begin{cases} X \in \{x_1, x_2, \ldots, x_n\} \\ Y \in \{y_1, y_2, \ldots, y_m\} \end{cases} \quad n, m \in \mathbb{N} / \{0,1\}$$

5.1 Funzione di massa di probabilità congiunta

$\varphi: \mathbb{R}^2 \longrightarrow [0,1]$

$$\varphi(a, b)=P(X=a \cap Y=b)=P(X=a, Y=b)$$

Proprietà

1. $\varphi(a, b) \in [0,1]$
2. $\sum\limits_{k=1}^{n} \sum\limits_{j=1}^{m} \varphi(x_k, y_j)=1$

5.2 Funzioni di massa marginali

$$\rho_X(a)=P(X=a)=P(X=a, Y \text{ qualsiasi})=\sum_{j=1}^{m} P(X=a, Y=y_j)$$

$$\rho_Y(b)=P(Y=b)=P(X \text{ qualsiasi}, Y=b)=\sum_{k=1}^{n} P(X=x_k, Y=b)$$

5.3 Variabili casuali indipendenti

Due variabili casuali X e Y si dicono *indipendenti* se $\forall \mathcal{A}, \mathcal{B} \subset \mathbb{R}$

$$P(X \in \mathcal{A}, Y \in \mathcal{B})=P(X \in \mathcal{A}) \cdot P(Y \in \mathcal{B})$$

N.B. La definizione è semplice ma non operativa, ovvero non è possibile verificarla $\forall \mathcal{A}, \mathcal{B} \subset \mathbb{R}$ (bisognerebbe definire ogni sottoinsieme dei numeri reali).

5.3.1 Riconoscere due variabili indipendenti

Per riconoscere facilmente l'indipendenza di una coppia di variabili continue di utilizzano i 3 teoremi seguenti:

Teorema 1

Condizione necessaria e sufficiente affinché due variabili casuali siano indipendenti è che

$$F(a, b)=F_x(a) \cdot F_Y(b) \qquad \forall a, b \in \mathbb{R}$$

Teorema 2

Condizione necessaria e sufficiente affinché 2 variabili casuali discrete siano indipendenti è che la loro funzione di massa di probabilità

$$p(a,b) = p_X(a) \cdot p_Y(b) \qquad \forall a,b \in \mathbb{R}$$

Teorema 3

Condizione necessaria e sufficiente affinché due variabili congiuntamente continue siano indipendenti è che la loro funzione di densità di probabilità

$$f(a,b) = f_X(a) \cdot f_Y(b) \qquad \forall a,b \in \mathbb{R}$$

Esempio

Y/X	0	1	2
0	1/4	1/4	0
1	0	1/4	1/4

X e Y sono indipendenti? Calcolo le funzioni di massa marginali sommando i termini sulle righe per p_Y, i termini sulle colonne per p_X.

Y/X	0	1	2	
0	1/4	1/4	0	$p_Y(0) = 1/2$
1	0	1/4	1/4	$p_Y(1) = 1/2$
	$p_X(0) = 1/4$	$p_X(1) = 1/2$	$p_X(2) = 1/4$	

Dal Teorema 2 X e Y sono indipendenti se e solo se

$$p(a,b) = p_X(a) \cdot p_Y(b)$$

$$p(0,1) \stackrel{?}{=} p_X(0) \cdot P_Y(1) \Longrightarrow 0 = p(0,1) \stackrel{?}{=} p_X(0) \cdot P_Y(1) = \frac{1}{4} \cdot \frac{1}{2} \Longrightarrow 0 \neq \frac{1}{8}$$

X e Y sono dipendenti.

N.B. In generale bisogna controllare per ogni valore.

6 Introduzione ai modelli di variabili casuali discrete

6.1 Variabile casuale di Bernoulli

$$X \sim B_e$$

Questa scrittura si legge ”X si comporta come una variabile casuale di Bernoulli”.
p è il parametro della variabile casuale di Bernoulli, $0 < p < 1$.

Si compie un esperimento in cui l’evento A può verificarsi oppure no

$$X = \begin{cases} 1 & \text{se si verifica } A \rightarrow P(X=1) = P(A) \\ 0 & \text{altrimenti} \rightarrow P(X=0) = 1 - P(A) = 1 - p = q \end{cases}$$

6.2 Variabile casuale binomiale

$$X \sim B(n,p)$$

n e p sono i parametri della binomiale, dove $n \in \mathbb{N}/\{0\}$ e $0 < p < 1$.

Si ripete n volte l’esperimento in maniera identica e indipendente.
X rappresenta il numero di prove in cui si verifica A, $X \in \{0,1,2,...,n\}$ se $P(A) = p$, se k è il numero di volte in cui si verifica A nelle n prove $k = 0,1,...,n$ allora $n-k$ sarà il numero di volte in cui non si verifica A:

$$p(k) = P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Il coefficiente binomiale si usa per calcolare tutte le possibili combinazioni di k elementi sugli n totali.

Proprietà

$$\sum_{k=0}^{n} p(k) = 1$$

6.3 Variabile casuale geometrica

$$X \sim G(p) \qquad\qquad o < p < 1$$

Si ripete un esperimento in maniera identica e indipendente fino ad osservare A ($p = P(A)$ probabilità di verificarsi di A in una prova), X è il numero di tentativi necessari.

$X \in \{1, ..., +\infty\}, k \in \mathbb{N}/\{0\}$

$$\begin{aligned} p(k) &= P(X = k) \\ &= P(\text{'nelle prime } k-1 \text{ prove si verifica sempre } A^c \text{ e nella } k-sima \text{ prova si verifica A'}) \\ &= (1-p)^{k-1} \cdot p \end{aligned}$$

Proprietà

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

Esempio

Se ci sono n chiavi e 1 sola apre la porta, e se le chiavi già aperte vengono rimesse nel mazzo

$$A = \text{ 'scelta della chiave giusta'}$$

$P(A) = p = \dfrac{1}{n}$ e $X = n^o$ di tentativi per aprire la porta

$$P(X = k) = (1-p)^{k-1} \cdot p = \left(1 - \frac{1}{n}\right)^{k-1} \cdot \frac{1}{n}$$

N.B. Se le chiavi no vengono rimesse nel mazzo il numero di tentativi non è una geometrica.

7 Valor medio, varianza e covarianza

7.1 Valor medio

Data una variabile casuale X si definisce, se esiste, il suo valor medio la seguente quantità

$$E[X] = \begin{cases} \displaystyle\sum_{k=1}^{n} x_k \cdot p(x_k) & \text{se } X \text{ variabile discreta e } X \in \{x_1, x_2, ..., x_n\} \\ \\ \displaystyle\int_{-\infty}^{+\infty} x \cdot f(x)\, dx & \text{se } X \text{ variabile casuale continua} \end{cases}$$

Esempio

$X \sim B_e(p), E[X] = ?$

$$X \in \begin{cases} 1 & p(1) = p \\ 0 & p(0) = 1 - p = q \end{cases} \qquad E[X] = \sum_{k=1}^{2} x_k \cdot p(x_k) = 0 \cdot p(0) + 1 \cdot p(1) = 1 \cdot p = \underbrace{p}_{P(X=A)}$$

Proprietà

1. Se $Y = h(X)$ con X variabile casuale nota

$$E[Y] = E[h(X)] = \begin{cases} \displaystyle\sum_{k=1}^{n} h(x_k) \cdot p(x_k) & \text{se } X \text{ variabile discreta e } X \in \{x_1, x_2, ..., x_n\} \\ \\ \displaystyle\int_{-\infty}^{+\infty} h(x) \cdot f(x)\ dx & \text{se } X \text{ variabile casuale continua} \end{cases}$$

2. Caso particolare: $Y = \alpha x + \beta, \quad \alpha \in \mathbb{R}, \beta \in \mathbb{R}$

$$E[Y] = \alpha E[X] + \beta$$

3. Se $Z = g(X, Y)$ con X, Y variabili casuali note

$$E[Z] = E[g(X,Y)] = \begin{cases} \displaystyle\sum_{k=1}^{n}\sum_{j=1}^{n} g(x_k, y_j)p(x_k, y_j) & \text{se } X, Y \text{ variabili casuali discrete con} \\ & X \in \{x_1, ..., x_n\}, Y \in \{y_1, ..., y_n\} \\ \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\ dx\ dy & \text{se } (X, Y) \text{ congiuntamente continue} \end{cases}$$

4. Date X, Y variabili casuali

$$E[X + Y] = E[X] + E[Y]$$

4bis. Se $X_1, X_2, ..., X_N$ variabili casuali con valor medio definito

$$E[X_1 + X_2 + ... + X_N] = E[X_1] + ... + E[X_N]$$

"il valor medio della somma di più variabili casuali (con media definita) è uguale alla somma dei valori medi"

Esempio

Lancio di un dado equilibrato di forma cubica

$$X \in \{1, 2, 3, 4, 5, 6\} \qquad p(1) = p(2) = ... = p(6) = \frac{1}{6}$$

Siamo nel caso discreto, quindi calcolo il valore medio con la sommatoria:

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2} = 3,5$$

Osservazione

In generale $E[X]$ può assumere un valore no compreso tra i valori che esso può presentare ma deve sempre essere compreso tra il più piccolo e il più grande dei valori che esso può rappresentare.

