Riassunto CPS

Probabilità di un evento

$$P(\text{evento}) = \frac{\text{numero casi favorevoli}}{\text{numero casi possibili}}$$

Permutazione n elementi → In quanti modi posso cambiare l'ordine di n elementi

$$\text{Perm}(n) = n!$$

Permutazione n elementi con ripetizione:

$$\text{Perm}_n^{\alpha} = \frac{n!}{\alpha!} \quad \begin{array}{l} \to \text{Numero tot. di elementi} \\ \to \text{Numero di elementi ripetuti} \end{array}$$

$$\text{Perm}_n^{\alpha,\beta,\gamma} = \frac{n!}{\alpha!\,\beta!\,\gamma!}$$ → se ho diversi elem. ripetuti

Esempio: $\text{Perm}(\text{MASCARA}) = \frac{7!}{3!} = 7 \cdot 6 \cdot 5 \cdot 4$

In quanti modi posso scegliere k elementi da un insieme di n, $k \leq n$ (coefficiente binomiale)

Non importa l'ordine!!

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

n = numero di elementi

k = numero di elementi nel sottoinsieme

es: In quanti modi posso estrarre 3 carte da un mazzo di carte composto da 52 carte

Proprietà:

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

2) $\binom{n}{1} = \binom{n}{n-1} = n$

3) $\binom{n}{k} = \binom{n}{n-k}$

4) $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

Disposizioni, modi di prendere k oggetti ordinati tra n elementi. Importa l'ordine!

$$D_{n,k} = S_{n,k} \cdot k! = \frac{n!}{(n-k)!}$$

Probabilità condizionata P(succeda A sapendo che è succ. B)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \xRightarrow{A \text{ e } B \text{ indipendenti se}} P(A|B) = P(A)$$

"Qual'è la probabilità che succeda A sapendo B?"

$$\Updownarrow$$

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

$$\Downarrow$$

$\forall$ scelta di indici $i_1, \dots, i_k$ con $k \leq n$

abbiamo: $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$

SPAZIO PROBABILIZZATO: (Ω, F, P)

a) Ω = Insieme degli eventi elementari (spazio campione)

Es: per un dado $\{1, 2, 3, 4, 5, 6\}$

b) F = Sottoinsieme delle $\mathbb{P}(\Omega)$ $\quad (F \subseteq \mathbb{P}(\Omega))$

$\Downarrow$

Insieme di tutti i possibili insiemi partendo dagli elementi di Ω

Esempio: $A = \{0, 1\}$ $\quad \mathbb{P}(A) = \{\emptyset, 0, 1, \{0,1\}\}$

$\downarrow$

$\#\mathbb{P}(A) = 2^{\#A} = 4$

L'insieme F deve essere chiuso per queste operazioni :
- Unione
- Intersezione
- Passaggio al complementare

$\left(A,B \in F \Rightarrow \begin{matrix} A \cup B \\ A \cap B \\ A^c, B^c \end{matrix} \in F \right)$

Condizione aggiuntiva : Se $A_1, ..., A_n \in F \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in F$

(unione degli insiemi da A_1 a A_n sono in F)

σ - Algebra (tribù)

(Insieme che rispetta la proprietà di unione numerabile e passaggio al complementare)

c) P = Probabilità $\quad P : F \Rightarrow [0,1]$

Proprietà :
- $P(\Omega) = 1$
- $A_1, ..., A_n$ disgiunti $A_i \cap A_j \neq 0 \quad \forall i,j \Rightarrow P(\bigcup_n A_n) = \sum_n P(A_n)$
- $P(A) \geq 0$

Se non sono verificate, lo spazio non è probabilizzato

P = Probabilità assegnata a ogni possibile evento compreso in Ω

Formule : (spazio probabilistico)

1) $A \cap A^c = \emptyset \quad A \cup A^c = \Omega$ e $P(\Omega) = 1 \Rightarrow P(A^c) = 1 - P(A)$

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

Se $A \cap B = 0$
(eventi indipendenti)

"Qual'è la probabilità che succeda A o che succeda B ?"

3) $A^c = \Omega \setminus A$

4) $A \subseteq B \Rightarrow P(A) \leq P(B)$

5) $P(A|B) = \dfrac{P(A\cap B)}{P(B)} \underset{\text{Se } A\cap B = 0}{\Longrightarrow}$ (eventi indipendenti) $P(A|B) = P(A) \Rightarrow$ eventi A e B indipendenti

Prob condizionata

6) $P(A\cap B) = P(A)\cdot P(B|A) \underset{\text{Se } A\cap B = 0}{\Longrightarrow}$ (eventi indipendenti) $P(A\cap B) = P(A)\cdot P(B) \Rightarrow$ eventi A e B indipendenti

"Qual'è la probabilità che succeda A e che succeda B?"

7) $P(A\setminus B) = P(A) - P(B)$ se $B \subseteq A$

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

9) $P(\bigcup_{n\in\mathbb{N}} A_n) = 1 - P(\bigcap_{n\in\mathbb{N}} A_n)$

Densità di probabilità discreta

↳ Funzione che ha 2 proprietà:
- $P_n \geq 0$
- $\sum_{n\in\mathbb{N}} P_n = 1$

$\{P_n\}_n \in \mathbb{N}$ ↳ succ di numeri

Se P_n è un densità di probabilità $\Rightarrow \{\mathbb{N}, \mathcal{P}(n), P\}$ = spazio probabilizzato

$P(k) = P_n$ ↳ Probabilità

Se A, B indipendenti $\Rightarrow$ A, B^c; A^c, B^c; A^c, B indipendenti $\Rightarrow A = (A\cap B)\cup(A\cap B^c)$

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

$$\Updownarrow$$

$$P(A \cap B^c) = P(A) - P(A)\cdot P(B) = P(A)[1 - P(B)] = P(A)\cdot P(B^c)$$

Formula di Fattorizzazione

Ipotesi:

- Ω Insieme Fondamentale
- $B_1 \dots B_n$ partizione finita di Ω con:
  - $B_i \in F$
  - $B_i \cap B_j \neq 0 \quad \forall i,j$
  - $\bigcup_{i=1}^{n} B_i = \Omega$

Teorema: Sia $\{B_i\}_{1,\dots,n}$ Partizione finita di Ω

allora: $P(B_i) > 0 \Rightarrow P(A) = \sum_{i=1} \overbrace{P(A|B_i)\cdot P(B_i)}^{P(A \cap B_i)}$

Esempio: Sapendo di avere tre porte e la probabilità che il concorrente vinca sia:

$P(A|B_1) = 1$ se apre la porta 1

$P(A|B_2) = 1/2$ se apre la porta 2

$P(A|B_3) = 0$ se apre la porta 3

$P(B_1) = P(B_2) = P(B_3) = 1/3$

$$P(A) = \sum_{i=1}^{3} P(A|B_i)\cdot P(B_i) = \frac{1}{3} + \left(\frac{1}{2}\cdot\frac{1}{3}\right) + 0 = \frac{1}{3} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

Condizionamento ripetuto

Se $A_1, \dots, A_n$ sono eventi

$$\Rightarrow P(A_1 \cap \dots \cap A_n) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1 \cap A_2)\cdot \dots \cdot P(A_n|A_{n-1} \cap \dots \cap A_1)$$

Esempio: Quale è la probabilità che escano il numero 1, 2 e 3 consecutivamente su 100 numeri

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2)$$

Formula di Bayes:

Se $P(A)>0,\ P(B)>0 \implies P(A|B) = \dfrac{P(B|A) \cdot P(A)}{P(B)}$

Osservazione:

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

2) $P(A|B^c) \neq 1 - P(A|B)$

$$P(A_k | B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B|A_k)\, P(A_k)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$$

Lessico: Se considero un test di medicina che indica se una persona è positiva o negativa ad una malattia, avrò dei test che non potrò dare buoni (un campione ogni tot)

Quindi ho:

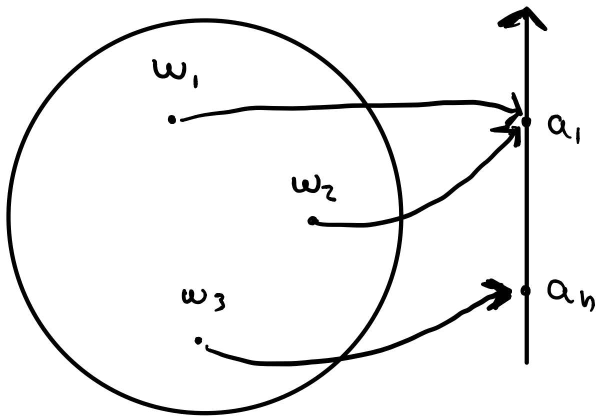
- Specificità: Probabilità che una persona sana abbia un test negativo
  - ↳ Alta: Pochi falsi positivi
- Sensibilità: Probabilità che una persona malata abbia un test positivo
  - ↳ Alta: Pochi falsi negativi

- Valore predictivo negativo: Probabilità che una persona con il test negativo sia sana
- Valore predictivo positivo: Probabilità che una persona con il test positivo sia malata

---

Variabile aleatorie discrete → Funzione che fa corrispondere un numero reale a ogni caso di un esperimento

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



$$p_J = \boxed{P(X = a_J)} = \boxed{P(X^{-1}(a_J))}$$

Probabilità che X prenda la controimmagine a_J

legge di probabilità di X

$$\boxed{p_J \text{ Densità di Probabilità}}$$

1) $p_J \geq 0$

2) $\sum_{J=1}^{n} p(J) = 1$

Proprietà:

1) $X^{-1} a_J$ sono tutti disgiunti perché parte un solo collegamento da ogni w

2) $X^{-1} a_J$ coprono tutto Ω

Legge di Bernoulli

Se considero 2 risultati di un esperimento: successo / insuccesso

X: successo → 1

X: insuccesso → 0

Se $p_1 = p \Rightarrow p_0 = 1 - p$

Legge Binomiale $B(n,p)$ (con reimbussolamento)

Se voglio fare un esperimento più volte e contare i successi $X: (0,1)^n \to (0, \dots, n)$

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

Prob che ottenga k successi con n esperimenti

Densità di probabilità

1) $p_k \geq 0$

2) $\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$

Proprietà: $B(n,p) + B(m,p) = B(n, m+p)$

Somma di variabili aleatorie

Se prendo due variabili aleatorie $X: \Omega \to \mathbb{R}$ Densità P_J

$Y: \Omega \to \mathbb{N}$ Densità Q_J

Ottengo che $Z = X + Y$ cioè posso sommarle

Indipendenza variabili aleatorie

Due variabili aleatorie X_1, X_2 sono indipendenti se

$\forall\ I_1, I_2 \subseteq \mathbb{R}$ (intervalli o semirette)

$$\Rightarrow P((X \in I_1) \cap (X_2 \in I_2)) = P(X \in I_1) \cdot P(X_2 \in I_2)$$

Variabili aleatorie congiunte

$$P(X = m \cap Y = n) = P_{n,m}$$

Densità di probabilità

1) $P_m \geq 0$
2) $\sum_m P_m = \sum_m \sum_n P_{m,n} = 1$

Se ho $P_{n,m}$ posso ricavarmi la prob. di P_n o P_m

$$P_m = P(X = m) = P(X = m \wedge \{\bigcup_n Y = n\}) = \sum_n P(X = m \wedge Y = n) = \sum_n P_{n,m}$$

$$P_n = P(Y = n) = P(Y = n \wedge \{\bigcup_m X = m\}) = \sum_m P(Y = n \wedge X = m) = \sum_m P_{n,m}$$

Viceversa se ho P_n e P_m non posso ricavarmi $P_{n,m}$

$$- P_{n,m} = P(X = m \cap Y = n) = \underbrace{P(X = m \mid Y = n)}_{\text{non so quanto vale}} \cdot P(Y = n)$$

Però se P_m e P_n sono indipendenti posso ricavarmi $P_{n,m}$

$$- P_{n,m} = P(X = m \mid Y = n) \cdot P(Y = n) = P(X = m) \cdot P(Y = n)$$

$$P(X=m \mid Y=n) = P(X=m)$$

Densità di probabilità

1) $P_{m,n} \geq 0$

2) $\sum_m \sum_n P_{m,n} = \sum_m \sum_n (P_m \cdot P_n) = \left(\sum_m P_m\right) \cdot \left(\sum_n P_n\right) = 1$

Formula di convoluzione

Se x e y sono indipendenti e $z = x+y$, la prob. che z sia un certo numero n è pari a:

$$P(z=n) = \sum_{i=0}^{n} P(x=i) \cdot P(y=n-i)$$

Variabile geometrica

La variabile geometrica conta la probabilità che al k-esimo tentativo io abbia un successo dopo $k-1$ insuccessi

$$P_k = P(\text{Geom}(p) = k) = (1-p)^{k-1} \cdot p$$

Densità di Probabilità →

1) $P_k \geq 0$

2) $\sum_{k=1}^{\infty} (1-p)^{k-1} p = p \cdot \sum_{k=0}^{\infty} (1-p)^{k-1} = p \cdot \frac{1}{1-(1-p)} = \cancel{p} \cdot \frac{1}{\cancel{p}} = 1$

Caratteristica:

$\text{Geom}(p)$ non ha memoria → Non si ricorda di aver sbagliato n tentativi, quindi l'$n+1$ esimo avrà la stessa probabilità

$$\Downarrow$$

$$P(\text{geom}(p) = n+m \mid \text{geom}(p) > n\} = P(\text{geom}(p) = m)$$

Probabilità che ho fatto $n+m$ esperimenti sapendo che i primi n sono andati male

Variabile ipergeometrica: Ho m elementi, k di un tipo e n estrazioni

(senza reimbussolamento)

$$P(H(m,k,n)=x) = \frac{\binom{k}{x}\cdot\binom{m-k}{n-x}}{\binom{m}{n}}$$

Probabilità che il numero di elementi di k estratti sia uguale a x

Binomiale negativa: $NB(n,p)$ ⇒ Conta i fallimenti prima dell'ennesimo successo

(Distribuzione di Pascal)

⇓

$N+NB$ esperimenti

$$P(NB(n,p)=k) = \binom{n+k-1}{k} p^n (1-p)^k$$

n = successi

k = insuccessi

Probabilità che in n successi abbia k insuccessi

Speranza/Valore atteso di x: (Media ponderata)

$$E[x] = \sum_{k=1}^{\infty} k\cdot P(x=k) = \sum_{k=1}^{\infty} k\cdot P_k$$

$$\Updownarrow$$

$$E[x] = \sum_{n=0}^{\infty} P(x>k) = \sum_{i=n+1}^{\infty} P(x=i) = \sum_{i=n+1}^{\infty} P_i$$

Valore atteso $f(x) = E[f(x)] = \sum_{k=1}^{\infty} f(k)\cdot P_k$

$$- E[B(n,p)] = \sum_{k=1}^{n} k\cdot\binom{n}{k}p^k(1-p)^{n-k} = n\cdot p$$

- $E[\text{Geom}(p)] = \sum_{k=0}^{n} k \cdot (1-p)^k \cdot p = \frac{1}{p}$
- $E[\text{Ber}(p)] = p$
- $E[NP(p,n)] = \frac{n}{p} - n$
- $E[\text{Poisson}(\mu)] = \mu$
- $E[H(m,k,n)] = \frac{n \cdot k}{m}$

Distribuzione di Poisson :

Conta i successi in un intervallo di tempo data la media di successi μ nell'intervallo.

$$\mu = E[B(n,p)] = n \cdot p$$

$$\boxed{P(\text{Poisson}(\mu) = k) = \frac{\mu^k}{k!} \cdot e^{-\mu}}$$

Densità di probabilità $\rightarrow$

1) $p_k \geq 0$
2) $\sum_k \frac{\mu^k}{k!} \cdot e^{-\mu} = e^{\mu} \cdot e^{-\mu} = 1$

Esempio: Ogni anno muoiono 12 cavalieri per un incidente a cavallo $\Rightarrow$ Probabilità che ne siano morti 7 l'anno 1861

$$P(\text{Poisson}(12) = 7) = \frac{12^7}{7! \cdot e^{12}} \simeq 0{,}04$$

La media è lineare

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

se x e y indipendenti:

$$E(x \cdot y) = E[x] \cdot E[y]$$

$$\Downarrow$$

se $E[x \cdot y] - (E(x) \cdot E[y]) \neq 0 \Rightarrow$ x, y sono dipendenti

2) $E[aX + b] = a\,E[x] + b$ $\quad a, b \in \mathbb{R}$, X var aleatoria

3) $E[\alpha X + \beta Y] = \alpha\,E[x] + \beta\,E[y]$

Momento di ordine superiore

$x : \Omega \to \mathbb{N}$ in generale

$$E[x^2] = \sum_{k=1}^{\infty} k^2 p_k \longrightarrow \boxed{E[x^n] = \sum_k k^n p_k}$$

Osservazione:

$E[x^2] \neq (E[x])^2$ perché X non è indipendente da se stessa

Varianza

la varianza serve a misurare la variabilità dei valori assunti dalla variabile. Nello specifico, indica di quanto si discostano i valori rispetto alla media $E[x]$

Sia x una var. aleatoria con $E[x] = \mu$

$$\boxed{Var(x) = E[(x - \mu)^2] = \sum_k (k - \mu)^2 p_k}$$

Oss.:

$$\mathrm{Var}[X] = E[X^2] - [E(X)]^2$$

Varianza Var. Aleatorie

$\mathrm{Var}(\mathrm{Bern}(p)) = p(1-p)$

$\mathrm{Var}(B(n,p)) = n \cdot p \cdot (1-p)$

$\mathrm{Var}(\mathrm{Geom}(p)) = \dfrac{1-p}{p^2}$

$\mathrm{Var}(H(m,k,n)) = \dfrac{n(m-n) \cdot k(m-k)}{m^2(m-1)}$

$\mathrm{Var}(\mathrm{Poisson}(\mu)) = \mu$

Disuguaglianze sulle var. aleatorie

Holder

$$E[X \cdot Y] \leq E[X^p]^{\frac{1}{p}} \cdot E[Y^q]^{\frac{1}{q}}$$

con $\frac{1}{p} + \frac{1}{q} = 1$

p e q due numeri positivi

Markov

Sia $X \geq 0$ $\quad a > 0$

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

Chebyshev

$$P(|X-\mu| > a) \leq \boxed{\frac{E[(X-\mu)^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}}$$

Oss.

$\mathrm{Var}(X) = 0 \Rightarrow X = \mu$

σ Deviazione standard

$$\boxed{\sigma(X) = \sqrt{\mathrm{Var}(X)}}$$

$$\sigma(\alpha X + \beta) = |\alpha| \, \sigma(X)$$

Proprietà della varianza:

X var aleatoria, media μ, $\alpha, \beta \in \mathbb{R}$

- $Var(\alpha X + \beta) = \alpha^2 Var(X)$
- $\sigma(\alpha X + \beta) = |\alpha| \sigma(X)$
- X, Y indipendenti $\Rightarrow Var(X+Y) = Var(X) + Var(Y)$

Processo stocastico: Successione di variabili aleatorie indicizzabili da una variabile temporale

$\{X_t\}_{t \in \underset{\text{tau}}{\overset{\parallel}{\tau}} \subseteq \mathbb{R}}$ $\quad \forall t \in X_t$ è una var. aleatoria discreta

t insieme finito o numerabile

legge marginale / marginale

Processo di Markov: Processo stocastico in cui il futuro è influenzato solo dal presente ($k-1$ tentativi precedenti li scarto, tengo solo il $k°$)

t_k = successione di tempi

e_k = successione di valori X_{t_k}

$$P(X_{t_{k+1}} = e_{k+1} \mid X_{t_k} = e_k \wedge X_{t_{k-1}} = e_{k-1} \wedge \dots \wedge X_{t_0} = e_0) \simeq P(X_{t_{k+1}} = e_{k+1} \mid X_{t_k} = e_k)$$

Esempi di processi di Markov

A) Estraggo, reimbusolo, estraggo ...

