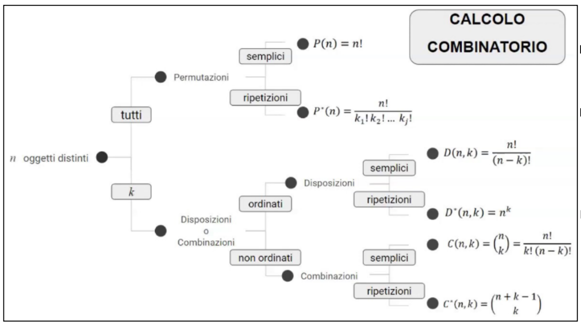


PERM SEMPLICI: modi di mescolare gli elementi (quanti modi di disporre 10 persone in fila) $10!$ — n scelte per il primo; n-1 scelte per il secondo; n-2 scelte per il 3, fino a n-(n-1)

PERM CON RIP: anagrammi della parola MAMMA (M ripetuta 3 volte e A ripetuta 2 volte) $\frac{5!}{3! \cdot 2!}$

DISP SEMPLICI: parola lunghe 2 con lettere diverse nella coppia con l'alfabeto {A, B, C, D}: $\frac{4!}{(4-2)!}$ — urna senza reinserimento — n scelte per il primo; n-1 scelte per il secondo; n-2 scelte per il 3, fino a n-k

DISP CON RIP: parola lunghe 2 con l'alfabeto {A, B, C, D}: 4^2 — urna con reinserimento

COMBINAZIONI SEMPLICI: parola lunghe 3 con lettere diverse nella coppia e tra le coppie con l'alfabeto {A, B, C, D}: $\frac{4!}{3!(4-3)!}$

spazio campionario: lancio di un dado a 6 facce Ω={1,2,3,4,5,6}
(insieme dei possibili risultati)

Assiomi della probabilità:
- **positività**: la probabilità è un valore compreso tra 0 e 1: $0 \leq \mathbb{P} \leq 1$
- **normalizzazione**: la probabilità che accada un evento dello spazio campionario è 1 (probabilità certa) $\mathbb{P}[\Omega] = 1$
- **additività**: la probabilità dell'unione di tutti gli eventi disgiunti è uguale alla somma delle singole probabilità degli eventi $\mathbb{P}\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{\infty} \mathbb{P}[A_n]$

Misura dell'incertezza
La probabilità di un evento è un numero compreso tra 0 e 1 che rappresenta con quale possibilità un evento casuale possa accadere, vicino ad 1 probabilità maggiore mentre vicino a 0 probabilità minore

probabilità dell'unione: $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

Legge della probabilità totale: La probabilità di un evento A è possibile scriverla come la somma della probabilità delle sue partizioni $\mathbb{P}[A] = \sum_i \mathbb{P}[A \cap C_i]$

Popolazioni:

	Popolazioni:	Esempio:
N	numero di elementi della popolazione	urna con 7 palline
m	elementi della popolazione con la caratteristica di interesse	4 bianche
$N-m$	elementi della popolazione senza la caratteristica	3 nere
n	numero di elementi estratti casualmente	3 palline estratte a caso
k	numero di elementi con la caratteristica tra gli n estratti	voglio 2 palline bianche
$n-k$	numero di elementi senza la caratteristica tra gli n estratti	quella rimanente sarà nera

Soluzione con reinserimento
(ci interessa l'ordine)

numero di elementi dello spazio campionario	$\#\Omega = N^n$	$\#\Omega = 7^3$
numero di modi per ottenere k elementi con la caratteristica tra gli n estratti	$\#A_k = \binom{n}{k} m^k (N-m)^{n-k}$	$\#A_2 = \binom{3}{2} 4^2 (7-4)^{3-2}$
probabilità di un modo per ottenere k elementi con la caratteristica tra gli n estratti	$\mathbb{P}[A_k] = \binom{n}{k}\left(\frac{m}{N}\right)^k\left(1-\frac{m}{N}\right)^{n-k}$	$\mathbb{P}[A_2] = \binom{3}{2}\left(\frac{4}{7}\right)^2\left(1-\frac{4}{7}\right)^{3-2}$

