Dal discreto al continuo

Le distribuzioni di probabilità già viste servono a rappresentare delle caratteristiche discrete

L'insieme dei valori possibili è formato da punti isolati che possono essere "contati" cioè posti in corrispondenza biunivoca con l'insieme dei numeri naturali.

Sono inadatte per descrivere il comportamento di quantità i cui valori ricadono in un intervallo di numeri reali: Distanze,pesi, altezze, etc.

Per affrontare questi aspetti è necessario ampliare il nostro bagaglio di strumenti d'analisi

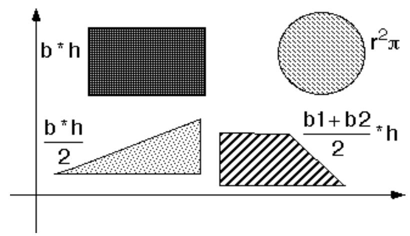
Variabili casuali continue

Le V.C. continue hanno valori in un intervallo di numeri reali (a,b).

Non è possibile osservare PUNTUALMENTE il risultato di una prova, ma solo per intervalli.

La definizione della distribuzione di probabilità delle V.C. continue è DIVERSA da quella delle V.C. discrete. Qui si parla di FUNZIONE DI DENSITA'

Nelle prossime lezioni si farà uso del concetto di area e delle tecniche di integrazione (a livello elementare, ma rigoroso)

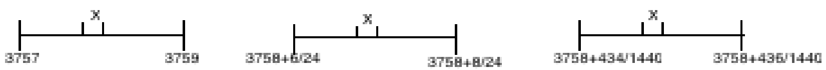


Universo degli eventi continuo

Ciò che una esperienza può provare è che il valore di una variabile continua sia il valore centrale di un intervallo di ampiezza determinata dalla precisione degli strumenti:

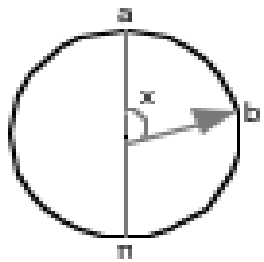
$$X_i - \frac{\Delta X_i}{2} \le X_i \le X_i + \frac{\Delta X_i}{2}$$

All'aumentare della capacità di dettaglio, gli estremi dell'intervallo saranno più vicini rimanendo però sempre distinti.



Esempio

Un esperimento consiste nel girare una freccia lungo una ruota in senso orario:



L'evento è l'intervallo E={x|a≤x≤b} formato dal punto finale "b" raggiunto dalla freccia (dopo un'impulso casuale) con il punto iniziale "a"; l'universo degli eventi è S={x|0≤x≤2π}.

Se è vera l'equiprobabilità è ragionevole assegnare:

P(E)=Arco(b-a)/2π

cioè come rapporto tra arco dell'evento a favore ed arco degli eventi possibili

Inconvenienti/1

Assegnare probabilità per intervalli è possibile, ma non ci possiamo basare sull'identità *Intervallo=Evento* perché crea problemi:

L'intersezione di una successione finita o enumerabile di intervalli è un intervallo, ma l'unione di due intervalli non sempre è un intervallo.

Siano : $a=x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ punti distinti di R. Allora:

$$\bigcup_{i=1}^{n}(x_{i-1},x_i] = (a,b] \ \ con \ \ (x_{i-1},x_i]\cap(x_i,x_{i+1}] = \varnothing; \ \ \bigcup_{i=1}^{n}[x_{i-1},x_i] = [a,b] \ \ con \ \ [x_{i-1},x_i]\cap[x_i,x_{i+1}] \neq \varnothing$$

$$\bigcup_{i=1}^{n}(x_{i-1},x_i) \neq (a,b)$$

per ottenere l'uguaglianza occorre aggiungere i punti singoli interni che non sono intervalli.

Inconvenienti/2

Tutte le σ-algebre sono incluse tra i seguenti casi estremi

$$La \ \ più \ \ povera: \ W_* = \{\varnothing, S\}; \qquad La \ \ più \ \ ricca: \ W^* = \{E | E \supseteq S\}$$

Se S è finito o enumerabilmente infinito la W^* è gestibile con i postulati di Kolmogorov (estesi agli eventi enumerabili).

Se invece S è continuo, W^* è troppo ampia per poter essere probabilizzata nel suo complesso e W_* è troppo semplice per essere utile.

Esiste una σ-algebra abbastanza grande da contenere eventi di interesse e, nello stesso tempo, tanto piccola da non creare problemi ai postulati di Kolmogorov?