7.2 Momento n-simo di una variabile casuale

Se esiste, il *momento n-simo* di X è

$$E[X^n] = \begin{cases} \displaystyle\sum_{k=1}^{n} x_k^n \cdot p(x_k) & X \text{ v.c. discreta} \\ \\ \displaystyle\int_{-\infty}^{+\infty} x^n \cdot f(x)\ dx & X \text{ v.c. continua} \end{cases}$$

Questa definizione può essere vista come un caso particolare della proprietà 1.

7.3 Covarianza di una variabile casuale

Date X e Y variabili continue con $E[X] = \mu_X$ e $E[Y] = \mu_Y$ si definisce *covarianza*

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Proprietà

Se X e Y variabile continue discrete

$$X \in \{x_1, x_2, ..., x_m\} \qquad\qquad Y \in \{y_1, y_2, ..., y_n\}$$

$$Cov(X,Y) = \sum_{k=1}^{m}\sum_{j=1}^{n}(x_k - \mu_X)(y_j - \mu_Y)p(x_k, y_j)$$

Se X e Y variabile continue continue

$$Cov(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - \mu_X)(y - \mu_Y)f(x,y)\ dxdy$$

Proprietà

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X,X) = E[(x - \mu_X)(x - \mu_X)] = E[(x - \mu_X)^2] = Var(X)$
3. $Cov(X,Y) = E[XY] - E[X]E[Y]$
4. $\forall \alpha \in \mathbb{R}$
$$Cov(\alpha X, Y) = Cov(X, \alpha Y) = \alpha Cov(X,Y)$$
5. $X_1, X_2, ..., X_N$ e $Y_1, Y_2, ..., Y_M$ variabili casuali
$$Cov\left(\sum_{k=1}^{N} X_k, \sum_{j=1}^{M} Y_j\right) = \sum_{k=1}^{N}\sum_{j=1}^{M} Cov(X_k, Y_j)$$
6. $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

Osservazione

$Cov(X,Y) > 0$ se X cresce (o decresce) insieme a Y
$Cov(X,Y) < 0$ se X decresce quando Y cresce e viceversa
$Cov(X,Y) = 0$ quando X e Y sono scorrelate

7.3.1 Coefficiente di correlazione

$$Cov(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \qquad \in [-1, 1]$$

Teorema

Se X e Y sono variabili casuali con $E[X] = \mu_X, E[Y] = \mu_Y$

$$E[XY] = E[X]E[Y]$$

7.4 Varianza di una variabile casuale

Data una variabile casuale X con valor medio $E[X] = \mu$ definito, si dice varianza di X (se esiste)

$$Var(X) = E[(X - E[X])^2]$$

Osservazione

Se X variabile casuale discreta con $X \in \{x_1, ..., x_n\}$

$$Var(X) = \sum_{k=1}^{n} (x_k - \mu)^2$$

Se X variabile continua

$$Var(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\ dx$$

7.4.1 Proprietà

1. $$Var(X) = E[X^2] - (E[X])^2$$

2. $\forall \alpha, \beta \in \mathbb{R}$
$$Var(\alpha x + \beta) = \alpha^2 \cdot Var(X)$$

   Conseguenze:
   Se $\alpha = 0$
$$Var(0 \cdot x + \beta) = Var(\beta) = 0^2 \cdot Var(X) = 0 \qquad \forall \beta \in \mathbb{R}$$

   Se $\alpha \neq 0$
$$Var(\alpha x + \beta) = \alpha^2 \cdot Var(X) \qquad \forall \beta \in \mathbb{R}$$

3. Date X e Y variabili casuali con valor medio e varianza definiti
$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

3bis. Date N variabili casuali $X_1, X_2, ..., X_N$

$$Var\Big(\sum_{k=1}^{N} x_k\Big) = \sum_{k=1}^{N} x_k Var(X_k) + \sum_{k=1}^{N}\sum_{j=1}^{N} Cov(X_k, Y_j)$$

7.5 Applicazioni con variabili casuali binomiali

$X \sim B_e(n, p)$
$X = Y_1 + ... + Y_n$ con $Y_k \sim B_e(p)$

$$Y_k = \begin{cases} 1 & \text{se nel k-simo esperimento si verifica } A \\ 0 & \text{altrimenti} \end{cases}$$

$E[X] = E[Y_1 + Y_2 + ... + Y_n] = E[Y_1] + E[Y_2] + ... + E[Y_n] = n \cdot p$

$$Var(X) = Var\left(\sum_{k=1}^{n} Y_k\right) \underbrace{=}_{Y_k\ indipendenti} \sum_{k=1}^{n} Var(Y_k) = \sum_{k=1}^{n} pq = npq$$

Esempio

Lancio di una moneta equilibrata.

$X =$ 'n^o di uscite di T'
$X \sim B(2, \frac{1}{2})$

$$E[X] = np = 2 \cdot \frac{1}{2} = 1$$

$$Var(X) = npq = 2 \cdot \frac{1}{2} \cdot (1 - \frac{1}{2}) = \frac{1}{2}$$

8 Variabili casuali identicamente distribuite

Le *variabili casuali identicamente distribuite* sono variabili con la stessa funzione di ripartizione (o distribuzione) di probabilità.

Nota Bene

- nel caso di variabili casuali discrete devono avere la stessa funzione di massa di probabilità;
- nel caso di variabili casuali continue devono avere la stessa funzione di densità di probabilità

8.1 Funzioni generatrici di momenti

Momento di ordine n:

$$E[X] = \begin{cases} \displaystyle\sum_{k=1}^{n} x_k^n \cdot p(x_k) \\ \\ \displaystyle\int_{-\infty}^{+\infty} x^n f(x)\ dx \end{cases}$$

Data una variabili casuale X, si definisce funzione generatrice dei momenti di X (se esiste)

$$\Phi(t) = E[e^{tx}] \qquad \Phi : \mathbb{R} \to \mathbb{R}$$

8.1.1 Proprietà

0. $\Phi(0) = E[1] = 1 \qquad \forall v.c.\ X$

1. $$\begin{aligned} \left.\frac{d\Phi(t)}{dt}\right|_{t=0} &= \left.\frac{d}{dt}E[e^{tX}]\right|_{t=0} \\ &= \left.E[Xe^{tX}]\right|_{t=0} \\ &= E[X] \end{aligned}$$

2. $$\frac{d^k\Phi(E)}{dt^k} = E[X^k], \qquad k > 1$$

3. Se X e Y v.c. indipendenti
$$\Phi_{X+Y}(t) = \Phi_X(t) \cdot \Phi_Y(t)$$

Esempio 1

$X \sim B_e(p)$ (variabile Bernoulliana)

$$\begin{aligned} \Phi(t) &= E[e^{tX}] \\ &= e^{t\cdot 0}(1-p) + e^{t\cdot 1} \cdot p \\ &= q + e^t \cdot p \end{aligned}$$

per $n \leq 1$

$$\begin{aligned} E[X^n] &= \left.\frac{d^n\Phi(t)}{dt^n}\right|_{t=0} \\ &= \left.e^t \cdot p\right|_{t=0} \\ &= p \end{aligned}$$

Esempio 2

$X \sim B_e(n,p)$ (variabile Binomiale)
$X \in \{0,1,...,n\}$

$$P(k) = \binom{n}{k} = p^k(1-p)^{n-k} \qquad k = 0,1,...,n$$

$$\Phi(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tX}\binom{n}{k}p^k \cdot q^{n-k}$$

Questo è il metodo standard per calcolare la funzione generatrice di momenti; ma se che una variabile binomiale è la somma di bernoulliane: $X = \sum_{j=1}^{n} Y_j$ dove $Y_j \sim B_e(p)$ con

$$Y_j = \begin{cases} 1 & \text{se A si verifica nel } j-simo \text{ tentativo} \\ 0 & \text{altrimenti} \end{cases}$$

$Y_1, Y_2, ..., Y_n$ corrispondono agli n esperimenti svolti in maniera indipendente e identica, quindi $Y_1, ..., Y_n$ sono variabili casuali indipendenti e identicamente distribuite (i.i.d.)

$$\Phi_X(t) = \Phi_{Y_1,Y_2,...,Y_n}(t) = E[e^{t(Y_1+Y_2,...+Y_n)}]$$

siccome le Y_j sono indipendenti posso scrivere quest'espressione come

$$\Phi_{Y_1}(t) + \Phi_{Y_2}(t) + ... + \Phi_{Y_n}(t) = \underbrace{(e^t \cdot p + q)\cdot(e^t \cdot p + q)\cdot ... \cdot(e^t \cdot p + q)}_{n\ volte} = (e^t \cdot p + q)^n$$

