Distribuzioni di probabilità

Variabile aleatoria

Una variabile aleatoria (v.a.) è una funzione *X* che associa un numero $X(\omega)$ ad ogni risultato $\omega \in S$ di un esperimento.

Useremo lettere maiuscole *X* per indicare variabili aleatorie e lettere minuscole *x* per indicare i valori che una v.a. può assumere.

Eventi generati da una v.a.

Consideriamo l'evento:

$$\{X \leq x\}$$

Questa notazione rappresenta un sottoinsieme di *S* costituito da tutti i risultati sperimentali ω tali che $X(\omega) \leq x$.

L'evento

$$\{x_1 \leq X \leq x_2\}$$

indica un sottoinsieme di S formato da tutti i risultati sperimentali ω tali che $x_1 \leq X(\omega) \leq x_2$.

Eventi generati da una v.a. (cont.)

L'evento:

$$\{X = x\}$$

è un sottoinsieme di S costituito da tutti i risultati sperimentali ω tali che $X(\omega) = x$.

Noi siamo interessati ad associare probabilità ad eventi $P\{X \leq x\}$, dove assumiamo che:

$$P\{X = +\infty\} = 0 \text{ e } P\{X = -\infty\} = 0.$$

Funzione di distribuzione

Gli elementi di S che sono contenuti nell'evento $\{X \leq x\}$ cambiano al variare di x. Quindi la probabilità $P\{X \leq x\}$ dell'evento $\{X \leq x\}$ è un numero che dipende da x.

La *funzione di distribuzione* (c.d.f.) della v.a. X è la funzione

$$F(x) = P\{X \leq x\}$$

definita per ogni x da $-\infty$ *a* $+\infty$.

Proprietà

La c.d.f. *F* di una v.a. *X* gode delle seguenti proprietà:

1. $0 \leq F(x) \leq 1$, per ogni $x \in R$;
2. $F(x)$ è una funzione non descrescente di x;
3. $F(x)$ tende a 0 per $x \to -\infty$ e tende a 1 per $x \to +\infty$;
4. $P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1)$.

Funzione densità di probabilità

La *funzione densità di probabilità* (p.d.f.) *f(x)* della v.a. *X* è la derivata di *F(x)*:

$$f(x) = \frac{dF(x)}{dx}$$

La p.d.f. è anche chiamata *funzione frequenza*.

Proprietà

La p.d.f. *f* di una v.a. *X* gode delle seguenti proprietà:

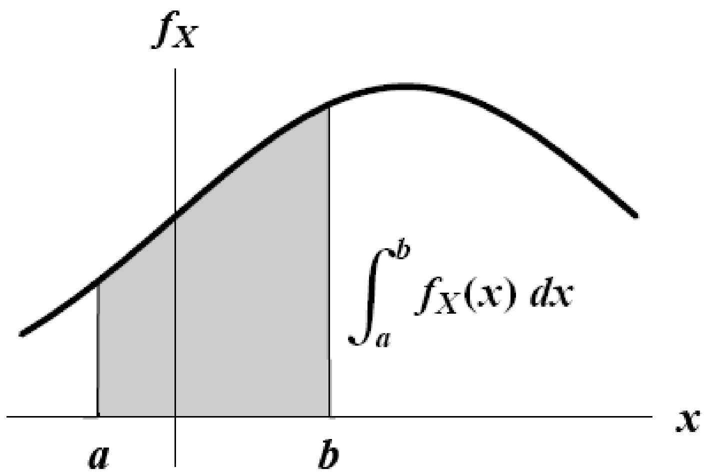
$$1)\, f(x) \geq 0, x \in R$$

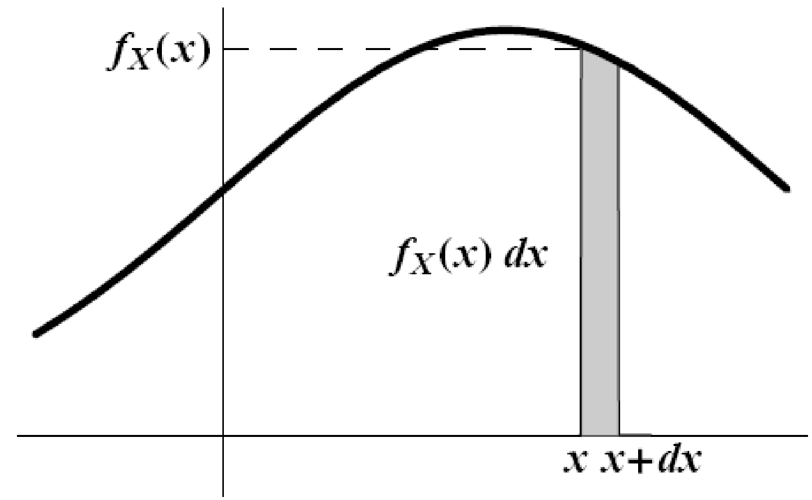
$$2)\, F(x) = \int_{-\infty}^{x} f(\tau) d\tau$$

$$3)\, \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$4)\, P\{x_1 \leq X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$$

Interpretazione





L'area tra *a* e *b* al di sotto della p.d.f. *f(x)* è la probabilità che la v.a. *X* assuma valori tra *a* e *b*. La quantità *f(x)dx* rappresenta la probabilità infinitesima che *X* stia nell'intervallo [*x*, *x*+*dx*].