C) Estraggo, estraggo, reimbusolo la prima, estraggo la terza, reimbusolo la seconda ...

(Ad ogni estrazione lascio fuori solo l'ultima estratta)

Processi di Markov omogenei (CATENE DI MARKOV):

Processi di Markov t.c la

$$P\left(X_{t_{k+1}} = i \mid X_{t_k} = J\right) = P\left(X_1 = i \mid X_0 = J\right)$$

Sapendo i primi 2 risultati posso rappresentare tutti gli altri.

↳ due esempi sopra sono entrambi omogenei

Esempio Processo di Markov non omogeneo

A e C ma a ogni estrazione aggiungo 1 nera

RAPPRESENTAZIONE CATENA DI MARKOV

1) MATRICE DI TRANSIZIONE

$$Q = \left(q_{i \to J}\right) \forall_{i,J}$$

Matrice delle transizioni

per 2 stati

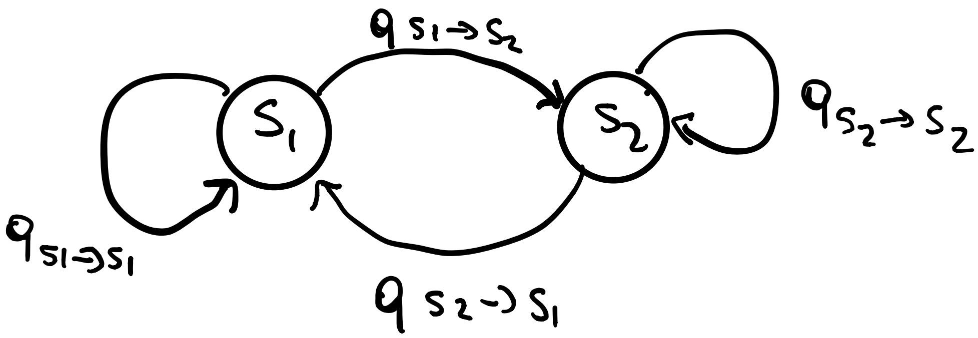
$$\begin{array}{c} \\ \#\text{stati} \end{array} \begin{array}{c} \overbrace{\qquad\qquad\qquad}^{\#\text{ stati}} \\ \begin{pmatrix} q_{0\to0} & q_{0\to1} \\ q_{1\to0} & q_{1\to1} \end{pmatrix} \end{array}$$

→ la somma degli elementi di Q per riga DEVE FARE "1"

È regolare se contiene tutti valori positivi

$$q_{i \to J} = \underbrace{P\left(X_1 = J \mid X_0 = i\right)}_{\text{Probabilità di passare da uno stato } i \text{ a uno stato } J}$$

2) RAPPRESENTAZIONE A GRAFO



Passeggiata aleatoria : Formalizzazione dell'idea di prendere passi successivi in direzioni casuali

$\{X_t\}$ = Coordinata partendo dall'origine e muovendosi lungo x

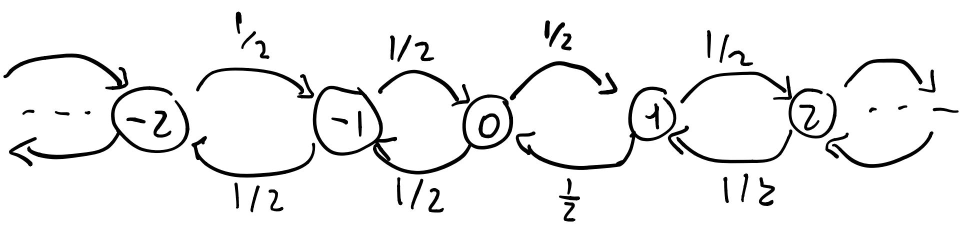
$\hookrightarrow \begin{cases} x_{k+1} = x_k + y_k & n \geq 1 \\ x_0 = 0 \end{cases}$

y_k = lancio di un moneta $\begin{cases} \nearrow \text{Testa} : -1 \\ \searrow \text{croce} : +1 \end{cases}$

$\hookrightarrow P(y_k = 1) = P(y_k = -1) = \frac{1}{2}$

Quindi ho:
- $P(x_{k+1} = J+1 \mid x_n = J) = \frac{1}{2}$
- $P(x_n = J-1 \mid x_n = J) = -\frac{1}{2}$

$\Big\}$ catena di Markov omogenea



B e R

Leggi Marginali di una catena di Markov

- Sia dato $\underline{P(x_0 = J)} = v = (v_J)$ vettore riga $\quad J \in$ (stati)

  $\hookrightarrow$ legge marginale iniziale

- Sia data $Q = (q_{iJ})_{iJ}$ matrice di transizione

  $\hookrightarrow q_{iJ} = q_{i \to J} = P(x_n = J \mid x_{n-1} = i) = P(x_1 = J \mid x_0 = i)$

Allora si può calcolare la legge marginale X_k come segue

$$\underbrace{P(X_k = i)}_{\text{Marginale al tempo } K} = (v_J) \cdot (Q)^k_J = v \cdot \underbrace{Q \cdot Q \cdot \ldots \cdot Q}_{K \text{ volte}}$$

$J \in$ Stati

$\Downarrow$

riga · colonna

Valore atteso:

- $(X_k)_{k=0,\ldots,n}$ catena di Markov $\quad X_n : \Omega \to E$
- $f : E \to \mathbb{R}$ Funzione che associa 1 valore $\in \mathbb{R}$ a uno stato

E: Stati

$\Downarrow$

1) Associamo ad f un vettore colonna

$$\vec{f}_J = f(J)$$

a) $E(f(X_k) \mid X_0 = i) = (Q^k \cdot \vec{f})_i \quad$ con $i \in E$

E: Stati

b) Se devo calcolare E per il tempo $K = 1$

$$E(f(X_1) \mid X_0 = 0) = \sum_i \Big(f(i) \cdot P(X_1 = i \mid X_0 = J) \Big) \quad \text{con } i, J \in E$$

Distribuzione invariante:

Un vettore riga $\vec{\mu}$ è detto distribuzione invariante per Q se

1) $\vec{\mu}_i > 0 \quad \forall i$

2) $\sum_i \vec{\mu}_i = 1$

3) $\vec{\mu}_i \cdot Q = \vec{\mu}_i \quad \Rightarrow (\vec{\mu}_i)^T$ Autovettore con autovalore 1 per matrice Q^T