Esempio

X v.c. continua

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

$$\begin{aligned} \Phi(t) &= E[e^{tX}] \\ &= \int_{-\infty}^{+\infty} e^{tx} f(x)\ dx \\ &= \underbrace{\int_{-\infty}^{0} e^{tx}\cdot 0\ dx}_{0} + \int_{0}^{+\infty} e^{tx}\cdot e^{-x}\ dx \\ &= \int_{0}^{+\infty} e^{-(1-t)x}\ dx \end{aligned}$$

$$\begin{aligned} \int_0^{+\infty} e^{-(1-t)x}\ dx &= \begin{cases} \Big[-\dfrac{1}{1-t}\cdot e^{-(1-t)x}\Big]_0^{+\infty} & 1-t>0 \\ \text{diverge} & 1-t \leq 0 \end{cases} \\ &= \begin{cases} \dfrac{1}{1-t} & 1-t>0 \\ \text{diverge} & 1-t \leq 0 \end{cases} \end{aligned}$$

Quindi se $t < 1,\ \Phi(t) = \dfrac{1}{1-t}$

$$E[X] = \frac{d\Phi(t)}{dt}\Big|_{t=0} = \frac{1}{(1-t)^2}\Big|_{t=0} = 1$$

oppure

$$E[X] = \int_{+\infty}^{-\infty} x \cdot f(x)\ dx = \int_{+\infty}^{0} xe^{-x}\ dx$$

$$E[X^2] = \int_0^{+\infty} x^2 e^{-x}\ dx$$

$$
\begin{aligned}
E[X^2] &= \frac{d^2\Phi(t)}{dt}\Big|_{t=0} \\
&= \frac{d^2}{dt}(\frac{1}{1-t})\Big|_{t=0} \\
&= \frac{d}{dt}\left(\frac{1}{(1-t)^2}\right)\Bigg|_{t=0} \\
&= -2(1-t)^{-3}(-1)\big|_{t=0} \\
&= 2(1-t)^-3\big|_{t=0} = 2
\end{aligned}
$$

$$Var(X) = 2 - 1^1 = 1$$

8.1.2 Proprietà di riproducibilità delle variabili casuali binomiali

$$\left.\begin{aligned} X &\sim B(n,p) \\ T &\sim B(m,p) \end{aligned}\right\} \text{indipendenti} \implies X + T \sim B(n+m, p)$$

8.2 Disuguaglianza di Markov

Data una variabile casuale $X \geq 0$ (con valori non negativi) e dato $a \in \mathbb{R}^+$

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Con questa formula si riesce a stimare $P(X \geq a)$ usando solo $E[X]$. Affinché la disuguaglianza di una stima sia utile deve valere $\frac{E[X]}{a} < 1$

8.3 Disuguaglianza di Čebyšëv

Data una variabile casuale X (con media $\mu = E[X]$ e varianza $Var(X) = \sigma^2$) e dato $r \in \mathbb{R}$

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

disuguaglianza utile se $\frac{\sigma^2}{r^2} < 1$

Immaginiamo di associare ad ogni esperimento una variabile casuale che rappresenti l'esito dell'esperimento. Se l'esperimento viene ripetuto in maniera identica e indipendente N volte allora abbiamo N variabili casuali independentie identicamente distribuite: $\underbrace{X_1, X_2, ..., X_N}$ campione *i.i.d.*. Si definisce così la media aritmetica delle variabili casuali (detta anche media campionaria)

$$\bar{X} = \frac{X_1, X_2, ..., X_N}{N}$$

$\bar{X}$ è una variabile casuale, quindi è possibile calcolarne il valore medio e la varianza

$$E[\bar{X}] = E[X_k] = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{N} \qquad k = 1, ..., N$$

8.3.1 Scarto quadratico medio e deviazione standard

$$\sigma = \sqrt{Var(X_k)} \qquad \sqrt{Var(\bar{X})} = \frac{\sigma}{\sqrt{N}}$$

8.4 Legge dei grandi numeri

Data una successione di variabili casuali i.i.d. $X_1, X_2, ..., X_N$ (con $E[X] = \mu$ e $Var(X_k) = \sigma^2$ con $k = 1, ..., N$) allora $\forall \epsilon > 0$ piccolo a piacere

$$P(|\bar{X} - \mu| \geq \epsilon) \underset{N \to +\infty}{\longrightarrow} 0$$

ovvero la media campionaria converge in probabilità alla media teorica.

L'idea della legge è che μ rappresenti la quantità fisica da misurare. *La legge dei grandi numeri* garantisce che su ∞ prove c'è probabilità nulla che $\bar{X}$ si discosti da μ. Ergo fare più esperimenti riduce l'errore.

8.5 Corollario di Bernoulli

$$\lim_{N \to +\infty} P(|f_A - p| \geq \epsilon) = \lim_{N \to +\infty} P\left(\left|\frac{\sum\limits_{k=1}^{N} x_k}{N} - p\right| \geq \epsilon\right) = 0$$

f_A = frequenza relativa di A, ovvero la definizione frequentista di probabilità

9 Modelli di variabili casuali discrete

9.1 Variabile casuale di Bernoulli

In un esperimento l'evento A (successo) si verifica con $P(A) = p$

$$X \sim Be(p) \qquad X = \begin{cases} 1 & \text{se A si verifica} \\ 0 & \text{altrimenti} \end{cases}$$

$$p(1) = p \qquad p(0) = 1 - p = q$$

$$E[X] = p \qquad Var(X) = pq \qquad \Phi(t) = pe^t + q$$

9.2 Variabile casuale Binomiale

Si ripete un esperimento in maniera identica e indipendente n volte. L'evento A si verifica in una prova con probabilità $P(A) = p$.

X = 'n^o di esperimenti in cui si verifica A'

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

$0 \leq k \leq n \quad (k \in \mathbb{N})$

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$$

$Y_1, Y_2, .., Y_n \sim Be(p)$ indipendenti tra loro

$$X = \sum_{j=1}^{n} Y_j$$

N.B. X è la somma di v.c.i.i.d.

$$E[X] = np \qquad Var(X) = npq \qquad \Phi(t) = (pe^t + q)^n$$

Riproducibilità della binomiale:

$$X \sim B(n, p), Z \sim B(m, p) \text{ indipendenti} \Longrightarrow X + Z \sim B(n + m, p)$$

9.3 Variabile casuale di Poisson (degli eventi rari)

Una variabile di Poisson rappresenta in genere il numero di eventi che si osservano in un certo intervallo temporale/spaziale a patto che il singolo evento abbia una probabilità bassa di verificarsi (evento raro).

Ad esempio il numero di accessi al sito internet, il numero di persone che si mettono in coda in posta, il numero di telefonate al centralino. etc...

$$X \sim P_o(\lambda) \qquad\qquad \lambda \in \mathbb{R}^+$$

$X \in \{0, 1, 2, ...\}$ ovvero $X \in \mathbb{N}$

$$p(k) = P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

$$E[X] = \lambda \qquad Var(X) = \lambda \qquad \Phi(t) = e^{\lambda(e^t-1)}$$

9.3.1 Proprietà di riproducibilità

$$X \sim P_o(\lambda), Y \sim P_o(\mu) \text{ indipendenti} \Longrightarrow X + Y \sim P_o(\lambda + \mu)$$

infatti

$$\Phi_{X+Y}(t) = \Phi_X(t) \cdot \Phi_Y(t) = e^{\lambda(e^t-1)}e^{\mu(e^t-1)} = e^{(\lambda+\mu)(e^t-1)}$$

9.4 Variabile casuale geometrica

Si ripete un esperimento in maniera identica e indipendente fino ad osservare un successo (l'evento A con $P(A) = p$).

$X \sim G(p)$ se $X =$ 'n^o di prova per osservare un successo'
$X \in \{1, 2, ..., +\infty\}$ ovvero $X \in \mathbb{N}/\{0\}$, $k = 1, 2, ..., +\infty$

$$p(k) = (1-p)^{k-1} \cdot p = q^{k-1} \cdot p$$

$$E[X] = \frac{1}{p}$$

$$Var(X) = \frac{q}{p^2}$$

9.5 Esempi

9.5.1 Esempio 1

Una compagnia di assicurazioni riceve in media 5 richieste di rimborso al giorno.

1. Che probabilità c'è che arrivino meno di 3 chiamate?
   $X =$ 'n^o di richieste in un giorno'
   $X \sim P_o(\lambda)$ con $E[X] = 6 = \lambda$

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{5^0}{0!}e^{-5} + \frac{5^1}{2!}e^{-5} + \frac{5^2}{2!}e^{-5} = e^{-5}[1 + 5 + \frac{25}{4}] = \frac{49}{4}e^{-5}$$

2. Con che probabilità in 5 giorni arriveranno 20 richieste?
   $X_k =$ 'n^o di richieste nel k-simo giorno
   $X_k \sim P_o(5)$
   $Y =$ 'n^o di richieste in 5 giorni'

$$Y = \sum_{k=1}^{5} X_k \sim P_o(\lambda + \lambda + \lambda + \lambda + \lambda) = P_o(5\lambda) = P_o(25)$$

$$P(Y = 20) = \frac{25^{20}}{20!}e^{-25}$$

9.5.2 Esempio 2

Da studi statistici si sa che il voto di uno studente all'esame di fine corso è una v.c. di media 75.