Soluzione senza reinserimento
(non ci interessa l'ordine)

numero di elementi dello spazio campionario	$\#\Omega = \binom{N}{n}$	$\#\Omega = \binom{7}{3}$
numero di modi per ottenere k elementi con la caratteristica tra gli n estratti	$\#A_k = \binom{m}{k}\binom{N-m}{n-k}$	$\#A_2 = \binom{4}{2}\binom{7-4}{3-2}$
probabilità di un modo per ottenere k elementi con la caratteristica tra gli n estratti	$\mathbb{P}[A_k] = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$\mathbb{P}[A_2] = \frac{\binom{4}{2}\binom{7-4}{3-2}}{\binom{7}{3}}$

probabilità composta: $\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$

proprietà della prob. composta: $\mathbb{P}[\bar{A}|B] = 1 - \mathbb{P}[A|B]$
(nota che quando si parla di eventi composti vengono scritti senza la P: A|B, C|D, ecc...)

Eventi indipendenti: $\mathbb{P}[A|B] = \mathbb{P}[A]$ $\quad \mathbb{P}[A|\bar{B}] = \mathbb{P}[A]$ $\quad \mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$

Eventi disgiunti (o incompatibili): $\mathbb{P}[A \cap B] = 0$ — due eventi disgiunti sono anche indipendenti solo se la probabilità di almeno uno dei due è 0

formula delle probabilità composte
$$P(A_1 \cap A_2 \cap A_3 \cdot \ldots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P[A_n|A_1 \cap \ldots \cap A_n - 1]$$
usata quando si vuole utilizzare le popolazioni senza reinserimento ma in cui ci interessa l'ordine

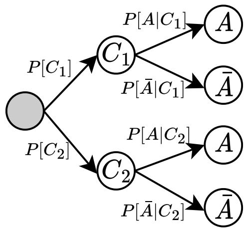
componenti indipendenti in serie
la probabilità che il sistema funzioni è dato dal prodotto delle probabilità che ogni componente funzioni oppure dal prodotto della probabilità che ogni componente non sia guasto

componenti indipendenti in parallelo
la probabilità che il sistema funzioni è dato da 1 - "la probabilità che il sistema funzioni" quindi 1 - il prodotto della probabilità che ogni componente si guasti

Formula di Bayes $\mathbb{P}[C_m|A] = \frac{\mathbb{P}[C_m \cap A]}{\mathbb{P}[A]}$ in cui

$\mathbb{P}[C_m \cap A] = \mathbb{P}[C_m] \cdot \mathbb{P}[A|C_m]$ la probabilità della strada per arrivare ad A passando dalla partizione richiesta

$\mathbb{P}[A] = \sum_i \mathbb{P}[C_i] \cdot \mathbb{P}[A|C_i]$ somma le probabilità di tutte le strade per raggiungere A



Variabile aleatoria Una variabile casuale (o aleatoria) è una funzione che associa agli elementi dello spazio campionario un valore numerico.

Funzione di ripartizione $F(x) = \mathbb{P}[X \leq x]$ La funzione di ripartizione in un dato punto x è la probabilità che la variabile casuale assuma valori minori o uguali ad x.

Proprietà:
- crescente (> o =)
- continua a destra
- a -infinito vale 0 e a +infinito vale 1

Varianza la varianza è un valore che misura quanto si discostano i valori assunti dalla variabile aleatoria rispetto al valore atteso
$\mathrm{Var}[a] = 0$
$\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X]$
$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[x]^2$

Valore atteso Il valore atteso è la media pesata dei valori che assume la variabile aleatoria. Il peso è dato dalla probabilità di ogni valore che assume la variabile.
$\mathbb{E}[a] = a$
$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$
$\mathbb{E}[X - \mathbb{E}[X]] = 0$

Costanti caratteristiche un numero associato ad una variabile aleatoria (o alla sua distribuzione) che sintetizza una informazioni di interesse sul fenomeno analizzato.