Variabili aleatorie discrete

Una v.a. *X* si dice discreta se può assumere un numero finito (o infinito numerabile) di valori x_k con k=1,2,…,n (oppure $n \in N$).

Nota. Nel caso di una v.a. *X* discreta, la *f(x)* è una somma di impulsi, mentre la *F(x)* è una funzione a scala (costante a tratti).

$$f(x) = \sum_i p_i \delta(x - x_i) \quad dove \quad p_i = P\{X = x_i\}.$$

Esempio

La funzione indicatore $I_A(\omega)$ dell'evento A definita:

$$I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A \end{cases}$$

è una v.a. discreta che può assumere solo i valori 0 e 1.

Uso di farmaci, prescritti e non, in donne che hanno partorito in un ospedale negli anni 1980-1982.

Numero di farmaci	Frequenza	$P\{X = x\}$	$P\{X \leq x\}$
0	1425	0.3405	0.3405
1	1351	0.3228	0.6633
2	793	0.1895	0.8528
3	348	0.0832	0.9360
4	156	0.0373	0.9733
5	58	0.0139	0.9872
6	28	0.0067	0.9939
7	15	0.0036	0.9975
8	6	0.0014	0.9989
9	3	0.0007	0.9996
10	1	0.0002	0.9998
11	1	0.0002	1
Totale	4185	1	

Rivedere le proprietà di *f(x)* e di *F(x)* della v.a. *X* e grafici.

1) Qual è la *P* che una donna abbia usato 3 farmaci?

$P\{X = 3\} = 0.0832.$

2) Qual è la *P* che una donna abbia usato uno o due farmaci?

$P\{\{X = 1\} \cup \{X = 2\}\} = P\{X = 1\} + P\{X = 2\} =$

$0.3228 + 0.1895 = 0.5123.$

3) Qual è la *P* che una donna abbia usato un numero di farmaci minore o uguale a due?

$P\{X \leq 2\} = 0.8528.$

4) Qual è la *P* che una donna abbia usato un numero di farmaci maggiore o uguale a cinque?

$P\{X \geq 5\} = 1 - P\{X \leq 4\} = 1 - 0.9733 = 0.0267.$

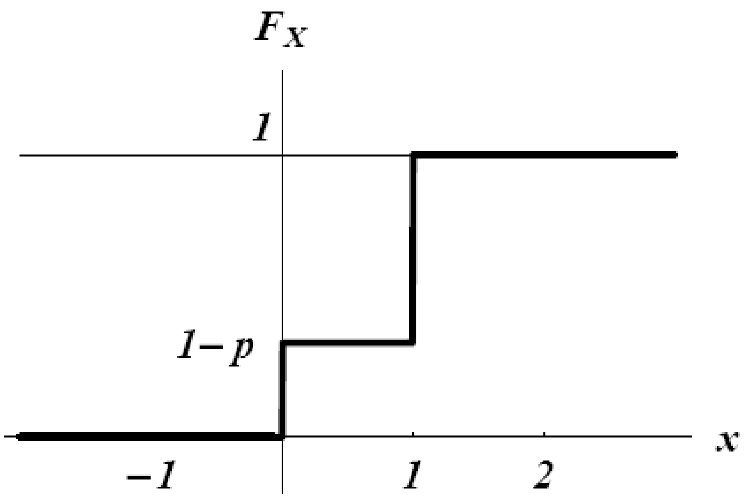
4) Qual è la *P* che una donna abbia usato un numero di farmaci compreso tra 3 e 5, estremi inclusi?

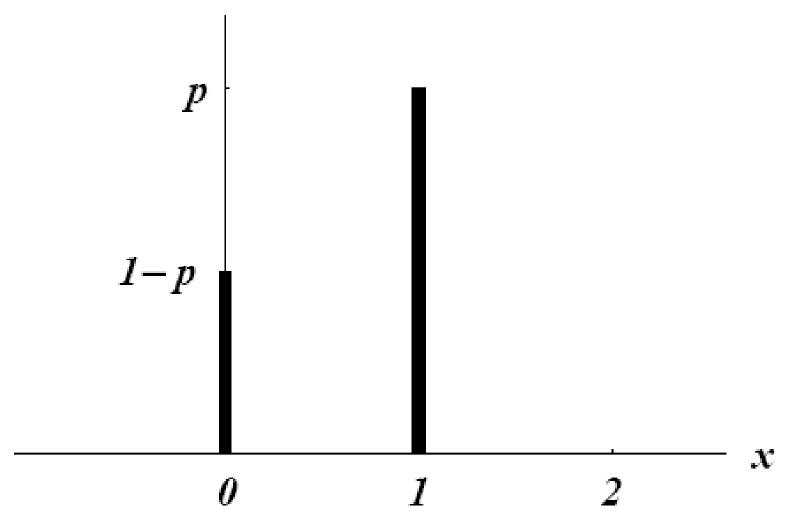
$P\{3 \leq X \leq 5\} = P\{X \leq 5\} - P\{X \leq 3\} = 0.9872 – 0.8528 = 0.1344.$

Variabile aleatoria di Bernoulli

Una v.a. X è distribuita secondo la legge di Bernoulli $B(1, p)$ quando essa assume due possibili valori: 1 con probabilità p e 0 con probabilità $1\text{-}p$, con $0 \leq p \leq 1$. Quindi:

$$P\{X = 1\} = p \quad \text{e} \quad P\{X = 0\} = 1\text{-}p$$





Distribuzione binomiale

In un esperimento l'evento A si verifica con probabilità p: $P(A) = p$ e $P(\bar{A}) = q$, $p + q = 1$.