1. Dare un limite superiore/stimare la probabilità che il punteggio di uno studente scelto a caso non sia inferiore a 85
   $X =$ 'punteggio di uno studente'
   X v.c. con $E[X] = 75$
   $$P(X \geq 85) \leq \frac{E[X]}{85} = \frac{75}{85}$$

2. Quanti studenti devono sostenere l'esame affinché vi sia una probabilità non più grande di 0.10 che la media dei punteggi della sessione disti non meno di 5 da 75, se la $Var(X_k) = 25$ $X_1, X_2, ..., X_n$ punteggi degli n studenti
   $$\bar{X} = \frac{\sum_{k=1}^{n} X_k}{n} \Longrightarrow E[\bar{X}] = E[X_k] = 75$$
   $$Var(\bar{X}) = \frac{Var(X_k)}{n} = \frac{25}{n}$$
   $$P(|\bar{X} - 75| \geq 5) \leq 0.10$$
   Usiamo la disuguaglianza di Čebyšëv
   $$P(|\bar{X} - 75| \geq 5) \leq \frac{Var(\bar{X})}{5^2} = \frac{25}{n} \cdot \frac{1}{25} = \frac{1}{n} \Longrightarrow \frac{1}{n} = 0.10 \Longrightarrow n = 10$$

3. Qual è la probabilità che uno studente ottenga un punteggio compreso tra 64 e 85 (estremi esclusi)?
   $$P(65 < X < 85) = P(|X - \underbrace{75}_{E[X]}| < 10) = 1 - P(|X - 75| \geq 10)$$
   per la disuguaglianza di Čebyšëv
   $$P(|X - 75| \geq 10) \leq \frac{Var(X)}{10^2} = \frac{25}{100} = \frac{1}{4} \Longrightarrow 1 - P(|X - 75| \geq 10) \geq 1 - \frac{1}{4} = \frac{3}{4}$$
   Senza altre informazioni è possibile fare solo una stima della probabilità.

9.6 Variabile casuale binomiale negativa

Si ripete un esperimento in maniera identica e indipendente fino a osservare l'evento A (successo) r volte ($r \in \mathbb{N}/\{0\}$)

$$X = \text{'}n^o \text{ di esperimenti necessari per osservare } A \; r \text{ volte'}$$

$p = P(A)$ probabilità del verificarsi di A in un singolo esperimento.
$X \sim NB(r, p)$
$X \in \{r, r+1, r+2, ..., +\infty\}$
$k = r, r+1, ...$

$$p(k) = P(X = k) = \binom{k-1}{n-1} \cdot p^r \cdot (1-p)^{k-r}$$

$X \sim NB(r, p)$ è descrivibile come

$$X = \sum_{j=1}^{r} Y_j \qquad \text{con } Y_j \sim G(p) \text{ e } Y_1, Y_2, ..., Y_j \text{ indipendenti}$$

$$E[X] = \frac{r}{p} \qquad Var(X) = \frac{rq}{p^2}$$

10 Modelli di variabili casuali continue

10.1 Variabili casuali uniformi

$$X \sim U(\alpha, \beta) \qquad\qquad \alpha, \beta \in \mathbb{R}, \alpha < \beta$$

$$f(x) = \begin{cases} k & \text{se } x \in [\alpha, \beta] \\ 0 & \text{altrimenti} \end{cases}$$

10.1.1 Proprietà

$f(x) \geq 0 \Rightarrow k \geq 0$

$k = \dfrac{1}{\beta - \alpha}$

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{se } x \in [\alpha, \beta] \\ \\ 0 & \text{altrimenti} \end{cases}$$

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\, dx = \begin{cases} 0 & a < \alpha \\ \\ \dfrac{a - \alpha}{\beta - \alpha} & \alpha \leq a \leq \beta \\ \\ 1 & a > \beta \end{cases}$$

$$E[X] = \frac{\beta + \alpha}{2} \qquad\qquad Var(X) = \frac{(\beta - \alpha)^2}{12}$$

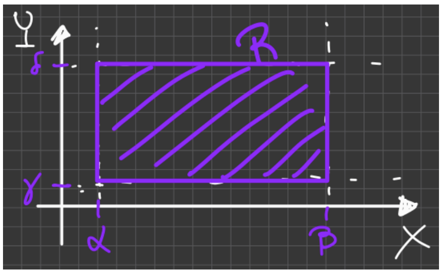
Più grande è l'intervallo più grande è la varianza.

Nota: le funzioni *random* usano variabili casuali uniformi.

10.2 Coppie di variabili casuali uniformi

$$\left.\begin{matrix} X \sim U(\alpha, \beta) \\ Y \sim U(\gamma, \delta) \end{matrix}\right\} \textit{indipendenti} \qquad\qquad \alpha, \beta \in \mathbb{R}, \alpha < \beta \ \text{ e } \ \gamma, \delta \in \mathbb{R}, \gamma < \delta$$

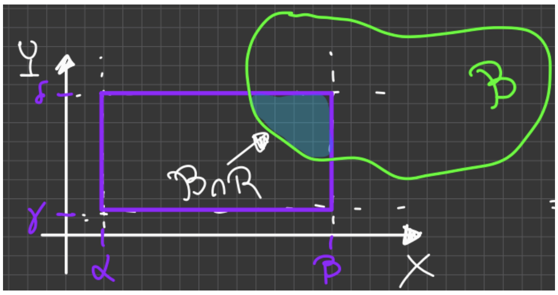
$$f(x, y) = f_X(x) \cdot f_Y(y) = \begin{cases} \dfrac{1}{\beta - \alpha} \cdot \dfrac{1}{\delta - \gamma} & x \in [\alpha, \beta], y \in [\gamma, \delta] \\ \\ 0 & \text{altrove} \end{cases}$$



R rettangolo di definizione di X e Y

$$Area(R) = (\beta - \alpha) \cdot (\delta - \gamma)$$

Immaginiamo di avere un sottoinsieme qualsiasi $\mathcal{B} \subset \mathbb{R}^2$



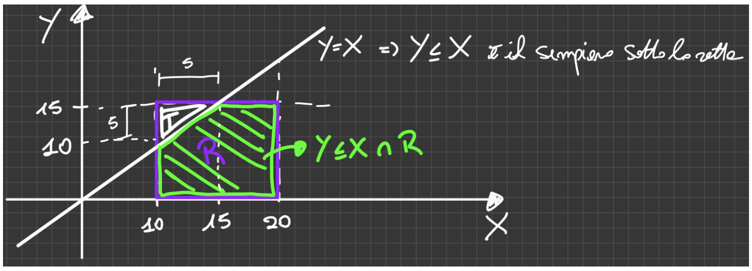
$$P((X,Y) \in \mathcal{B}) = \frac{Area(\mathcal{B} \cap R)}{Area(R)}$$

Esempio

L'autobus arriva alla fermata in un istante casuale tra le 8:10 e le 8:20. Anna arriva alla fermata in un orario casuale tra le 8:00 e le 8:15, indipendentemente dall'autobus. Qual è la probabilità che Anna prenda l'autobus?

$$\left.\begin{array}{l} X = \text{'arrivo dell'autobus'} \sim U(10,20) \\ Y = \text{'arrivo di Anna'} \sim U(0,15) \end{array}\right\} \textit{indipendenti}$$

$P(\text{Anna prende l'autobus}) = P(Y \leq X)$



$$P(Y \leq X) = \frac{Area(Pentagono)}{Area(R)} = \frac{Area(R) - Area(T)}{Area(R)} = \frac{(10 \cdot 15) - (\frac{5 \cdot 5}{2})}{10 \cdot 15} = \frac{150 \cdot 25/2}{150} = \frac{300 - 25}{300} = \frac{275}{300} = \frac{11}{12}$$

10.2.1 Metodi Monte Carlo

Metodi che sfruttano la casualità/probabilità per calcolare una quantità deterministica.

Esempio: misura della della superficie di un lago attraverso il lancio di palle di cannone.
Si lanciano delle palle di cannone all'interno del rettangolo (di lati l_1 e L_2) che contiene il lago. (X,Y) = punto di arrivo delle palle di cannone.

$$\left.\begin{array}{l} X \sim U(0,l_1) \\ Y \sim U(0,l_2) \end{array}\right\} \textit{indipendenti}$$

$$P(\text{palla di cannone finisce nel lago}) = P((X,Y) \in L) = \frac{Area(L)}{Area(R)}$$

Se l'esperimento di lancio viene ripetuto n volte ($n >> 1$) allora

$$P(\text{palla di cannone finisce nel lago}) \approx f_L$$

dove $f_l = \frac{n_L}{n} = \frac{n^o \text{ di lanci che finiscono nel lago}}{n^o \text{ di lanci totali}}$.
Per il corollario di Bernoulli se $n >> 1$

$$\frac{n_l}{n} = f_L \approx P(\text{palla di cannone finisce nel lago}) = \frac{Area(L)}{Area(R)} \Longrightarrow Area(L) = \frac{Area(R) \cdot n_L}{n}$$

10.2.2 Problema degli spilli di Buffon (1717)

Il problema degli spilli di Buffon, antecedente al metodo Monte Carlo, si basa su un esperimento in cui abbiamo degli spilli di lunghezza $2a$ e delle rette parallela distanti tra loro $2b$ con $2a < 2b$. Gli spilli vengono lanciati in maniera casuale sul piano.
Buffon propone di stimare il π considerando la frequenza relativa dell'evento