La classe di Borel

Poniamo S=R cioè l'asse reale (-∞, +∞)

Quale può essere l'elemento base della σ-algebra per questo universo?

Scegliamo l'intervallo semiaperto a sinistra e chiuso a destra.

$$(a_i,b_i] = \{x \in R | a_i < x \leq b_i\}; \quad a_1 < a_2 < \dots < a_n; \ b_1 < b_2 < \dots < b_n \ -\infty \leq a_1; \ b_n < \infty$$

Qui (a,∞] è inteso come (a,∞) e il complemento di (-∞,b] è pure un intervallo semiaperto a sinistra e chiuso a destra.

A partire da questi "mattoni", costruiremo un'algebra con le operazioni di unione, intersezione e negazione di intervalli disgiunti del tipo: (,]

$$Se \ (a_i,b_i],(a_j,b_j] \in S \ \ con \ (a_i,b_i]\cap(a_j,b_j] = \varnothing \ \ per \ \ i \neq j \ \Rightarrow \ E_n = \bigcup_{i=1}^{n}(a_i,b_i] \in W$$

La classe di Borel/2

Per ottenere una σ-algebra dobbiamo aggiungere i limiti delle loro successioni monotone.

Condizione necessaria e sufficiente affinché un'algebra diventi una σ-algebra è che essa includa anche le successioni crescenti o decrescenti di eventi.

In tal caso W è la σ-algebra B minimale dell'asse reale (classe di Borel)

PROBLEMA
L'evento E={xε[0,1]|x è un irrazionale} non è un intervallo cioè non è un elemento della classe di Borel per cui il sistema assiomatico non è in grado di dare una probabilità al suo verificarsi.

Spazi di probabilità e Var. Cas. Cont.

Si parte e si arriva nello stesso tipo di dominio.

Dobbiamo però risolvere il problema di assegnare le probabilità agli intervalli secondo le indicazioni di una particolare funziione di insieme P(.);

Il calcolo integrale risponde bene a questa esigenza.

Sia S=(0,∞), B la classe di Borel di S e siano E1,E2 ε B;

$$P(E_1 \in B) = \int_{E_1} e^{-x}dx;\quad P(E_1 \in B) = \int_{E_2} e^{-x}dx;\quad P(E_1 \cup E_2) = \int_{E_1 \cup E_2} e^{-x}dx = \int_{E_1} e^{-x}dx; + \int_{E_2} e^{-x}dx;$$

E' evidente che P(E)≥0; P(S)=1 e che P(.) è σ-additiva

Considerazioni

Sia S={xεRl0≤ x ≤4} ed E={xεSlx è irrazionale}.
Ad E dobbiamo per forza assegnare probabilità uno dato che nell'intervallo vi sono sicuramente √2 e π. Però l'evento E è diverso dall'evento certo.

La probabilità che la X assuma un valore specifico X=c è:

$$P(X = c) = F(c) - \lim_{x \to c-} F(x) = F(c) - F(x-)$$

La funzione di ripartizione ha un salto se e solo se P(X=c)>0 e che, se X è continua allora P(X=c)=0.

Inoltre, si ha:

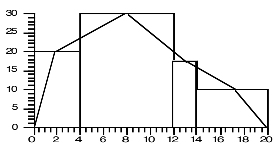
$$P\big[X \in (R - \{c\})\big] = 1$$

che ripropone la tematica di eventi con probabilità uno diversi dal certo.

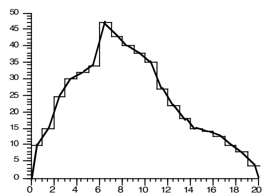
La funzione di densità

ESEMPIO:
Un campione di schede logiche di un elettrodomestico è stato raggruppate in classi annuali di durata:

Durata		Schede	nh_i
0	4	80	50.0
4	12	240	30.0
12	14	35	17.5
14	20	60	10.0
		415	



Durata		Schede	Durata		Schede
0	1	10	10	11	35
1	2	15	11	12	27
2	3	25	12	13	22
3	4	30	13	14	18
4	5	32	14	15	15
5	6	34	15	16	14
6	7	47	16	17	13
7	8	43	17	18	10
8	9	40	18	19	8
9	10	38	19	20	4
					480



Suddividiamo le classi in sottoclassi di ampiezza unitaria.

L'istogramma sarà formato da rettangoli di base uno e di altezza pari alla frequenza relativa.

inoltre, l'area sottesa al poligono coinciderà con quella totale dell'istogramma.