Vogliamo determinare la probabilità $p_n(k)$ che in n ripetizioni indipendenti dell'esperimento, l'evento A si verifichi k volte in qualsiasi ordine.

Spiegazione

n ripetizioni di un esperimento possono essere rappresentate con una stringa binaria di lunghezza n (es. 00110010) con un 1 nella i-esima posizione se *A* si è verificato nella i-esima ripetizione e 0 altrimenti. Dobbiamo quindi contare il numero di stringhe di lunghezza n con k ‘1’.

A tale scopo, estraiamo a caso k posizioni della stringa ed a queste assegniamo 1, alle altre 0.

La 1ª posso estrarla in n modi diversi;

la 2ª posso estrarla in n-1 modi diversi; …

la kª posso estrarla in n-k+1 modi diversi.

Quindi $n(n-1)(n-2) \cdot \cdot \cdot (n-k+1)$ è il numero di possibili modi di estrarre k posizioni dalle n disponibili.

Poiché non siamo interessati all'ordine con cui estraiamo le posizioni, allora questo numero deve essere diviso per il numero di modi diversi in cui possiamo disporre i *k* oggetti estratti in *k* posizioni, ossia:

$$\frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} =$$

$$\frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \cdot \frac{(n-k)(n-k+1)\cdots 1}{(n-k)(n-k+1)\cdots 1} =$$

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

Il coefficiente binomiale ${}_nC_k$ conta il numero di modi diversi in cui l'evento A può verificarsi in n ripetizioni di un esperimento.

Poiché le prove sono indipendenti allora:

$$P\{A \text{ si verifica } k \text{ volte in un ordine specifico}\} = p^k q^{n-k}$$

Inoltre l'evento

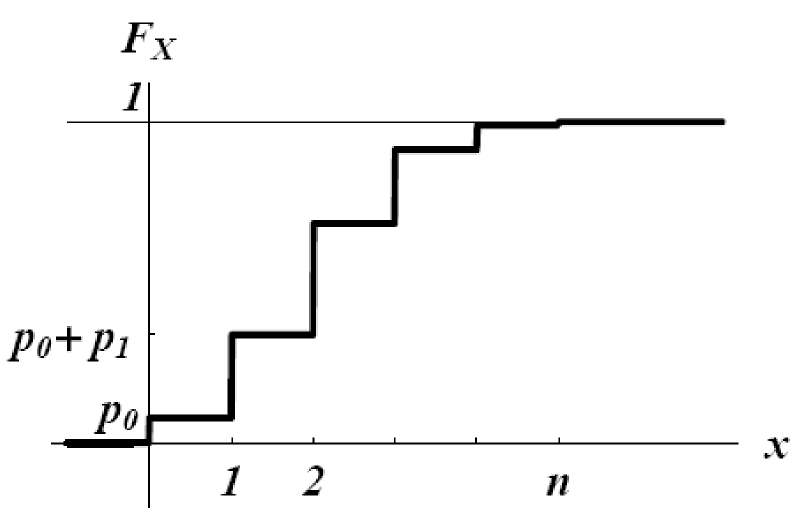
$$\{A \text{ si verifica } k \text{ volte in qualunque ordine}\}$$

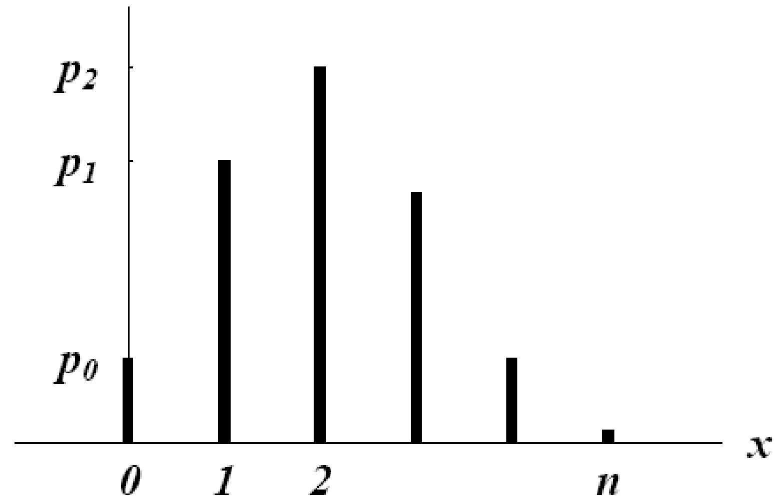
è l'unione di ${}_nC_k$ eventi

$$\{A \text{ si verifica } k \text{ volte in un ordine specifico}\}$$

e questi sono mutuamente esclusivi. Allora:

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





Quindi la v.a. *X numero di successi in n ripetizioni indipendenti di verifica di un evento A con P(A) = p* è una v.a. binomiale con parametri *n* e *p*, *B(n, p)*. Nota che:

$$p_n(k) \geq 0, \qquad k = 0,1,\ldots,n$$

$$\sum_{k=0}^{n} p_n(k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

Esempio

Il 30% di una certa popolazione è immune da una malattia. Se si estrae un campione di dimensione 10 da questa popolazione, qual è la probabilità che esso contenga esattamente 4 persone immuni?