'lo spillo lanciato cade intersecando una retta'

- M = punto medio dello spillo
- X = 'distanza di M dalla retta più vicina'

$$X \sim U(0, b)$$

- θ = 'angolo tra la retta di direzione comune passante per M e lo spillo'

$$\theta = U(0, \pi)$$

Si ha intersezione tra la retta e lo spillo se

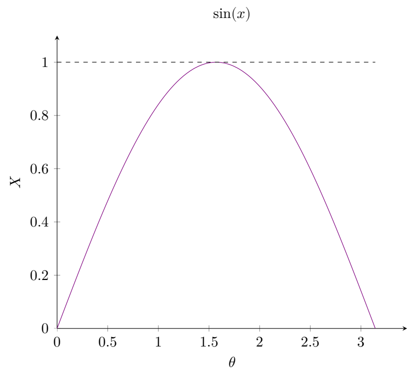
$$X \leq a \sin\theta \Longrightarrow P(\text{spillo interseca la retta } r) = P(X \leq a \sin\theta)$$

Ipotesi: X e θ indipendenti.
Quindi se sono indipendenti possiamo considerare X, θ una coppia di v.c. continue indipendenti

$$f(x, y) = f_X(x) \cdot f_\theta(y)$$

$$f_X(x) = \begin{cases} \frac{1}{b} & x \in [0, b] \\ \\ 0 & \text{altrove} \end{cases} \qquad f_\theta(y) = \begin{cases} \frac{1}{\pi} & y \in [0, \pi] \\ \\ 0 & \text{altrove} \end{cases}$$

$$P(X \leq a \sin\theta) = \int_0^{\pi} \int_0^{a \sin\theta} f(x, y)\ dxdy$$



$$\begin{aligned} a \sin\theta &= \int_0^{\pi} \int_0^{a \sin\theta} \frac{1}{b} \cdot \frac{1}{\pi}\ dxdy \\ &= \int_0^{\pi} \frac{1}{b\pi} \Big[x\Big]_0^{a \sin y} \\ &= \frac{1}{b\pi} \int_0^{\pi} a \sin y\ dy \\ &= \frac{a}{b\pi} \Big[-\cos y\Big]_0^{\pi} \\ &= \frac{2a}{b\pi} \end{aligned}$$

quindi:

$$P(\text{spillo interseca una retta}) = \frac{2a}{b\pi}$$

Ripetendo n volte il lancio dello spillo ($n >> 1$) per il *corollario di Bernoulli*

$$f_A = \frac{n_a}{n}$$

con $n_A = n^o$ di esperimenti in cui lo spillo cade intersecando una retta

$$\frac{2a}{b\pi} \approx \frac{n_A}{n} \Longrightarrow \pi \approx \frac{2an}{bn_A}$$

A livello pratico questo esperimento è difficile da realizzare, e chi ha provato a farlo ha ottenuto risultati deludenti. Si potrebbe eseguire al computer, ma per farlo dobbiamo dare a θ un valore random tra 0 e π, quindi si deve conoscere il valore di π a priori, rendendo l'esperimento insensato.

10.2.3 Esempio 1 sulle variabili casuali uniformi

Immaginiamo di avere 3 dispositivi in serie D_1, D_2, D_3 indipendenti

10.3 Variabile casuale esponenziale

$$X \sim \varepsilon(\lambda) \qquad\qquad \lambda \in \mathbb{R}$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrove} \end{cases} \qquad\qquad F(a) = \begin{cases} 0 & a < 0 \\ 1 - e^{-\lambda a} & a \geq 0 \end{cases}$$

$$\Phi(t) = \int_{-\infty}^{+\infty} e^{tx} \cdot f(x)\ dx = \begin{cases} \text{diverge} & t \geq \lambda \\ \dfrac{\lambda}{\lambda - t} & t < \lambda \end{cases}$$

$$E[X] = \frac{1}{\lambda} \qquad\qquad E[X^2] = \frac{2}{\lambda^2} \qquad\qquad Var(X) = \frac{1}{\lambda^2}$$

Le v.c. esponenziali non sono riproducibili

10.3.1 Proprietà

1. $X \sim \mathcal{E}(\lambda), c \in \mathbb{R}^+$

$$Y = cX \sim \mathcal{E}\Big(\frac{\lambda}{c}\Big) \Longrightarrow \Phi_Y(t) = \Phi_X(tc) = \frac{\lambda/c}{\lambda/c - t} \qquad\qquad t < \frac{\lambda}{c}$$

2. Dispositivi in serie $D_1, D_2, ..., D_n$, ciascuno con un tempo di funzionamento di tipo esponenziale

$$D_k \text{ ha tempo di funzionamento } X_k \sim \mathcal{E}(\lambda k) \qquad\qquad \lambda, k \in \mathbb{R}^+$$

$X_1, X_2, ..., X_n$ indipendenti, $T =$ *'tempo di funzionamento del sistema'*

$$F_T(t)$$

10.4 Variabile casuale Gaussiana o Normale

$$X \sim \mathcal{N}(\mu, \sigma^2) \qquad\qquad \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^2$$

$$f(x) == \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

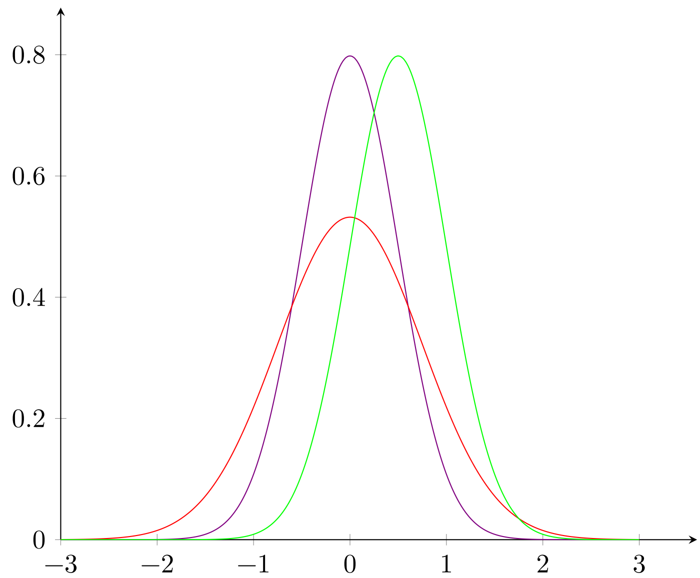
$$E[X] = \mu \qquad\qquad Var(X) = \sigma^2$$

$$I = \sqrt{2\pi} \qquad\qquad I^2 = 2\pi$$

$$\int_{-\infty}^{+\infty} f(x)\ dx = 1 \qquad\qquad F(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ dx$$

$$\Phi(t) = e^{t\mu + \frac{t^2}{2}\sigma^2}$$

Grafico variabile casuale Gaussiana



In viola la funzione di densità di probabilità con parametri $(0, 0.5)$, in rosso con parametri $(0, 0.75)$, in verde con parametri $(0.5, 0.5)$.

Il picco della Gaussiana cresce o cala in proporzione inversa a σ e la forma della campana è più schiacciata oppure più concentrata intorno a $x = \mu$, mantenendo l'area unitaria.

10.4.1 Proprietà

1. $X \sim \mathcal{N}(\mu, \sigma^2)$ e $Y = \alpha X + \beta$ con $\alpha \in \mathbb{R}/\{0\}, \beta \in \mathbb{R}$

$$Y \sim \mathcal{N}(\alpha\mu + \beta, \alpha^2\sigma^2)$$

10.4.2 Applicazione delle proprietà 1 e v.c. normale standard

$\alpha = \frac{1}{\sigma}, \beta = -\frac{\mu}{\sigma}$

$$X \sim \mathcal{N}(\mu, \sigma) \Longrightarrow Y \sim \mathcal{N}\Big(\frac{1}{\sigma}\mu + \Big(-\frac{\mu}{\sigma}\Big), \frac{1}{\sigma^2} \cdot \sigma^2\Big)$$

allora se definisco una variabile $X = \frac{Z - \mu}{\sigma}$

$$Z \sim \mathcal{N}(0, 1)$$

questa variabile viene chiamata **variabile casuale normale standard**.

$$\begin{aligned} F_X(a) &= P(X \leq a) \\ &= P(\underbrace{\frac{1}{\sigma}X - \frac{\mu}{\sigma}}_{Z\sim\mathcal{N}(0,1)} \leq \frac{1}{\sigma}a - \frac{\mu}{\sigma}) \\ &= P(Z \leq \frac{a - \mu}{\sigma}) \\ &= F_Z(\frac{a - \mu}{\sigma}) \end{aligned}$$

2. $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ e $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ v.c. indipendenti

$$T = X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Esempio

Le precipitazioni annuali in una città americana hanno distribuzione normale con $\mu = 12.08$ pollici, $\sigma = 3.1$ pollici.
Si suppone l'indipendenza delle precipitazioni tra anni diversi.