Per trovare

μ bisogna risolvere il sistema omogeneo

$$(Q^T - id) \cdot \vec{\mu}^T = 0$$

OSS: $\vec{\mu}$ contiene $\# E$ elementi

$\downarrow$ stati

OSS 2: Se ho per esempio $\vec{\mu} = (\mu_1, \mu_2, \mu_3)$

$$\Rightarrow \quad \mu_1 = P(X_0 = E_1)$$
$$\mu_2 = P(X_0 = E_2)$$
$$\mu_3 = P(X_0 = E_3)$$

Dove E_1 = Stato 1, E_2 = Stato 2, E_3 = Stato 3

Catena stazionaria: Una catena di markov è detta stazionaria se tutte le sue leggi marginali sono uguali ossia:

$$P(X_k = i) = P(X_0 = i) \quad \forall i, k$$

OSS: Posto $\vec{\mu}$ corrispondente alla legge X_0 la catena è stazionaria se e solo se

$$\vec{\mu} \cdot Q^k = \vec{\mu}$$

Quindi se (X_k) è stazionaria $\Rightarrow$ $\vec{\mu} = P(X_0 = i)$ è distribuzione per Q

Variabile Aleatoria con densità → la densità deve essere sempre positiva

Una var. aleatoria X sullo spazio (Ω, A, P) si dice con densità se $\exists\ f: \mathbb{R} \to \mathbb{R}^+$ t.c

$$P(X \in A) = \int_A f(x)\,dx$$

$f(x)$ è detta densità di X

Se X ha densità $\Rightarrow P(X = x) = \int_{\{x\}} f(t)\,dt = 0$

OSSERVAZIONE: $$\int_{-\infty}^{+\infty} f(x)\,dx = P(\Omega) = 1$$

Funzione di ripartizione

La funzione di ripartizione F di una v.a con densità X è:

$$F(x) = P(X \leq x) = P(X \in [-\infty, x]) = \int_{-\infty}^{x} f(t)\,dt$$

Proprietà Funz. di ripartizione

1) $s < t \Rightarrow F_X(s) \leq F_X(t)$

2) $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$

3) $F(x) = \lim_{y \to x^+} F(y)$

V.a indipendenti

1) X e Y indipendenti $\iff P\{X \in A, Y \in B\} = P\{X \in A\} \cdot P\{Y \in B\}$

2) // $\iff p(x_i, y_j) = p_1(x_i) \cdot p_2(y_j)$

3) // $\iff f(x,y) = f_1(x) \cdot f_2(y)$

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

4) Se X e Y indipendenti e $Z = X+Y$ $\Rightarrow$ densità di Z

$$\Downarrow$$

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

Valore Atteso / Speranza

$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)\,dx \quad \text{se e solo se} \quad \int_{-\infty}^{+\infty} |x|\, f(x)\,dx < +\infty$$

N.B: se $f(x) = \begin{cases} 0 & \text{per } x \leq 0 \\ \dots & \text{per } x > 0 \end{cases}$ allora $E[X] = \int_0^{+\infty} x\, f(x)\,dx =$

- $\nearrow$ $+\infty$ (Non esiste)
- $\searrow$ numero reale (esiste)

$$\boxed{E[X^k] = \int_{-\infty}^{+\infty} x^k f(x)\,dx}$$

esiste se $\int_{-\infty}^{+\infty} |x|^n f(x)\,dx < +\infty$

Teorema: Supponiamo di avere $1 \leq m < n$

Se esiste il momento di ordine n, esiste anche quello di ordine m

N.B Se ho v.a X con densità f e Y che dipende da X
$(Y = g(X))$

$$\Rightarrow \boxed{E[g(X)] = \int g(x) \cdot f(x)} \, dx \text{ se esiste } \int |g(x)| f(x) \, dx < +\infty$$

Proprietà

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

2) $E[aX] = a E[X]$

3) Se $X \geq 0$, $E[X] \geq 0$ $\Rightarrow$ $X \leq Y \Rightarrow E[X] \leq E[Y]$

4) Se X e Y indipendenti $\Rightarrow$ $E[X, Y] = E[X] \cdot E[Y]$

Varianza → Misura quanto i valori si discostano dalla speranza

Se X ha un momento secondo $E[X^2]$ finito

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

Quindi:

$$Var(X) = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x) \, dx$$

Proprietà:

$$Var(x) = E[x^2] - E^2[x] = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{+\infty} x f(x)\,dx\right)^2$$

Covarianza: Misura quanto due variabili variano insieme.

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

Due variabili con $covar(X,Y) = 0$ sono indipendenti

Oss.

1) $Covar(X,X) = Var(X)$

2) $Var(X+Y) = Var(X) + Var(Y) + 2\,covar(X,Y)$

Disuguaglianza:

1) Markov: $a P(X > a) \leq E[X]$ per X positivi e $a > 0$

2) Chebishev: $P(|X - E[X]| > a) \leq \frac{Var(x)}{a^2}$ per $a > 0$

↳ $Var(x) = 0$ se x costante

Calcolare densità funzione $y = h(X)$
$\hookrightarrow$ X con densità f

Metodo 1:

1) Si calcola $F_Y(y) = P(Y \leq y)$

2) Si deriva

Metodo 2:

- Supponiamo che f_X è diverso da 0 su un intervallo $A = \,]a,b[$
- h biunivoca derivabile da A su B

$$f_Y(y) = \begin{cases} 0 & \\ f_X\left(h^{-1}(y)\right) \cdot \left| \dfrac{dh^{-1}(y)}{dy} \right| & y \in B \end{cases}$$

Densità uniforme su $[a,b]$ e Funzione di ripartizione $F(x)$

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \text{altrove} \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{se } x \leq a \\ \dfrac{x-a}{b-a} & \text{se } a < x < b \\ 1 & \text{se } x \geq b \end{cases}$$