Deviazione standard (scarto quadratico medio) simile alla varianza, la deviazione standard è una misura di dispersione dei dati rispetto alla media, la differenza è che essa appartiene alla stessa scala dei valori originali (la varianza ha una scala quadratica). $\mathrm{sd}(x) = \sqrt{\mathrm{Var}(x)}$

Moda La moda di una variabile aleatoria sono i punti in cui la variabile assume i valori di massimo (assoluto ma anche locale). — la moda è un valore della X e non la sua P[X]

Quantili È un indice di posizione. Avendo la funzione di ripartizione di una variabile aleatoria, il quantile di livello \alpha è il minimo valore che assume la variabile la cui probabilità è maggiore di \alpha: $F(q_\alpha) = \mathbb{P}[X \leq q_\alpha] \geq \alpha$

percentili: sono quantili espressi in percentuale
decili: sono quantili espressi in una scala di 10

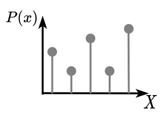
primo quartile: $\alpha = 0.25$
mediana: $\alpha = 0.5$
terzo quartile: $\alpha = 0.75$

coefficiente di variazione Il coefficiente di variazione serve per paragonare la deviazione standard rispetto alla media. più è grande più i valori sono distanti dalla media $\mathrm{cv}(x) = \frac{\mathrm{sd}(x)}{|E(x)|}$

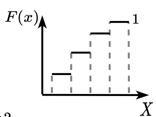
Variabile aleatoria discrete

- valori che assume la variabile si possono contare.
- La probabilità di ogni valore della variabile è compreso tra 0 e 1.
- la somma di tutte le probabilità risulta 1

Funzione di probabilità $P(x)$ probabilità che la variabile assuma un generico valore x



Funzione di ripartizione discreta Si calcola sommando le singole probabilità fino alla probabilità di x. Data la ripartizione si può trovare la probabilità di un valore facendo: $\mathbb{P}[x] = F(x) - F(x^-)$



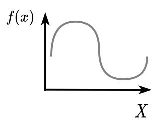
$$\mathbb{E}[X] = \sum_i (x_i \cdot p_i) \qquad \mathrm{Var}[X] = \sum_i (x_i^2 \cdot p_i) - [\mathbb{E}[X]]^2$$

Formule del cerchio
$\text{Perim.} = 2\pi r$
$\text{Area.} = \pi r^2$

Variabile aleatoria continua

- valori che assume la variabile non si possono contare.
- La probabilità di un **singolo valore** è 0 $\mathbb{P}[X = x] = 0$
- Si lavora tramite intervalli di valori

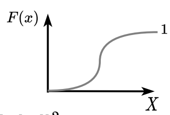
Funzione di densità $f(x)$ Andamento della probabilità al variare di x



$f(x) \geq 0$ funzione sempre positiva

$\int_{\mathbb{R}} f(x)dx = 1$ l'area totale sotto il grafico di f(x) risulta essere 1

Funzione di ripartizione continua $F(x) = \int_{-\infty}^{x} f(t)\,dt$
Data la ripartizione si può trovare la funzione di densità facendo la derivata



$$\mathbb{E}[X] = \int_{\mathbb{R}} x f(x)\,dx \qquad \mathrm{Var}[X] = \int_{\mathbb{R}} x^2 \cdot f(x)\,dx - [\mathbb{E}[X]]^2$$

Distribuzioni discrete

UNIFORME

$X \sim U\{x_1, \ldots, x_n\}$ — $\{x_1, \ldots, x_n\}$ tutti con stessa probabilità

$\mathbb{P}[X = x] = \frac{1}{n}$ $\quad \mathbb{E}[X] = \frac{x_1 + x_n}{2}$ $\quad \mathrm{Var}[X] = \frac{n^2 - 1}{12}$