1. Determinare la probabilità che le precipitazioni superino complessivamente i 25 pollici.
   X_1 = precipitazioni nel primo anno, X_2 = precipitazioni nel secondo anno

$$X_1 \sim \mathcal{N}(12.08, (3.1)^2) \qquad\qquad X_2 \sim \mathcal{N}(12.08, (3.1)^2)$$

   X_1, X_2 indipendenti

$$S = X_1 + X_2 \sim \mathcal{N}(12.08 + 12.08, (3.1)^2 + (3.1)^2)$$

$$\begin{aligned} P(S > 25) &= P\left(\frac{S - 24.16}{\sqrt{2(3.1)^2}} > \frac{25 - 24.16}{\sqrt{2(3.1)^2}}\right) \\ &= P(Z > 0.192) \\ &= P(Z \leq 0.192) \\ &\approx 1 - F_Z(0.19) \\ &\approx 0.42 \end{aligned}$$

2. Determinare la probabilità che le precipitazioni del prossimo anno superino quelle dell'anno seguente di almeno 3 pollici.

$$P(X_1 > X_2 + 3) = P(X_1 - X_2 > 3)$$

$$\begin{aligned} X_1 &\sim \mathcal{N}(12.08, (3.1)^2) \qquad\qquad X_2 \sim \mathcal{N}(12.08, (3.1)^2) \\ -X_2 &\sim \mathcal{N}(-12.08, (3.1)^2) \end{aligned}$$

$$T = X_1 - X_2 \sim \mathcal{N}(12.08 - 12.08, (3.1)^2 + (3.1)^2) = \mathcal{N}(0, 2 \cdot (3.1)^2)$$

$$\begin{aligned} P(T > 3) &= P\left(\frac{T - 0}{\sqrt{2 \cdot (3.1)^2}} > \frac{3 - 0}{\sqrt{2 \cdot (3.1)^2}}\right) \\ &= P(Z > 0.68) = 1 - F_Z(0.68) \\ &\approx 25\%(0.2482) \end{aligned}$$

11 Processi stocastici e processi stocastici di Poisson

Un processo stoacastico è una famiglia variabili casuali che dipendono da un parametro reale (spesso questo parametro è il tempo).

11.1 Processo stocastico di Poisson

Un proceso stocastico si dice di Poisson se rappresenta una famiglia di variabili casuali

$$\mathcal{N}(t) = n^o \text{ di eventi che si verificano in }]0, t]$$

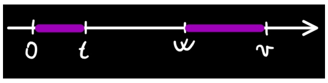
dove t è il tempo e $\mathcal{N}(t)$ sono v.c. discrete con le ipotesi

1. $\mathcal{N}(0) = 0$
2. **Stazionarietà degli incrementi:** il numero di eventi che si verificano in un certo intervallo di tempo dipende dalla lunghezza dell'intervallo, ma non dipende dalla posizione sull'asse reale, $t > 0, s > 0$

   Se prendiamo in esempio il grafico qui sopra la proprietà 2 ci dice se $\mathcal{N}(t)$ si comporta nello stesso modo nei due segmenti evidenziati in viola.

3. **Indipendenza degli incrementi:** il numero di eventi che si verificano in un dato intervallo è indipendente dal numero di eventi che si verificano in un intervallo disgiunto al primo



4.
$$\lim_{h\to 0} \frac{P(\mathcal{N}(h)=1)}{h} = \lambda \qquad \lambda \in \mathbb{R}^+$$

Se $h << 1,\ P(\mathcal{N}(h)=1) \approx \lambda h$

5.
$$\lim_{h\to 0} \frac{P(\mathcal{N}(h)\geq 2)}{h} = 0$$

Se $h << 1,\ P(\mathcal{N}(h)\geq 2) \approx 0$

11.1.1 Distribuzione di probabilità

Che distribuzione di probabilità presenta $\mathcal{N}(t)$?
Sappiamo che per definizione $\mathcal{N}(t) = n^o$ di eventi che si verificano in $]0,t]$

dividiamo l'intervallo $(0,t)$ in m sotto-intervalli (con $m >> 1$) di lunghezza t/m; in un intervallo piccolissimo la probabilità che in quell'intervallo ci sia più di un evento è 0.

$$P(\mathcal{N}(t)=k) \qquad k = 0,1,....$$

In ciascuno di questi sotto-intervalli si può verificare un evento con probabilità $\lambda t/m$ (per affermare questo usiamo la proprietà 2 e 3) e si possono verificare 2 o più eventi con probabilità 0.

$$P(\mathcal{N}(t)=k) = \binom{m}{k}\Big(\frac{\lambda\cdot t}{m}\Big)^k\Big(1-\frac{\lambda t}{m}\Big)^{m-k}$$

e questa non è altro che una distribuzione binomiale.

Per $m \to +\infty$
$$B\Big(m, \frac{\lambda t}{m}\Big) \longrightarrow P_o\Big(m\cdot\frac{\lambda t}{m}\Big) \Longrightarrow \mathcal{N}(t) \sim P_o(\lambda t)$$
quindi $\mathcal{N}(t)$ è formata da tante possoniane con distribuzioni diverse.

Esercizio proposto

(X,Y) v.c. continue con
$$f(x,y) = \begin{cases} a(x+y) & x\ in[0,1], y \in [0,1] \\ 0 & \text{altrove} \end{cases}$$

Determinare a.
Chiamiamo $Z = \max(X,Y)$ e calcoliamo $F_Z(z)$ e $f_Z(z)$.
$f(x,y)$ è una funzione di densità di probabilità se

- $$f(x,y) \geq 0 \Rightarrow a \geq 0$$
- $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\ dxdy = 1$$

$$\begin{aligned}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\ dxdy &= \int_0^1\int_0^1 a(x+y)\ dxdy = \\ &= a\int_0^1 \Big[\frac{x^2}{2}+yx\Big]_0^1 dy \\ &= a\int_0^1 (\frac{1}{2}+y)\ dy = a\Big[\frac{y}{2}+\frac{y^2}{2}\Big]_0^1 \\ &= a \Rightarrow a = 1\end{aligned}$$

$$
\begin{aligned}
F_Z(z) &= P(Z \leq z) \\
&= P(\max(X, Y) \leq z) \\
&= P(X \leq z, Y \leq z) \\
&= \int_{-\infty}^{z}\int_{-\infty}^{z} f(x,y)\ dxdy \\
&= \begin{cases} \displaystyle\int_{-\infty}^{0}\int_{-\infty}^{0} 0\ dx = 0 & z < 0 \\ \displaystyle\int_{0}^{z}\int_{0}^{z} (x+y)\ dxdy & 0 \leq z \leq 1 \\ 1 & z > 1 \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\int_{0}^{z}\int_{-\infty}^{z} (x+y)\ dxdy &= \int_{0}^{z}\left[\frac{x^2}{2} + xy\right]_0^z dy \\
&= \int_{0}^{z} \frac{z^2}{2} + zy\ dy \\
&= \left[\frac{z^2}{2}y + \frac{zy^2}{2}\right]_0^z \\
&= z^3
\end{aligned}
$$

$$
F_Z(z) = \begin{cases} 0 & z < 0 \\ z^3 & 0 \leq z \leq 1 \\ 1 & z > 1 \end{cases} \qquad f_Z(z) = \begin{cases} 0 & z < 0 \\ 3z^2 & 0 \leq z \leq 1 \\ 0 & z > 1 \end{cases}
$$

$$
E[Z] = \int_{-\infty}^{+\infty} z f_Z(z)\ dz = \int_0^1 3z^2 \cdot z\ dz = \left[\frac{3}{4}z^4\right]_0^1 = \frac{3}{4}
$$

$$
E[Z^2] = \int_{-\infty}^{+\infty} z^2 3z^2\ dz = \left[\frac{3}{5}z^5\right]_0^1 = \frac{3}{5}
$$

$$
Var(Z) = \frac{3}{5} - (\frac{3}{4})^2 = \frac{48-45}{80} = \frac{3}{80}
$$

11.1.2 Intertempi di un processo stocastico di Poisson

Definiamo gli intertempo $X_1, X_2, ...$ come

$$X_1 = \textit{'tempo che intercorre tra lo stato iniziale e il verificarsi del primo evento'}$$

Se X_1 è una v.c. continua quanto varrà $P(X_1 > s)$ (con $s > 0$)?

$$
\begin{aligned}
P(X_1 > s) &= P(\mathcal{N}(s) = 0)\} \rightarrow \text{ il numero di eventi tra }]0, s] \text{ è nullo} \\
\mathcal{N}(s) \sim P_o(\lambda s) \leftarrow \{ &= \frac{(\lambda s)^0}{0!} e^{-\lambda s} \\
&= e^{-\lambda s}
\end{aligned}
$$

$$
F_{X_1}(s) = \begin{cases} 1 - e^{-\lambda s} & s > 0 \\ 0 & s < 0 \end{cases}
$$

$F_{X_1}(s)$ non è altro che la funzione di distribuzione di una v.c. esponenziale di parametro λ.

$$X_2 = \textit{'tempo che intercorre tra il primo e il secondo evento'}$$

con s_1 = primo evento

$$P(X_2 > w) = P(\text{ci sono 0 eventi in }]0, w]) \qquad \} \rightarrow \text{per l'indipendenza degli eventi e la stazionarietà degli incrementi}$$

$$= P(\mathcal{N}(w) = 0) = \frac{(\lambda w)^0}{0!} e^{-\lambda w}$$

$$= e^{-\lambda w}$$

$$F_{X_2}(w) = \begin{cases} 1 - e^{-\lambda w} & w \geq 0 \\ 0 & w < 0 \end{cases}$$