Soluzione: p=0.3, n=10, k=4.

$$p_{10}(4) = \binom{10}{4} 0.3^4 0.7^6$$

$$= \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6!}{4! \cdot 6!} 0.0081 \cdot 0.1176 = 0.2001$$

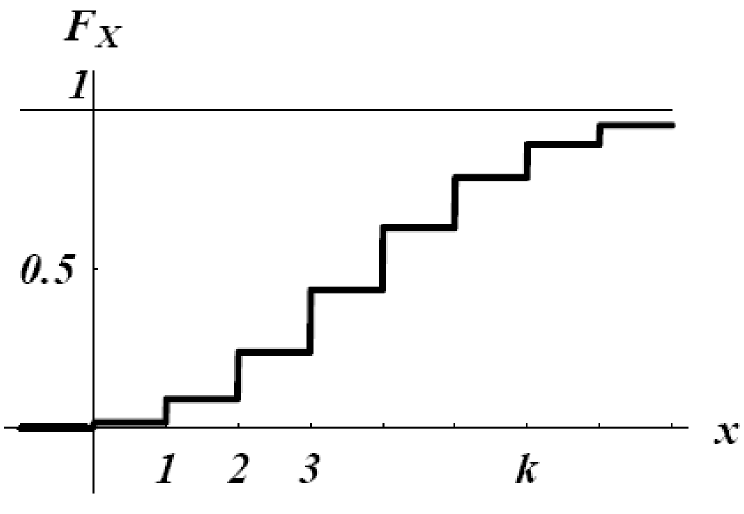
Teorema

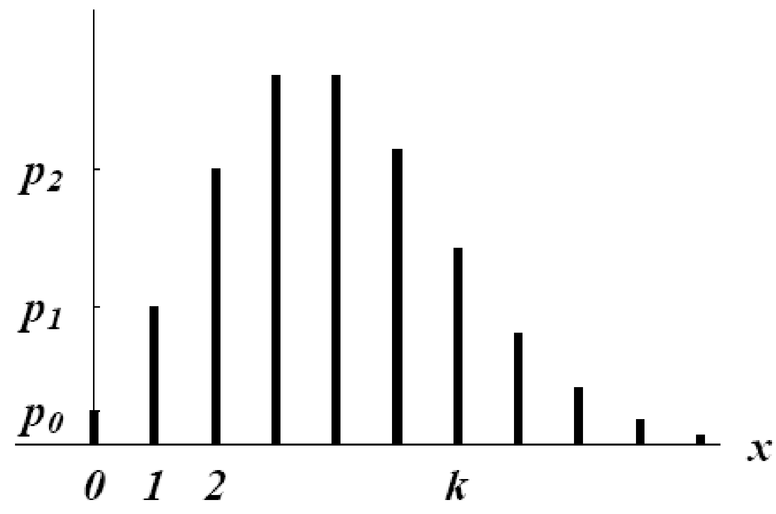
Se n v.a. $X_1, X_2, \ldots, X_n$ sono indipendenti ed identicamente distribuite secondo la legge di Bernoulli $B(1, p)$, la loro somma $X=X_1+X_2+\ldots+X_n$ è una v.a. distribuita secondo la legge Binomiale $B(n, p)$.

Distribuzione di Poisson

Una v.a. X segue la legge di Poisson $P(\lambda)$ con λ>0 quando essa assume tutti i valori interi $k \in N$ con le seguenti probabilità:

$$p_k = P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0,1,2,...$$





Osserva che $p_k \geq 0$. Inoltre, ricordando l'espansione in serie di Taylor della funzione esponenziale:

$$e^{\lambda} = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}$$

si ha:

$$1 = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{+\infty} p_k$$

Quindi p_k è una funzione densità di probabilità.

Osservazione

La legge di Poisson è particolarmente adatta a descrivere v.a. che rappresentano conteggi e che possono assumere un numero illimitato di valori: numero di telefonate che arrivano ad un centralino in un dato periodo di tempo; numero di clienti che si presentano allo sportello di un ufficio durante una giornata, ecc.

Teorema di Poisson

Consideriamo la legge Binomiale $B(n,p)$ con $p=\lambda/n$

$$p_k(n) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}, \quad k = 0,1,\ldots,n, \qquad n = 1,2,\ldots$$

Allora risulta che:

$$\lim_n p_k(n) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0,1,2,\ldots$$

Cioè $p_k(n)$ converge alla $P(\lambda)$ per ogni k.

$$p_k(n)=\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}=\frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$=\frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$=\frac{\lambda^k}{k!}\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$=\frac{\lambda^k}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}.$$

Osserva che:

$$\lim_n\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)=1,\qquad \lim_n\left(1-\frac{\lambda}{n}\right)^{-k}=1.$$

$$\left(1+\frac{\lambda}{n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\left(\frac{\lambda}{n}\right)^k = 1+n\frac{\lambda}{n}+\frac{n(n-1)}{2}\frac{\lambda^2}{n^2}+\cdots$$

$$=1+\lambda+\frac{\lambda^2}{2}-\frac{\lambda^2}{2n}+\cdots$$

Quindi si ha:

$$\lim_{n}\left(1+\frac{\lambda}{n}\right)^n = e^{\lambda}.$$

e questo dimostra il teorema.

Il teorema afferma che se *n* è molto grande e *p* è molto piccola allora una legge Binomiale *B(n, p)* è ben approssimata dalla legge di Poisson con parametro $\lambda = np$.