IPERGEOMETRICA

$X \sim Ig(N, m, n)$

- N Numero di elementi della popolazione
- m Elementi della popolazione con la caratteristica di interesse
- n Numero di estrazioni
- k Numero di elementi con la caratteristica tra gli n estratti

$\mathbb{P}[X = k] = \frac{\binom{m}{k} \cdot \binom{N-m}{n-k}}{\binom{N}{n}}$

$\mathbb{E}[X] = n\frac{m}{N}$ $\quad \mathrm{Var}[X] = n\frac{m}{N} \cdot \frac{N-m}{N} \cdot \frac{N-n}{N-1}$

$\mathbb{P}[X = k]$ dhyper(k, m, N-m, n)
$\mathbb{P}[X \leq k]$ phyper(k, m, N-m, n)

popolazioni senza reinserimento

Distribuzioni continue

UNIFORME

$X \sim U(a, b)$ — (a, b) intervallo in cui tutti hanno la stessa probabilità

$f(x) = \frac{1}{b-a}$

$F(x) = \mathbb{P}[X \leq x] = \frac{x-a}{b-a}$ $\quad \mathbb{E}[X] = \frac{a+b}{2}$ $\quad \mathrm{Var}[X] = \frac{(b-a)^2}{12}$

$\mathbb{P}[X \leq x]$ punif(q=x, min=a, max=b)

NORMALE

$X \sim N(\mu, \sigma^2)$ — μ "parametro di posizione", rappresenta la media — σ "parametro di scala", rappresenta la deviazione standard

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ $\quad \mathbb{E}[X] = \mu$ $\quad \mathrm{Var}[X] = \sigma^2$

Standard $Z = \frac{X - \mu}{\sqrt{\sigma^2}} \sim N(0, 1)$

n grande e prob. alta: $np(1-p) \geq 10$ — varianza

Appr. Binomiale $\mathrm{Bin}(n, p) \approx N(np, np(1-p))$
correzione di continuità: aggiungere e sottrarre 0.5 per includere gli estremi

$\mathbb{P}[X > x] = p$ qnorm(p, mean=μ, sd=σ)
$\mathbb{P}[X \leq x]$ pnorm(q=x, mean=μ, sd=σ)
$\mathbb{P}[a \leq X \leq b]$ pnorm(q=b, mean=μ, sd=σ) - pnorm(q=a, mean=μ, sd=σ)

occhio a non mettere la varianza

Bernoulli $X \sim \mathrm{Ber}(p)$

p probabilità del successo

$\mathbb{P}[X=x] = p^x(1-p)^{1-x} I_{\{0,1\}}(x)$ $\quad \mathbb{E}[X]=p$

$\mathbb{P}[X=k]$ dbinom(x=k, size=1, prob=p)

$\mathbb{P}[X\le k]$ pbinom(q=k, size=1, prob=p) $\quad \mathrm{Var}[X]=p(1-p)$

Binomiale $X \sim \mathrm{Bin}(n,p)$

n numero di prove

p probabilità del successo

$\mathbb{P}[X=x] = \binom{n}{x} p^x(1-p)^{n-x}$

$\mathbb{E}[X]=np \quad \mathrm{Var}[X]=np(1-p)$

$\mathbb{P}[X=k]$ dbinom(x=k, size=n, prob=p)

$\mathbb{P}[X\le k]$ pbinom(q=k, size=n, prob=p)

popolazioni con reinserimento

Poisson $X \sim \mathrm{Po}(\lambda)$

λ Numero medio di eventi contati in unità di tempo

k Eventi contati

$\mathbb{P}[X=k] = \frac{\lambda^k}{k!}\cdot e^{-\lambda}$

$\mathbb{E}[X]=\lambda \quad \mathrm{Var}[X]=\lambda$

contare qualcosa in una unità di tempo

$\mathbb{P}[X=k]$ dpois(x=k, lambda)

$\mathbb{P}[X\le k]$ ppois(q=k, lambda)