In modo simile si dimostra che

$$X_k = \text{'}tempo\ tra\ il\ (k-1) - simo\ e\ il\ k - simo\ evento\text{'} \Longrightarrow X_k \sim E(\lambda)$$

12 Variabile casuali note

12.1 Funzione di una variabile casuale nota

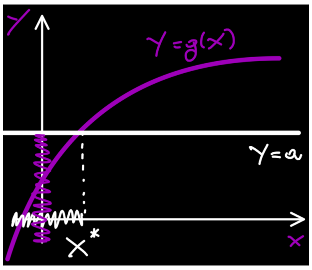
$Y = g(X)$ con X v.c. nota

$$E[Y] = E[g(X)] = \begin{cases} \displaystyle\sum_{j=1}^{n} g(x_j)p(x_j) & x \text{ v.c. discreta}, x \in \{x_1, x_2, ..., x_j\} \\ \displaystyle\int_{-\infty}^{+\infty} g(x)f(x)\ dx & x \text{ v.c. continua} \end{cases}$$

$$E[Y^2] = E[g^2(X)] = \begin{cases} \displaystyle\sum_{j=1}^{n} g^2(x_j)p(x_j) & x \text{ v.c. discreta}, x \in \{x_1, x_2, ..., x_j\} \\ \displaystyle\int_{-\infty}^{+\infty} g^2(x)f(x)\ dx & x \text{ v.c. continua} \end{cases}$$

Nel caso continuo come si può determinare la funzione di ripartizione o la funzione di densità di probabilità della Y?

12.1.1 Caso di una funzione g monotona crescente (una funzione monotona è invertibile)



X v.c. continua nota con funzione di densità di probabilità $f_X(x)$

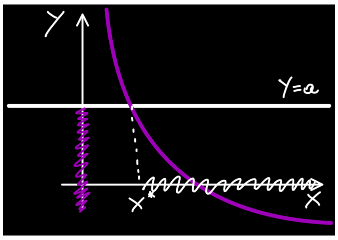
$$\begin{aligned} F_Y(a) &= P(Y \leq a) \\ &= P(g(X) \leq a) \\ &= P(X \leq x^*) \\ &= \int_{-\infty}^{x^*} f_X(x)\ dx \qquad g(x^*) = a \Rightarrow x^* = g^{-1}(a) \\ &= \int_{\infty}^{g^{-1}(a)} f_X(x)\ dx \end{aligned}$$

Quindi

$$F_Y(a) = \int_{\infty}^{g^{-1}(a)} f_X(x)\ dx$$

$$f_Y(a) = \frac{dg^{-1}(a)}{da} f_X(g^{-1}(a)) = \begin{cases} \frac{1}{\alpha} & \beta \leq a \leq \beta * \alpha \\ \\ 0 & \text{altrove} \end{cases} \sim U(\beta, \beta + \alpha)$$

12.1.2 Caso g funzione monotona decrescente



$$F_Y(a) = \int_{g^{-1}(x)}^{*\infty} f(x)\ dx$$

$$f_Y(a) = -\frac{dg^{-1}(a)}{da} f_X(g^{-1}(a))$$

N.B. g è monotona decrescente $\Rightarrow \dfrac{dg^{-1}(a)}{da} < 0$

Per scrivere un'unica formula per una funzione g monotona

$$f_Y(a) = \left|\frac{dg^{-1}(a)}{da}\right| f_X(g^{-1}(a))$$

Esempio

$$X \sim U(0,1) \qquad\qquad Y = -\ln(X)$$

$$f_Y(a) = -\frac{dg^{-1}(a)}{da} \cdot f_X(g^{-1}(a))$$

$$a = -\ln(x) = g(x) \Longrightarrow x = g^{-1}(a) = e^{-a}$$

$$\frac{dg^{-1}(a)}{da} = -e^{-a} \Longrightarrow -\frac{dg^{-1}(a)}{da} = e^{-a}$$

$$X \sim (0,1) \leftarrow \{f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{altrove} \end{cases} \Longrightarrow f_X(g^{-1}(a)) = \begin{cases} 1 & 0 \leq e^{-a} \leq 1 \quad (\text{per } a \geq 0) \\ 0 & \text{altrove} \end{cases}$$

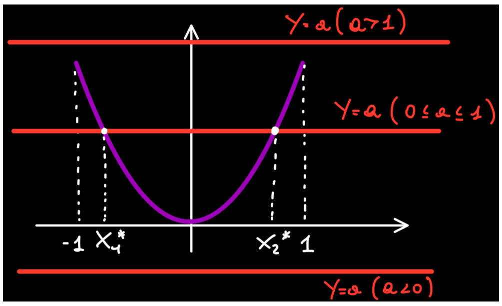
$$\begin{aligned} f_Y(a) &= e^{-a} \cdot \begin{cases} 1 & a \geq 0 \\ 0 & \text{altrove} \end{cases} \\ &= \begin{cases} e^{-a} & a \geq 0 \\ 0 & \text{altrove} \end{cases} \\ &\Longrightarrow Y \sim \varepsilon(1) \end{aligned} \qquad\qquad p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

12.1.3 Caso g funzione non monotona

Nel caso in cui $Y = g(X)$, g funzione non monotona non è possibile scrivere una formula generale, ma, come vedremo nel prossimo esempio, se g è una funzione 'semplice', e per semplice si intende che è crescente in un certo intervallo e decrescente in un altro, si può applicare la formula precedente.

Esempio

$$X \sim U(-1,1) \qquad\qquad Y = x^2$$



In questo caso il valor medio e la varianza esulano dalla funzione di densità di Y

$$E[Y] = E[X^2] = \int_{-1}^{1} \frac{1}{2}x^2\ dx = \left[\frac{x^3}{6}\right]_{-1}^{1} = \frac{1}{3}$$

$$E[Y^2] = E[X^4] = \int_{-1}^{1} \frac{1}{2}x^4\ dx = \left[\frac{x^5}{10}\right]_{-1}^{1} = \frac{1}{5}$$

$$Var(X) = E[Y^2] - E^2[Y] = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$$

$$F_Y(a) = P(Y \leq a) = \begin{cases} 0 & a < 0 \\ \displaystyle\int_{x_1^*}^{x_2^*} f_X(a)\ dx & 0 \leq a \leq 1 \\ 1 & a > 1 \end{cases}$$

$$\int_{x_1^*}^{x_2^*} f_X(a)\ dx = \int_{x_1^*}^{x_2^*} \frac{1}{2}\ dx = \left[\frac{x}{2}\right]_{x_1^*}^{x_2^*} = \frac{x_2^* - x_1^*}{2}$$

Siccome $Y = x^2$ possiamo scrivere x_1^* e x_2^* in funzione di a

$$\left.\begin{array}{l} (x_1^*)^2 = a \\ (x_2^*)^2 = a \\ x_1^* < 0 \\ x_2^* > 0 \end{array}\right\} \Longrightarrow \begin{cases} x_1^* = -\sqrt{a} \\ x_2^* = \sqrt{a} \end{cases}$$

quindi torniamo al calcolo dell'integrale

$$\int_{x_1^*}^{x_2^*} f_X(a)\ dx = \frac{\sqrt{a} - (-\sqrt{a})}{2} = \sqrt{a}$$

$$F_Y(a) = \begin{cases} 0 & a < 0 \\ \sqrt{a} & 0 \leq a \leq 1 \\ 1 & a > 1 \end{cases}$$

12.2 Variabile casuale Lognormale

$$X \sim \mathcal{N}(\mu, \sigma^2) \qquad \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+$$
$$Y = e^X \sim \text{Lognormal}(\mu, \sigma^2)$$

$$f_Y(a) = \begin{cases} \dfrac{1}{a\sigma\sqrt{2\pi}} \exp\left(-\dfrac{(\ln(x)-\mu)^2}{2\sigma^2}\right) & a > 0 \\ 0 & a \leq 0 \end{cases} \qquad F_Y(a) = \begin{cases} F_Z\left(\dfrac{\ln(x)-\mu}{\sigma}\right) & a > 0 \\ 0 & a \leq 0 \end{cases}$$

$$E[Y] = \exp\left(\mu + \frac{\sigma^2}{2}\right) \qquad E[Y^2] = \exp(2\mu + 2\sigma^2) \qquad Var(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right) - \exp(2\mu + 2\sigma^2)$$