Interpretazione

Consideriamo la v.a. X numero di telefonate che arrivano ad un centralino telefonico in un intervallo di tempo T. Si osserva che X è una v.a. discreta che può assumere tutti i valori k=0,1,2,…

Dividiamo l'intervallo T in n parti uguali di ampiezza T/n e scegliamo n tanto grande da poter supporre che in ogni intervallo di ampiezza T/n arrivi al più una telefonata. Questa ipotesi è sempre più realistica al limite per $n \to \infty$.

Associamo all' intervallo i-mo una v.a. X_i che assume valore 1 se nell'intervallo i-mo arriva una telefonata, 0 altrimenti. Inoltre sia $P\{X_i=1\}=p$ con $p=\lambda/n$. La probabilità che arrivi una telefonata nell'i-mo intervallo diminuisce al crescere di n.

Interpretazione (cont.)

Osserva che il valore di λ è costante e dipende dalla lunghezza dell'intervallo *T* e dalla intensità delle telefonate durante questo periodo.

Ogni X_i è una v.a. Bernoulli $B(1, \lambda/n)$ e le n v.a. $X_1, X_2, \dots, X_n$ sono indipendenti. Allora, per n grande, il numero X di telefonate che arriva sarà approssimato da $X_1+X_2+\dots+X_n$ che, per il teorema, è Binomiale $B(n, \lambda/n)$.

Il teorema di Poisson ci garantisce che per $n\to\infty$ la v.a. $X_1+X_2+\dots+X_n$ e quindi X è una v.a. di Poisson.

Esempio

Nella contea di Cook (Illinois-USA) la distribuzione mensile dei suicidi tra gli adolescenti nel periodo 1977-1987 era approssimata bene da una distribuzione di Poisson con λ=2.75. Qual è la probabilità che ci siano 3 suicidi in un mese?

Soluzione:

$$P\{X=3\}=\frac{2.75^3}{3!}e^{-2.75}=\frac{20.79}{6}0.063=0.22$$

Shotgun sequencing

Supponiamo di voler sequenziare un genoma *S* costituito da *g* basi. Quindi possiamo pensare ad *S* come una stringa di lunghezza *g* di lettere dall'alfabeto {A,C,G,T}.

Metodi di sequenziamento chimico possono essere applicati a filamenti relativamente corti di DNA, da 500 a 2000 basi.

Shotgun sequencing è una strategia per sequenziare genomi di lunghezza *g* dell'ordine di 10^5, 10^6 basi. Numerose copie di *S* vengono rotte a caso in frammenti di lunghezza *L*. *N* di questi frammenti vengono scelti a caso e sequenziati, dove *N* è un parametro scelto dallo sperimentatore.

Due frammenti che si sovrappongono contengono una sottosequenza di DNA in comune, e tale sovrapposizione è utilizzata per ricostruire *S*.

Shotgun sequencing (cont.)

Supponiamo di aver accesso ad un numero elevato N di frammenti di lunghezza L ($L \ll g$) del genoma tale che:

$NL=ag$

dove a è detto covering number. Vogliamo determinare a tale che tutte le basi del genoma vengano sequenziate almeno una volta.

Indichiamo con x un sito del genoma, $x \in [1,g]$, e con $i=1,\ldots,N$ un generico frammento. Consideriamo la v.a.

$$\xi_i^x = \begin{cases} 1 \text{ se il frammento } i \text{ contiene } x \\ 0 \qquad\qquad\qquad \text{altrimenti} \end{cases}$$

Shotgun sequencing (cont.)

Poichè $P\left\{\xi_i^x = 1\right\} = \dfrac{L}{g}$, allora ξ_i^x è una Bernoulli $B\left(1, \dfrac{L}{g}\right)$.

Inoltre, essendo $\xi_1^x, \ldots, \xi_N^x$ indipendenti allora la v.a.

$$K^x = \sum_{i=1}^{N} \xi_i^x$$

è una Binomiale $B\left(N, \dfrac{L}{g}\right)$ che conta il numero di frammenti

che contengono il sito x (successi) tra gli N estratti.
Calcoliamo:

$$P\left\{K^x > 0\right\} = 1 - P\left\{K^x = 0\right\} = 1 - \binom{N}{0}\left(1 - \frac{L}{g}\right)^N \approx 1 - e^{-N\frac{L}{g}} = 1 - e^{-a}.$$

Per $a = 6, P\left\{K^x > 0\right\} \approx 0.997$.

Variabili aleatorie continue

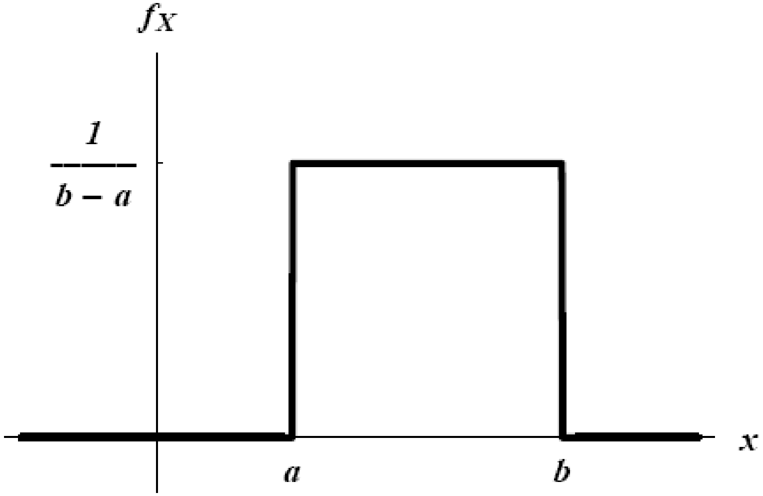
Una v.a. X si dice continua quando può assumere tutti i valori in un intervallo di numeri reali, o in tutto R.