$\mathbb{P}[X\ge k]$ 1 - ppois(q=k-1, lambda)

approssimazione della binomiale:

legge degli eventi rari

quando $n \ge 100 \wedge p \le 0.05$ allora: $\lambda = np$

Processo di Poisson

$X_t \sim \mathcal{P}(\lambda\cdot t)$

λ numero medio di eventi per unità di tempo

t intervallo di tempo custom

Le formule sono quelle di Poisson

si passa da Poisson che ha un tempo prefissato al processo tramite una proporzioni per determinare t

si può anche utilizzare l'esponenziale chiedendosi "tempo trascorso tra due eventi"

conta il numero di eventi nell'unità di tempo personalizzata

Gamma $X \sim \mathrm{Ga}(\alpha,\lambda)$

α "parametro di forma", numero di passi indipendenti

λ "parametro di frequenza", tempo di ogni passo

$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1}e^{-\lambda x} \quad \mathbb{E}[X]=\frac{\alpha}{\lambda} \quad \mathrm{Var}[X]=\frac{\alpha}{\lambda^2}$

questa distribuzione misura il tempo totale

funzione gamma: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$ se alpha è intero allora: $\Gamma(\alpha)=(\alpha-1)!$

$\mathbb{P}[X\le x]$ pgamma(q=x, shape=α, rate=λ)

$\mathbb{P}[a\le X\le b]$ pgamma(q=b, shape=α, rate=λ) - pgamma(q=a, shape=α, rate=λ)

Esponenziale $X \sim \mathrm{Exp}(\lambda)$

λ "parametro di frequenza", numero di eventi per unità di tempo

$f(x) = \lambda e^{-\lambda x} \quad \mathbb{E}[X]=\frac{1}{\lambda} \quad \mathrm{Var}[X]=\frac{1}{\lambda^2}$

ripartizione

$\mathbb{P}[X\le x] = 1-e^{-\lambda x}$

$\mathbb{P}[X> x] = e^{-x\lambda}$

mancanza di memoria $\mathbb{P}[X>t \,|\, X>t+s] = \mathbb{P}[X>s]$

$\mathbb{P}[X\le x]$ pexp(q=x, rate=λ)

$\mathbb{P}[a\le X\le b]$ pexp(q=b, rate=λ) - pexp(q=a, rate=λ)

usato per: tempo di attesa, durata di vita di componenti

Geometrica $X \sim \mathrm{Geo}(p)$

p probabilità del successo

$\mathbb{P}[X=x] = (1-p)^{x-1}p \quad \mathbb{E}[X]=\frac{1}{p} \quad \mathrm{Var}[X]=\frac{1-p}{p^2}$

$\mathbb{P}[X=k]$ dgeom(x=k-1, prob=p)

$\mathbb{P}[X\le k]$ pgeom(q=k-1, prob=p)

$\mathbb{P}[X> k]$ 1 - pgeom(q=k-1, prob=p)

numero di prove prima di ottenere il primo successo

proprietà di mancanza della memoria

$\mathbb{P}[X>m+n|X>m] = \mathbb{P}[X>n]$

$\mathbb{P}[X>k] = (1-p)^k$

Variabili congiunte discrete

$p(x,y) = \mathbb{P}[X=x, Y=y]$ —— usa la formula inversa

$p(x,y)$		X 1	2	3	4	$P_Y(y)$
Y	1	0	0.06	0.06	0.10	0.22
	2	0.10	0.10	0.04	0.04	0.28
	3	0.40	0.10	0	0	0.5
$P_X(x)$		0.5	0.26	0.1	0.14	1

marginale di Y

marginale di X

variabili indipendenti se vale: $p(x,y) = p_X(x)\cdot p_Y(y)$

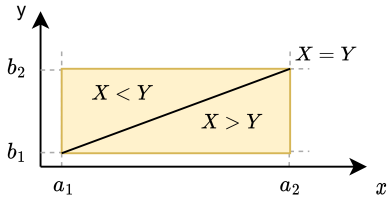
Probabilità condizionata $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)} \quad p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$