12.2.1 Proprietà

$$\left.\begin{aligned} Y_1 &\sim \text{Lognormal}(\mu_1, \sigma_1^2) \\ Y_2 &\sim \text{Lognormal}(\mu_2, \sigma_2^2) \end{aligned}\right\} \text{indipendenti}$$

$$\begin{aligned} Y_1 &= e^{x_1} & \qquad Y_2 &= e^{x_2} \\ X_1 &\sim \mathcal{N}(\mu_1, \sigma_1^2) & \qquad X_2 &\sim \mathcal{N}(\mu_2, \sigma_2^2) \end{aligned}$$

$$T = Y_1 \cdot Y_2 = e^{X-1} \cdot e^{x_2} = e^{x_1 + x_2}$$

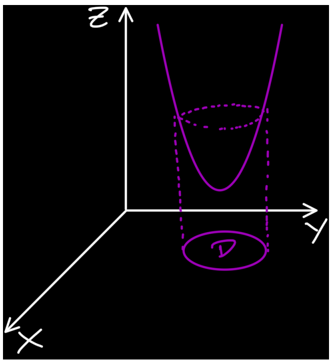
Per la riproducibilità delle v.c. gaussiane

$$W = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$
$$T = e^W \Longrightarrow T \sim \text{Lognormal}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

12.3 Funzioni di due variabili casuali continue note

$$Z = h(X, Y)$$

se (X, Y) è una coppia di v.c. note sarà nota anche $f(x, y)$



$$\begin{aligned} F_Z &= P(Z \leq a) \\ &= P(h(X,Y) \leq a) \\ &= \iint_{\mathcal{D}} f(x,y)\ dxdy \end{aligned}$$

12.3.1 Caso particolare di $Z = X + Y$

$$F_Z(a) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{-x+a} f(x,y)\ dy \right) dx \qquad F_Z(a) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{-y+a} f(x,y)\ dx \right) dy$$

$$f_Z(a) = \int_{-\infty}^{+\infty} f(x, -x+a)\ dx \qquad f_Z(a) = \int_{-\infty}^{+\infty} f(-y+a, y)\ dy$$

12.3.2 Caso particolare di variabili casuali indipendenti

$$(X, Y) \text{ v.c. indipendenti } \Longrightarrow f(x, y) = f_X(x) + f_Y(y)$$

$$f_Z(a) = \int_{-\infty}^{+\infty} f_X(-y+a) \cdot f_Y(y)\ dy \qquad\qquad f_Z(a) = \int_{-\infty}^{+\infty} f_Y(x) \cdot f_Y(-x+a)\ dx$$

questi vengono chiamati **prodotti di convoluzione**.

12.3.3 Esempio 1

$X \sim \mathcal{E}(1), Y \sim \mathcal{E}(2)$ indipendenti; se $Z = X + Y$ determinare $f_Z(a)$.

$$f_Z(a) = \int_{-\infty}^{+\infty} \underbrace{f_X(x) \cdot f_Y(a-x)}_{I}\ dx$$

$$f_X(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{altrove} \end{cases} \qquad\qquad f_Y(y) = \begin{cases} e^{-y} & y \geq 0 \\ 0 & \text{altrove} \end{cases}$$

$I \neq 0$ se $x \geq 0, a - x \geq 0 \Rightarrow x \geq 0, x \leq a\ (a \geq 0)$

$$\begin{aligned} I &= \int_0^a e^{-x} \cdot e^{-(a-x)}\ dx \\ &= \int_0^a e^{-x-a+x}\ dx \\ &= e^{-a} \int_0^a\ dx \\ &= ae^{-a} \end{aligned}$$

$$f_Z(a) = \begin{cases} ae^{-a} & a \geq 0 \\ 0 & \text{altrove} \end{cases}$$

$$F_Z(b) = \int_{-\infty}^{b} f_Z(a)\ da = \begin{cases} 1 - (b+1)e^{-b} & b \geq 0 \\ 0 & \text{altrove} \end{cases}$$

12.3.4 Esempio 2

$$\left.\begin{array}{l} X \sim U(0,1) \\ Y \sim U(0,1) \end{array}\right\} \text{indipendenti} \qquad\qquad Z = X + Y$$

$$F_Z(a) = P(Z \leq a) = P(X + Y \leq a) = \begin{cases} 0 & a < 0 \\ 1 & a > 2 \end{cases}$$

La somma di due variabili che vanno da 0 a 1 può assumere valori tra 0 e 2, quindi è logico che non può essere minore di 0 ed è sempre minore o uguale a 2.
Ora bisogna calcolare il caso in cui $0 \leq a \leq 2$. Discerniamo i due casi $0 \leq a \leq 1$ e $1 < a \leq 2$.

Se $0 \leq a \leq 1$

$$\begin{aligned} P(Y < -X + a) &= \frac{Area(T)}{Area(rettangolo)} \\ &= \frac{a \cdot a \cdot 1/2}{1^2} \\ &= \frac{a^2}{2} \end{aligned}$$

Se $1 < a \leq 2$

$$\begin{aligned} P(Z \leq a) &= P(Y + X \leq a) \\ &= P(Y \leq -X + a) \\ &= \frac{Area(Pentagono)}{Area(Rettangolo)} \\ &= \frac{Area(Rettangolo) - Area(T_2)}{Area(Rettangolo)} \\ &= \frac{1 - (2-a) \cdot (2-a) \cdot 1/2}{1} \\ &= 1 - \frac{(2-a)^2}{2} \end{aligned}$$

Riassumendo

$$F_Z(a) = P(X + Y \leq a) = \begin{cases} 0 & a > 0 \\ \frac{a^2}{2} & 0 \leq a \leq 1 \\ 1 - \frac{(2-a)^2}{2} & 1 < a \leq 2 \\ 1 & a > 2 \end{cases}$$

$$f_Z(a) = \frac{dF_Z(a)}{da} = \begin{cases} 0 & a < 0 \\ a & 0 \leq a \leq 1 \\ (2-a) & 1 < a \leq 2 \\ 0 & a > 2 \end{cases}$$

13 Teorema del limite centrale o teorema centrale del limite

Abbiamo $X_1, X_2, ..., X_n$ v.c indipendenti e identicamente distribuite con

$$\left. \begin{aligned} E[X_k] &= \mu \\ Var(X_k) &= \sigma^2 \end{aligned} \right\} \forall k = 1, ..., n$$

Se $S_n = \sum\limits_{k=1}^{n}$

$$\begin{aligned} E[S_n] &= E\left[\sum_{k=1}^{n} X_k\right] \\ &= \sum_{k=1}^{n} E[X_k] \\ &= \sum_{k=1}^{n} \mu \\ &= n\mu \end{aligned}$$

$$\begin{aligned} Var(S_n) &= Var\left(\sum_{k=1}^{n} X_k\right) \\ \text{indipendenza} \leftarrow \{ &= \sum_{k=1}^{n} Var(S_n) \\ &= \sum_{k=1}^{n} \sigma^2 \\ &= n\sigma^2 \end{aligned}$$

Qui sorge un problema: per n che tende a infinito, il valor medio e la varianza tendono a infinito, e questo non ha molto senso; per questo è meglio definire la somma delle nostre variabili S_n nel seguente modo:

$$\begin{aligned} Y_n &= \frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \\ &= \frac{S_n - n\mu}{\sigma\sqrt{n}} \end{aligned}$$

$$\begin{aligned} E[Y_n] &= E\left[\frac{S_n - n\mu}{\sigma\sqrt{n}}\right] \\ &= \frac{1}{\sigma\sqrt{n}} E[S_n] - \frac{n\mu}{\sigma\sqrt{n}} \\ &= \frac{n\mu}{\sigma\sqrt{n}} - \frac{n\mu}{\sigma\sqrt{n}} \\ &= 0 \end{aligned}$$

$$\begin{aligned} Var(Y_n) &= Var(\frac{S_n}{\sigma\sqrt{n}} - \frac{n\mu}{\sigma\sqrt{n}}) \\ Var(\alpha x + \beta) = \alpha^2 Var(X)\} &= \left(\frac{1}{\sigma\sqrt{n}}\right)^2 Var(S_n) \\ &= \frac{\sigma^2 n}{\sigma^2 n} \\ &= 1 \end{aligned}$$

In questo modo otteniamo un valor medio e una varianza unitari.
$\forall a \in \mathbb{R}$

$$\lim_{n\to+\infty} F_{Y_n} = F_Z(a) \qquad\qquad Z \sim \mathcal{N}(0,1)$$

perché Y_n converge in distribuzione alla normale standard.

13.1 Applicazioni

1

$$X \sim B(n,p) \Longrightarrow X = \sum_{k=1}^{n} T_k \qquad\qquad T_k \sim B_e(p)$$

$$Y_n = \frac{X - np}{\sqrt{npq}} \underset{n\to+\infty}{\longrightarrow} X \dot{\approx} \mathcal{N}(np, npq) \qquad\qquad n \geq 30$$

$\dot{\approx}$ vuol dire "si comporta circa come".
N.B questo teorema può essere applicato solo per $n \geq 0$ e più grande è n più precisa è l'approssimazione.

2

Se $X_1, X_2, ..., X_n$ sono v.c.i.i.d. allora la loro media aritmetica

$$\bar{X} = \frac{\sum_{k=1}^{n} X_k}{n} \qquad\qquad E[\bar{X}] = \mu \qquad\qquad Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \underset{n\to+\infty}{\longrightarrow} \text{converge in distribuzione alla normale standard} \qquad\qquad n \geq 30$$

3

$S_n = X_1 + X_2 + ... + X_n$ con $X_1, X_2, ..., X_n$ v.c.i.i.d.

$$S_n \dot{\approx} \mathcal{N}(n\mu, \sigma^2 n) \qquad\qquad n \geq 30$$