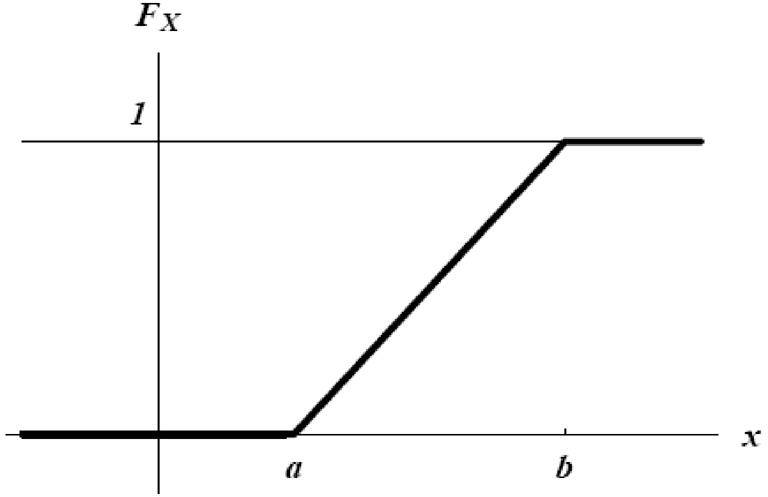
Osservazione. Se X è una v.a. continua con p.d.f. $f(x)$, allora $f(x)$ NON è la probabilità che X assuma valore x, in quanto $P\{X=x\}=0$. Inoltre f(x) può assumere valori maggiori di 1.

X uniformemente distribuita $U(a, b)$

$$f(x) = \begin{cases} 1/(b-a) & se\ a \le x \le b \\ 0 & altrimenti \end{cases}$$

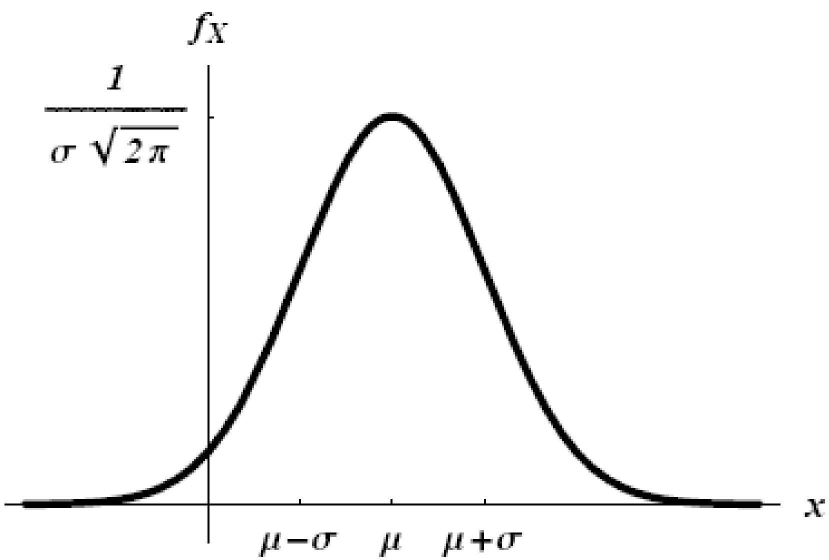
$$F(x) = \begin{cases} 0 & se\ x < a \\ (x-a)/(b-a) & se\ a \le x \le b \\ 1 & se\ x > b \end{cases}$$

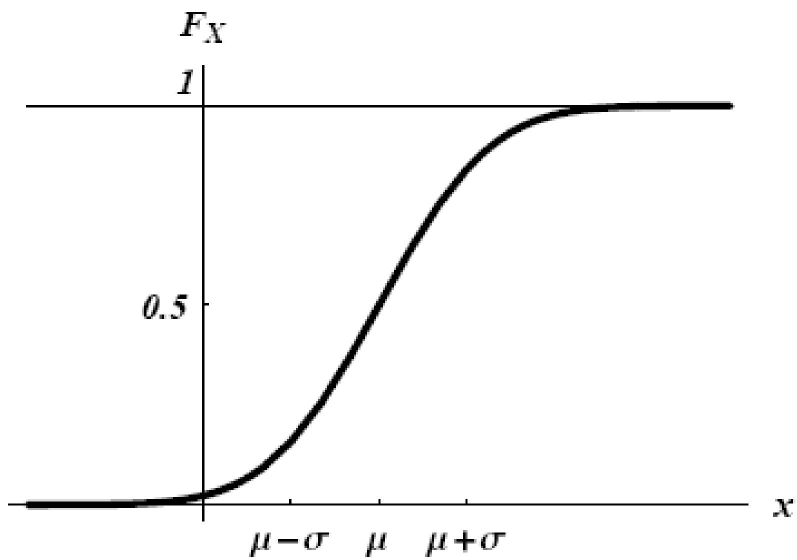




X normale o Gaussiana $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{e} \quad F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$





Proprietà

1) $f(x)>0$;