se indipendenti allora vale: $p_{X|Y}(x|y) = p_X(x) \quad p_{Y|X}(y|x) = p_Y(y)$

Valore atteso $\mathbb{E}[g(X,Y)] = \sum_y\sum_x g(x,y)p(x,y)$

Variabili congiunte continue

$f(x,y) = \mathbb{P}[a1\le X\le a2, b1\le Y\le b2]$



$\mathbb{P}[X>Y]$

$$\int_{a_1}^{a_2}\int_{b_1}^{x} f(x,y)\,dy\,dx$$

Variabili congiuntamente continue se:

$f(x,y)\ge 0 \quad \forall (x,y)\in\mathbb{R}^2 \qquad \int_{a_1}^{a_2}\left(\int_{b_1}^{b_2} f(x,y)dx\right)dy = 1$

funzione di ripartizione

$$F(x,y) = \mathbb{P}[X\le x, Y\le y] = \int_{-\infty}^{y}\int_{\infty}^{x} f(s,t)\,ds\,dt$$

probabilità marginali

marg. di X $\quad f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy$ occhio!

marg. di Y $\quad f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx$ occhio!

variabili indipendenti se scelta una x, cambiano i valori che può assumere la y, oppure più formalmente se vale $f(x,y) = f_X(x)\cdot f_Y(y)$

Probabilità condizionata $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} \quad f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$

$\mathbb{P}[X\in A|Y=y] = \int_A f_{X|Y}(x|y)\,dx \quad \mathbb{P}[Y\in B|X=x] = \int_B f_{Y|X}(y|x)\,dy$

se indipendenti allora vale: $f_{X|Y}(x|y) = f_X(x) \quad f_{Y|X}(y|x) = f_Y(y)$

Valore atteso $\mathbb{E}[g(X,Y)] = \int\int_{\mathbb{R}^2} g(x,y)f(x,y)\,dxdy$

$\mathbb{E}[X+Y] = \mathbb{E}[X]+\mathbb{E}[Y]$ se indipendenti: $\mathbb{E}[X\cdot Y] = \mathbb{E}[X]\cdot\mathbb{E}[Y]$

Varianza della somma

$\mathrm{Var}[X+Y] = \mathrm{Var}[X]+\mathrm{Var}[Y]+2\mathrm{Cov}[X,Y]$

se indipendenti $\mathrm{Var}[X+Y] = \mathrm{Var}[X]+\mathrm{Var}[Y]$

Covarianza $\mathrm{Cov}[X,Y] = \mathbb{E}[XY]-\mathbb{E}[X]\cdot\mathbb{E}[Y]$

proprietà covarianza

$\mathrm{Cov}[X,Y] = \mathrm{Cov}[Y,X]$

$\mathrm{Cov}[X,X] = \mathrm{Var}[X]$

$\mathrm{Cov}[aX,Y] = a\,\mathrm{Cov}[X,Y]$

$\mathrm{Cov}[X,a] = 0$

$\mathrm{Cov}\left[\sum_i X_i, \sum_j Y_j\right] = \sum_i\sum_j \mathrm{Cov}[X_i,Y_j]$

Se indipendenti allora: $\mathrm{Cov}[X,Y]=0$

Media campionaria

$\bar{X} = \frac{X_1+\ldots+X_n}{n}$

$\mathbb{E}[\bar{X}] = \mu \quad \mathrm{Var}[\bar{X}] = \frac{\sigma^2}{n}$

Correlazione misura di quanto sono dipendenti tra loro le due variabili

$\mathrm{Cor}[X,Y] = \rho_{X,Y} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\cdot\mathrm{Var}[Y]}}$

$\mathrm{Cor}[X,Y]\to 0$ quasi indipendenti

$\mathrm{Cor}[X,Y]\to 1$ molto dipendenti (all'aumentare di una aumenta l'altra)