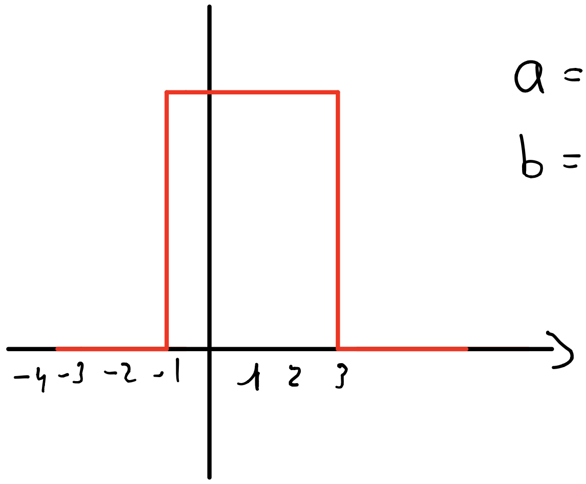
$$E[X] = \frac{a-b}{2}$$

$$E[X^2] = \int_a^b \frac{x^2}{b-a}\,dx$$

$$\text{Var}(X) = \frac{(a-b)^2}{12}$$

Dominio: $[a,b]$

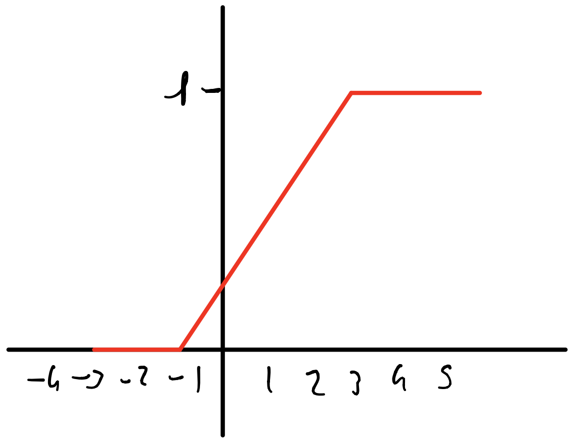
Densità:



$a = -1$

$b = 3$

F. Ripartizione



$a = -1 \quad b = 3$

Attenzione: Densità uniforme su $[a,b] \neq 0$ ha tutti i momenti

Densità Esponenziale e Funzione di ripartizione

↳ Parametro $\lambda > 0$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases} \qquad F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{se } x \leq 0 \\ 1 - e^{-\lambda x} & \text{se } x > 0 \end{cases}$$

$$E[X] = \frac{1}{\lambda} \qquad \text{Dominio: } \mathbb{R}^+$$

$$Var(X) = \frac{1}{\lambda^2}$$

Attenzione:

1) Esponenziale di $\lambda = 1$ ha tutti i momenti

2) X e Y esponenziali di parametro λ sono indipendenti

3) La v.a esponenziale non ha memoria $P(X > t+s \mid X > t) = P(X > s)$

Densità

F. Partizione

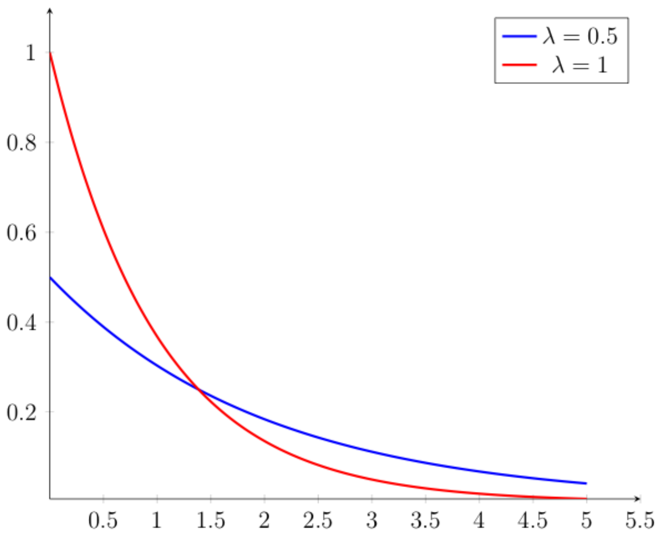
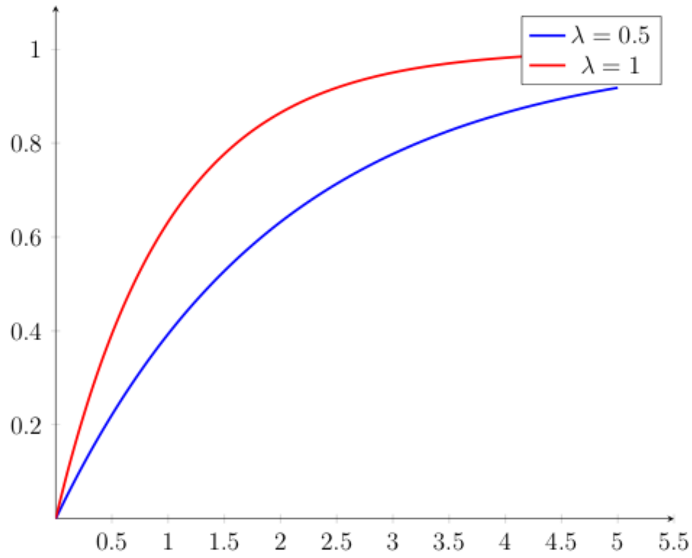


Figura 5.5: Densità di probabilità esponenziale



Densità di cauchy

$$f(x) = \frac{1}{\pi(1+x^2)}$$

Attenzione: <u>Non ha valore atteso</u>

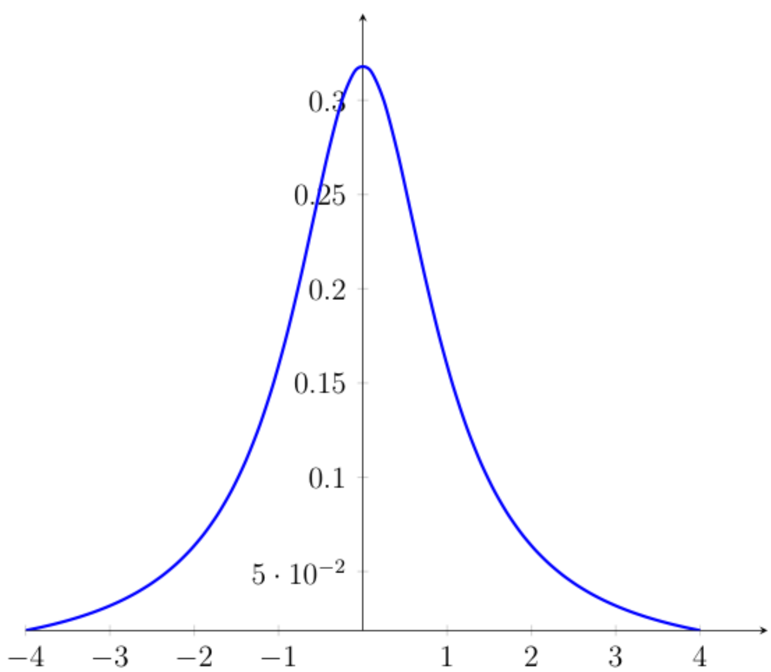


Figura 5.7: Densità di probabilità di Cauchy

Attenzione:
X con densità di cauchy e Y uniforme su [0,1] sono indipendenti

Densità Gamma

Gamma (n, λ)

Dominio: $\mathbb{R}^+$

$n > 0$
$\lambda > 0$

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{\Gamma(n)} \cdot \lambda^n x^{n-1} e^{-\lambda x} & x > 0 \end{cases}$$