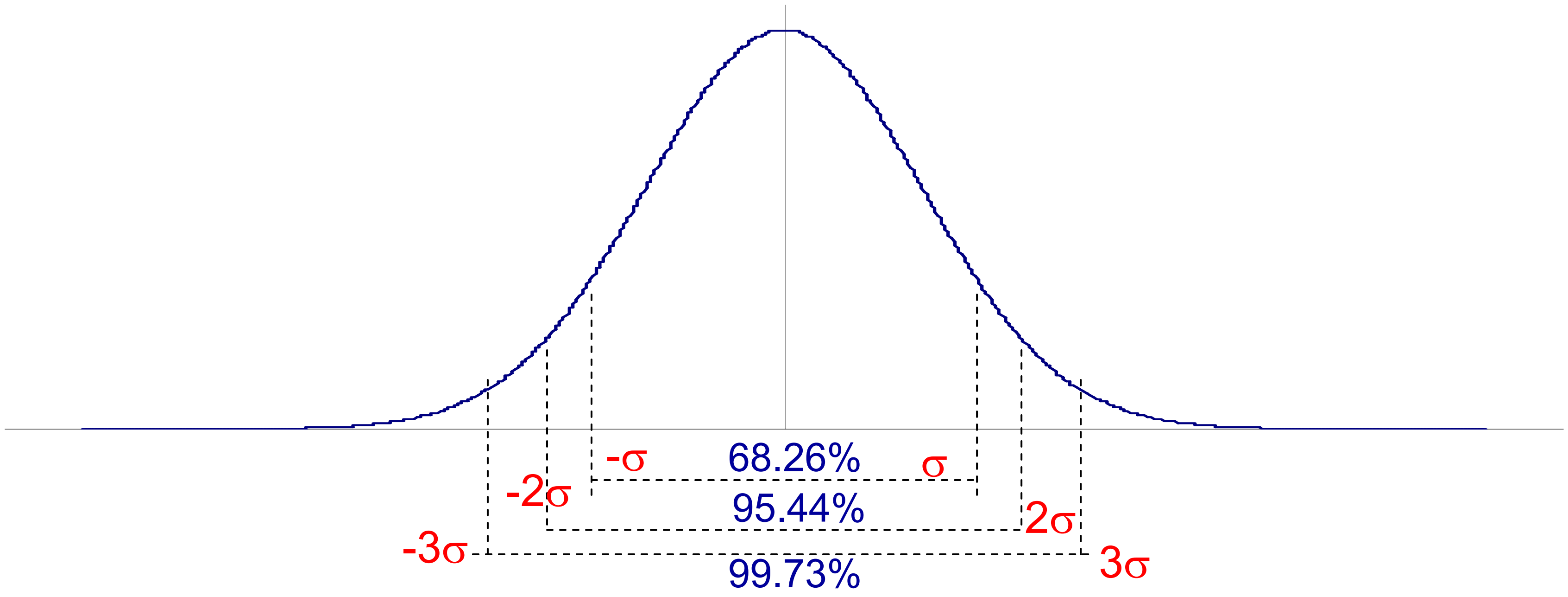
2) $f(x)$ è una curva a campana simmetrica intorno a $x=\mu$;

3) la larghezza della campana è regolata dal valore di σ;

4) $f(x)$ ha due punti di flesso in $x=\mu \pm \sigma$;

5) $F(X)$ ha una forma di S allungata con un flesso in $x=\mu$;

6) $F(X)$ diventa sempre più ripida per $\sigma \to 0$.



$$P\{\mu-\sigma \le X \le \mu+\sigma\} = F(\mu+\sigma) - F(\mu-\sigma) = 0.68$$

$$P\{\mu-2\sigma \le X \le \mu+2\sigma\} = F(\mu+2\sigma) - F(\mu-2\sigma) = 0.95$$

$$P\{\mu-3\sigma \le X \le \mu+3\sigma\} = F(\mu+3\sigma) - F(\mu-3\sigma) = 0.997$$

La distribuzione normale standard *N(0, 1)*

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{e} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

È facile verificare che:

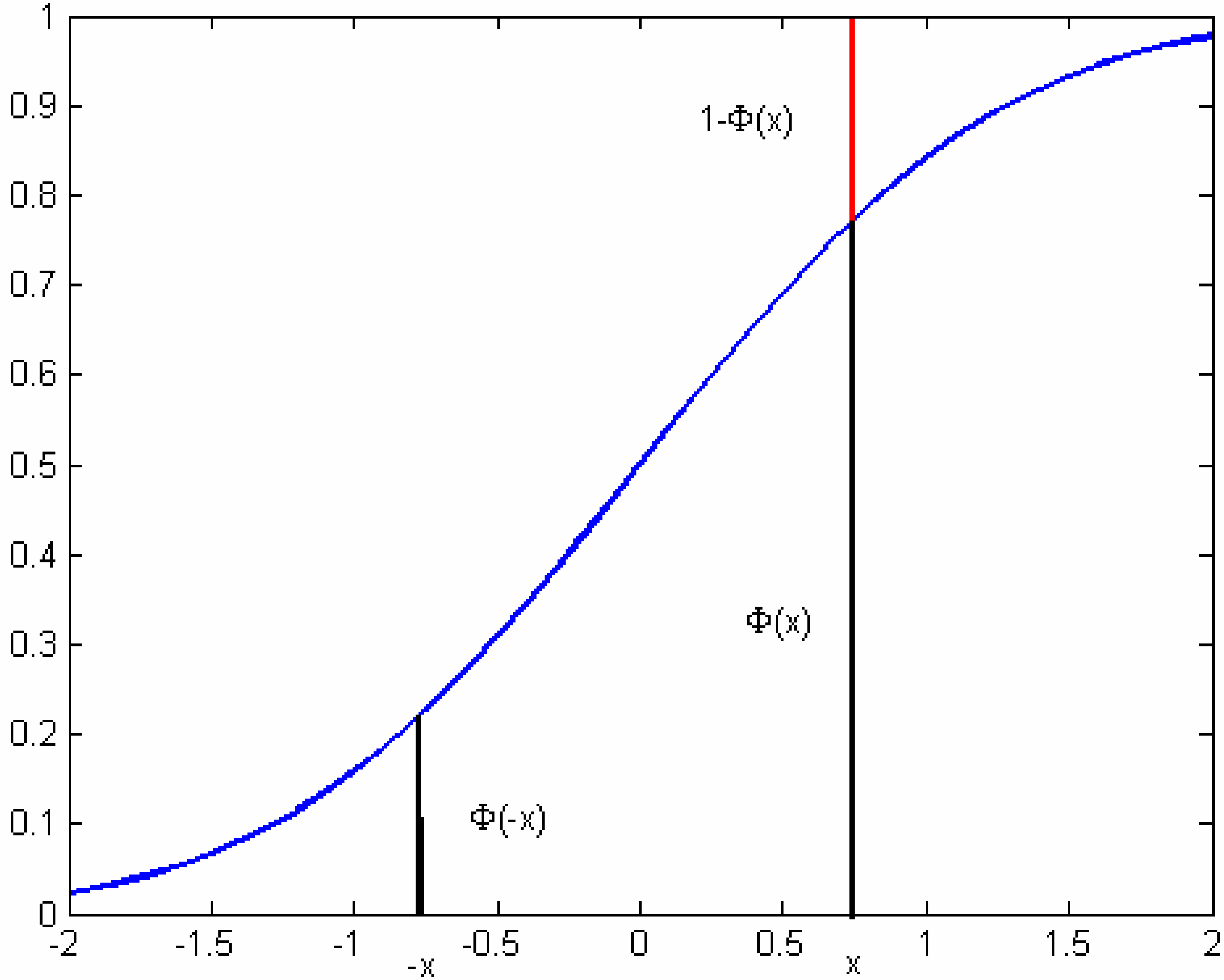
$$\varphi(-x) = \varphi(x) \quad \text{e} \quad \Phi(-x) = 1 - \Phi(x).$$

$$\Phi(-x) = \int_{-\infty}^{-x} \varphi(t) dt = 1 - \int_{-x}^{+\infty} \varphi(t) dt.$$

$$\text{Pongo } t = -y$$

$$= 1 + \int_{x}^{-\infty} \varphi(-y) dy = 1 - \int_{-\infty}^{x} \varphi(y) dy = 1 - \Phi(x).$$

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



Proprietà

Se Y è N(μ, σ^2) allora la v.a. $X = \dfrac{Y-\mu}{\sigma}$ è N(0, 1).

Viceversa, se X è N(0, 1), allora la v.a. Y = σX+μ è N(μ, σ^2).

Osservazione. Questa proprietà è utile per calcolare la probabilità di eventi di v.a. Gaussiane. Infatti:

$$P\{a \le Y \le b\} = P\{a \le \sigma X + \mu \le b\} =$$

$$P\{a-\mu \le \sigma X \le b-\mu\} = P\left\{\frac{a-\mu}{\sigma} \le X \le \frac{b-\mu}{\sigma}\right\}.$$

Quindi:

$$P\{a \le Y \le b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Esempio

Sia X una v.a. N(0, 1).

$P\{-\infty \le X \le 0.75\} = \Phi(0.75) = 0.77337.$

$P\{-2.55 \le X \le 2.55\} = \Phi(2.55) – \Phi(-2.55) = \Phi(2.55) – (1 - \Phi(2.55))$

$= 2\,\Phi(2.55) – 1 = 2 * 0.99461 – 1 = 0.98922.$

$P\{X \ge 2.71\} = 1 - P\{X \le 2.71\} = 1 – 0.99664 = 0.00336.$

Legge Normale standard $N(0, 1)$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861

Esempio

La pressione sistolica nella popolazione maschile degli USA in età [18,74] anni è approssimativamente normale:

X: N(129 mm Hg, 392.04 mm HG^2).

Qual è la percentuale di soggetti nella popolazione con una pressione sistolica maggiore di 150 mm Hg ?

Vogliamo calcolare la

$P\{X \geq 150\} = 1 - P\{X \leq 150\} =$

$1 - P\{(X - 129) / 19.8 \leq (150 - 129) / 19.8\} = 1 - \Phi(1.06) =$

$1 - 0.85543 = 0.145.$

Allora il 14.5% di soggetti di questa popolazione ha una pressione sistolica maggiore di 150 mm Hg.

Esempio

La statura di una certa popolazione di individui è approssimativamente normale, X: N(70 in, 9 in^2).

Qual è la probabilità che una persona estratta a caso da questa popolazione sia alta tra i 65 e 74 pollici ?

Vogliamo calcolare la

$P\{65 \leq X \leq 74\} =$

$P\{(65 - 70) / 3 \leq (X - 70) / 3 \leq (74 - 70) / 3\} =$

$\Phi(4/3) - \Phi(-5/3) = \Phi(4/3) - 1 + \Phi(5/3) =$

$\Phi(1.33) + \Phi(1.66) - 1 = 0.90824 + 0.95154 - 1 = 0.86.$

Allora l' 86% dei soggetti di questa popolazione ha una altezza compresa tra 65 e 74 pollici.