$\mathrm{Cor}[X,Y]\to -1$ molto dipendenti (all'aumentare di una, decrementa l'altra)

Teorema del limite centrale

considerando un sufficiente numero di variabili **i.i.d.**, indipendentemente dalla loro distribuzione, la distribuzione della media campionaria si può approssimare alla distribuzione normale standard

utilizzabile quando viene fornito: quantità di elementi, media e ds/varianza

$\bar{X}_n \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$

$Z = \frac{\bar{X}_n-\mu}{\sqrt{\frac{\sigma^2}{n}}}$

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

somme di distribuzioni note

$\mathrm{Bin}(1,p)+\mathrm{Bin}(1,p)$ indipendenti $\to \mathrm{Bin}(2,p)$

$\mathrm{Po}(\lambda_1)+\mathrm{Po}(\lambda_2)$ indipendenti $\to \mathrm{Po}(\lambda_1+\lambda_2)$

$\mathrm{Exp}(\lambda)+\mathrm{Exp}(\lambda)$ indipendenti $\to \mathrm{Ga}(2,\lambda)$

$\mathrm{N}(\mu_1,\sigma_1^2)+\mathrm{N}(\mu_2,\sigma_2^2)$ indipendenti $\to \mathrm{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

Disuguaglianza di Chebyshev

$\mathbb{P}[|X-\mathbb{E}[X]|>\epsilon] \le \frac{\mathrm{Var}[X]}{\epsilon^2} \quad \mathbb{P}[|X-\mathbb{E}[X]|<\epsilon] \ge 1-\frac{\mathrm{Var}[X]}{\epsilon^2}$

determina un limite superiore o inferiore alla probabilità che la variabile assuma valori che si discostano di *epsilon* dalla media

Legge dei grandi numeri più variabili si considerano più la media campionaria si avvicina alla media della distribuzione (\mu).

distribuzione marginale dopo n passi $\pi^{(n)} = \pi^{(0)}\cdot P^n$

matrice di transizione

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM}\end{pmatrix}$$

ogni p deve essere compresa tra 0 e 1

ogni riga deve sommare 1

Catena di Markov Le catene di Markov sono un modello che descrive una sequenza di eventi. In particolare, la probabilità che accada un evento è determinata solo dall'evento precedente.

$X(n)$ oppure $\{X_n\}$

Probabilità congiunta $P[X_0=i, X_1=j, X_2=k, X_3=t] = \pi_i^{(0)}\cdot p_{ij}\cdot p_{jk}\cdot p_{kt}$

catena regolare Una catena di Markov si dice regolare se ogni probabilità nella matrice di transizione a n passi risulta essere strettamente maggiore di 0.

per verificarlo fare il diagramma a stati e verificare che da ogni stato si possa raggiungere ogni altro stato (anche indirettamente)

Passeggiata aleatoria Una passeggiata aleatoria è una catena di Markov con spazio degli stati S = {numeri positivi e negativi} che ad ogni istante si muove di un passo a destra o a sinistra con probabilità rispettivamente p e 1 − p:

P^n per sapere la **probabilità condizionata** facendo n salti

Distribuzione stazionaria Se la catena è regolare, allora moltiplicando la matrice per se stessa un numero tendente ad infinito di volte si ottiene sempre la stessa distribuzione stazionaria π per ogni riga della matrice

sistema in cui: $\begin{cases}\pi\cdot P = \pi \\ \text{somma degli elementi della distribuzione} = 1\end{cases}$

Catena stazionaria se vale che la distribuzione iniziale è la distribuzione stazionaria

La **passeggiata aleatoria** con barriere **assorbenti**

si possono aggiungere alla passeggiata aleatoria due barriere assorbenti, in modo che, una volta raggiunte, il sistema non si muova più da lì.

La **passeggiata aleatoria** con barriere **non assorbenti**

Si possono aggiungere alla passeggiata aleatoria due barriere non assorbenti, in modo che, una volta raggiunte, il sistema rimbalzi allo stato precedente.