Dove $\Gamma(n)$ è la funzione gamma di Eulero:

$$\Gamma(n) = \int_0^{+\infty} x^{n-1} e^{-x} \, dx \qquad \text{per } n > 0$$

$$\Gamma(n) = \begin{cases} \Gamma(n) = 1 & \text{se } n = 1 \\ \Gamma(n) = (n-1)! & \text{se } n > 1 \end{cases}$$

Proprietà:

1) Per ogni $n > 0$ si ha che $\Gamma(n) = (n-1) \cdot \Gamma(n-1)$

$$\Downarrow$$

$$\Gamma(n) = (n-1)!$$

2) Esponenziale di parametro $\lambda \Rightarrow Gamma(1, \lambda)$

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

$$E[X^n] = \frac{\Gamma(r+n)}{\Gamma(r) \cdot \lambda^n}$$

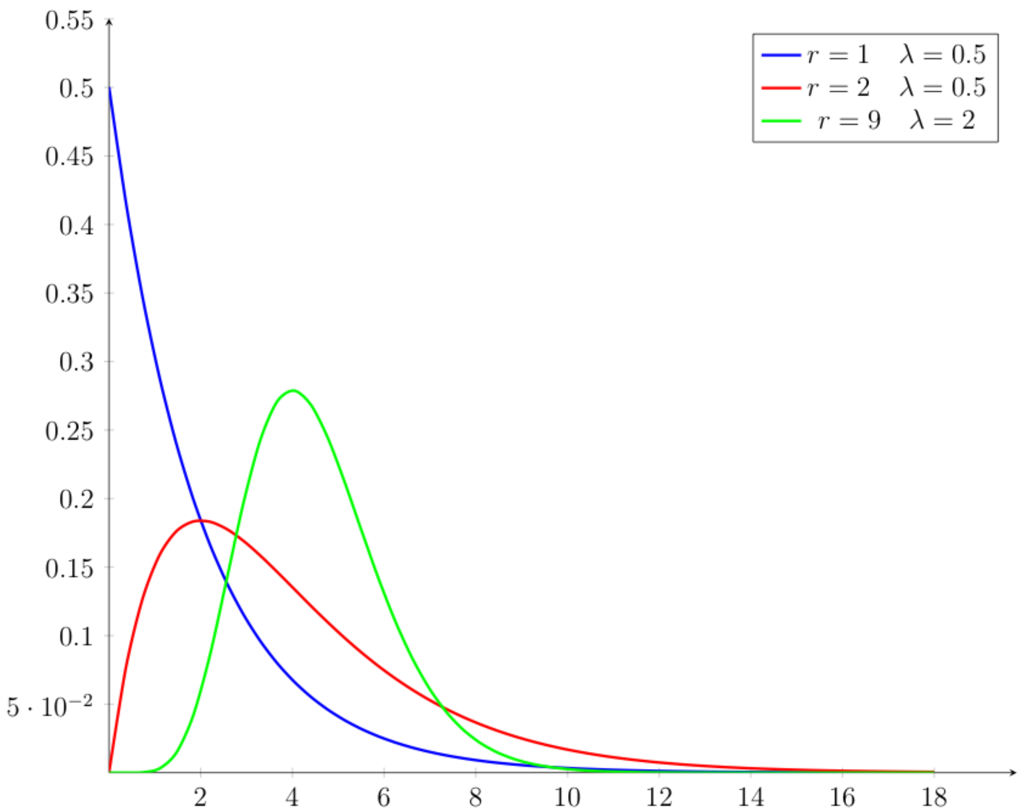


Figura 5.8: Densità di probabilità Γ

Osservazione:

Se $\left.\begin{array}{l} X \sim \text{Gamma}(1, \lambda) \\ Y \sim \text{gamma}(1, \lambda) \end{array}\right\rangle$ Indipendenti

$$X + Y \text{ è gamma}(2, \lambda)$$

$$\Updownarrow$$

$$(X+Y) \sim \Gamma(n_x + n_y, \lambda)$$

Densità Gaussiana $\rightarrow$ Dominio: $\mathbb{R}$

Se v.a var aleatoria $Y = \sigma X + m$ con

X var. aleatoria $X \sim N(0,1)$ e

Densita Gaussiana $\sim N(m, \sigma^2)$ $\begin{cases} m = E[Y] \\ \sigma^2 = Var(Y) \end{cases}$

$$f_Y(y) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(y-m)^2}{2\sigma^2}}$$

Proprietà:

1) Se $X \sim N(m, \sigma^2) \Rightarrow Y = \frac{X - m}{\sigma} \sim N(0,1)$

2) Se $X \sim N(0,1) \Rightarrow Y = \sigma X + m \sim N(m,\sigma^2)$

2) Posso sempre passare da una v.a $N(m,\sigma^2)$ nella forma $\sigma X + m$ con $X \sim N(0,1)$

Densità Gaussiana/normale standard $\rightarrow N(0,1)$

$E[X]=0$ $\quad Var(X)=1$

$$f(x) = \frac{1}{2\pi} e^{-\frac{x^2}{2}}$$

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

Funzione di ripartizione

Proprietà funzione di ripartizione

1) $\phi(-x) = 1 - \phi(x)$

2) X ha tutti i momenti

Gaussiana standard

se esponente pari

$$E[x^n] = (n-1) \cdot E[x^{n-2}]$$

$$E[x^n] = 0$$

se esponente dispari

Esempio: $- E[X] = 0$

$- E[X^2] = 1$

$- E[X^4] = 3$

$- E[X^6] = 15$

Somma di v.a con densità gaussiana

Se ho $X \sim N(m_1, \sigma_1^2)$ e $Y \sim N(m_2, \sigma_2^2)$

allora $Z = X + Y \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

Somma di v.a con densità Gaussiana standard

Supponiamo X e Y gaussiane standard indipendenti

$Z = X + Y \sim N(0,2)$ cioè ha densità

$$g(z) = \frac{1}{2\tilde{\pi}} e^{-\frac{z^4}{4}}$$

Teorema limite centrale (quando viene chiesta una valutazione approssimata)

Sia $X_1, \dots, X_n$ una succ di v.a indipendenti,

equidistribuite con $E[X_i] = \mu$ e $Var(X_i) = \sigma^2 > 0$

$$\lim_{n\to\infty} P\left\{ a \le \frac{X_1+\dots+X_n - n\mu}{\sigma\sqrt{n}} \right\} = \phi(b) - \phi(a)$$

Conclusione pratica

Se n è grande $(n \ge 50)$ abbiamo che

$$\frac{X_1+\dots+X_n - n\mu}{\sigma\sqrt{n}}$$ è approssimativamente una gaussiana standard

Dove $\mu = E[X_i]$ e $\sigma = Var(X_i)$

Esempi:

1) X binomiale$(195, 0.9)$ $\quad P(X > 180)$?

$$\frac{X - 195\cdot 0{,}9}{\sqrt{195\cdot 0{,}9\cdot 0{,}1}}$$ → App. gauss. standard

$$\frac{X - 175{,}5}{4{,}18}$$

$$P(X > 180) = P\left(\frac{X-175{,}5}{4{,}18} > \underbrace{\frac{180-175{,}5}{4{,}18}}_{1{,}07}\right)$$

$$\simeq P(z > 1{,}07) = 1 - \phi(1{,}07) = 1 - 0{,}857$$

Attenzione:

se X binomiale

$$\{X > 180\} = \{X \geq 181\}$$

$$\hookrightarrow \{X > 180{,}5\}$$ CORREZIONE DI CONTINUITÀ (è pref. questa forma)

2) $X_1, \ldots, X_{50}$ indipendenti esponenziali di parametro 2 $\quad \Gamma(1,2)$ $\quad r \swarrow \quad \hookrightarrow \lambda$

1) Stimare la $P\{X_1 + X_5 \geq 32\}$

$E[X_i] = \frac{1}{2} \qquad Var(X_i) = \frac{1}{4}$

Circa gaussiana

$$\frac{X_1 + \ldots + X_{50} - 25}{\sqrt{\frac{50}{4}}} = \frac{2(X_1 + \ldots + X_{50} - 25)}{\sqrt{50}}$$

$$P(X_1 + \ldots + X_{50} \geq 32) = P\left(\frac{X_1 + \ldots + X_{50} - 25}{\sqrt{50}} \cdot 2 \geq \frac{32-25}{\sqrt{50}} \cdot 2\right) \approx$$

1,98

$$\approx P(Z \geq 1.98) = 1 - \Phi(1.98) = 1 - 0{,}976$$

3) $X_1, \dots, X_{60}$ indipendenti, equidistribuite, uniformi su $[0,2]$

$$\overline{X} = \frac{X_1 + \dots + X_n}{n}$$

1) Stimare la $P\left(\frac{5}{6} \leq \overline{X} \leq \frac{7}{6}\right)$

2) Stimare il più piccolo x t.c. $P(\overline{X} > x) \leq 0.05$

1)

$$\frac{X_1 + \dots + X_{60} - 60 \cdot 1}{\sqrt{20}} \approx \underset{\text{Gaussiana standard}}{Z}$$

$E[X_i] = 1$

$Var(X_i) = \frac{2^2}{12} = \frac{1}{3}$

↓ perché

$\frac{(b-a)^2}{12} = Var(X)$ lim. su intervallo

$$\sigma\sqrt{n} = \frac{1}{\sqrt{3}} \cdot \sqrt{60} = \sqrt{20}$$

Finestra che non c'entra nulla con l'esercizio

X Poisson con $\lambda = n$ grande ($n \geq 80$)

come si può approssimare con T.L.C?

$$P\left(\frac{5}{6} \leq \frac{X_1 + \dots + X_{60}}{60} \leq \frac{7}{6}\right) =$$

$$= P\left(-\frac{1}{6} \leq \frac{X_1 + \dots + X_{60} - 60}{60} \leq \frac{1}{6}\right) =$$

$$= P\left(-\frac{1}{6} \cdot \frac{60}{\sqrt{20}} \leq \frac{X_1 + \dots + X_{60} - 60}{\sqrt{20}} \leq \frac{1}{6} \cdot \frac{60}{\sqrt{20}}\right)$$

$\frac{1}{6} \cdot \frac{60}{\sqrt{20}} \simeq 2{,}23$

$$\approx P(-2{,}23 \leq Z \leq 2{,}23) = \phi(2{,}23) - \phi(-2{,}23) =$$

$$= 2\phi(2{,}23) - 1 = 2(0{,}987) - 1$$

b) $P(\overline{X} \geq x) \leq 0.05 \Leftrightarrow P(\overline{X} \geq x) \approx 0.05$

$$\hookrightarrow P\left(60\frac{(\overline{X} - 1)}{\sqrt{20}} > \frac{(x-1) \cdot 60}{\sqrt{20}}\right)$$

$$P(Z > z) = 0{,}05 \Rightarrow \phi(z) = 0{,}95$$

Come faccio a trovare z tale che $\phi(z) = 0{,}95$?

$$z \approx 1{,}65 \quad \text{perchè} \quad \phi(1{,}65) \approx 0{,}95$$

$$\frac{(X-1)\cdot 60}{\sqrt{20}} = 1{,}65$$

Definizione:

Sia $X_1, \dots, X_n$ con funz di ripartizione $F_1(x), F_2(x)\dots$

e sia X limite con $F(x)$ continua

si dice che X_n con $n \geq 1$ converge in distribuzione a X se si ha $\forall x$

$$\lim_{n \to \infty} F_n(x) = F(x)$$

Esempio:

$X_1, X_2 \dots$ indipendenti, equidistribuite, $E[X_i] = \mu$ $Var(X_i) = \sigma^2$

Posto $Z_n = \dfrac{X_1 + \dots + X_n - n\mu}{\sigma \sqrt{n}}$ e Z = gaussiana standard la succ. Z_n converge in distribuzione a Z

Convergenza in probabilità

Se ho una successione $X_1, \dots, X_n$ e una v.a X

X_n converge a X in prob. se

$$\forall \varepsilon > 0 \Rightarrow P\{|X_n - X| > \varepsilon\} \mapsto 0$$

cioè

$$\lim_{n \to \infty} P\{|X_n - X| > \varepsilon\} = 0$$

Criterio convergenza:

Se $\lim_{n \to \infty} E[X_n] = c$ e $\lim_{n \to \infty} Var(X_n) = 0$

allora X_n converge a c

Legge (debole) dei grandi numeri

Se ho $X_1 \dots X_n$ indipendenti, equidistribuite,

$$E[X_i] = \mu \quad e \quad \sigma^2 = Var(X_i)$$

e $\bar{X} = \dfrac{X_1 + \dots + X_n}{n}$

allora $\bar{X}$ converge in prob. a μ

Attenzione!

Dire che X_n converge a X in prob è equivalente a dire

$(X_n - X)$ converge a 0 in prob.

cioè

$$\lim_{n \to \infty} E[X_n - X] = 0 \quad e \quad \lim_{n \to \infty} Var(X_n - X) = 0$$

$$X_n \longmapsto X$$

Serie importanti:

1) $\displaystyle\sum_{h=0}^{\infty} a^h = \frac{1}{1-a}$

2) $\displaystyle e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$