La funzione di densità/2

Supponiamo ora di estendere la rilevazione all'infinito, ma con classi pari ad un infinitesimo "dx" cioè punti-intervallo:

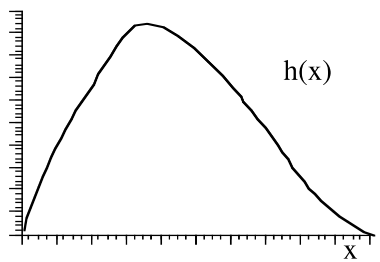
$$[x,\ x + dx]$$

ovvero l'intorno più piccolo diverso dal punto singolo.

L'istogramma non è più tracciabile e diventa simile al poligono di frequenza.

A questo punto è possibile sostituire puna curva matematica che esprima per ogni punto-intervallo la densità e racchiuda la stessa area.

Tanto più se la curva ricorda una funzione nota



Definizione formale

La funzione di densità di una V.C. continua "X" è una funzione definita su R per ogni valore della X

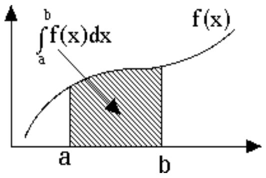
1. $f(X) \geq 0 \quad \text{se} \quad a \leq X \leq b$
2. $f(X) = 0 \quad \text{se} \quad (X < a) \cup (X > b)$
3. $\int_{-\infty}^{\infty} f(X)dx = 1$

Le 1^ deriva dal fatto che la probabilità non può essere negativa.

La 2^ estende a tutto l'asse reale il campo di variazione della V.C. L'intervallo (a,b) è detto SUPPORTO della V.C.

La 3^ deriva dal fatto che l'evento certo deve avere probabilità 1

il simbolo "$\int_a^b$" indica l'integrale della f(X) in (a,b) cioè la misura dell'area sottesa alla curva della funzione f(x) nell'intervallo (a,b)



N.B. La f(X) non dà la probabilità di X, ma è proporzionale alla probabilità che X ricada in un intervallo infinitesimale centrato su X.

Esempio

Verifichiamo che la seguente espressione descriva una funzione di densità:

$$h(x) = \begin{cases} \dfrac{3x^2 + 2x + 1}{36} & 1 \leq x \leq 3 \\ 0 & altrove \end{cases}$$

Notiamo subito che la funzione è positiva nel supporto. Controlliamo l'area sottesa:

$$\int_1^3 \frac{3x^2 + 2x + 1}{36} dx = \frac{1}{36}\left[\int_1^3 3x^2 dx + \int_1^3 2x dx + \int_1^3 1 dx\right] =$$

$$= \frac{1}{36}\left[x^3\Big|_1^3 + x^2\Big|_1^3 + x\Big|_1^3\right] = \frac{1}{36}[26 + 8 + 2] = 1$$

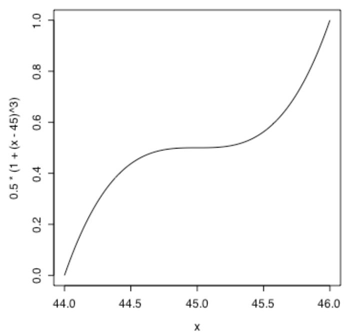
L'integrale si calcola limitatamente al supporto dato che, al di fuori di questo, il contributo della h(.) all'area è nullo.

Esempio

In una gara a cronometro si vuole battere il record della pista (31 chilometri di lunghezza) fissato a 45 minuti.
La durata del percorso è una variabile continua.

Poiché in ogni gara c'è o un po' di casualità, ogni lunghezza può essere associata alla probabilità di raggiungerla. Il modello potrebbe essere:

$$44 < x < 46; \quad F(x) = 0.5\left[1 + (x - 45)^3\right]$$



In questo esperimento i tempi ricadono in (44,46).

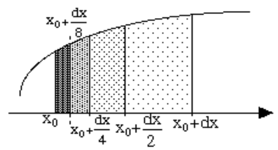
Ciò non è in contrasto con l'assegnazione di probabilità su tutto R dato che agli intervalli esterni a (44,46) è data probabilità zero.

Conseguenze della definizione

La probabilità che la V.C. continua assuma uno specifico valore è zero

$$P(X = x_0) = P(x_0 \leq X \leq x_0) = \int_{x_0}^{x_0} f(x)dx = 0$$

in quanto l'area di un segmento è zero per definizione



Da notare come il rettangolo costruito intorno a X_0 riduca la sua area sempre di più fino ad azzerarsi

Nel calcolo di probabilità per intervalli è indifferente includere o no gli estremi

$$(X \leq x_0) = (X < x_0) \cup (X = x_0) \Rightarrow P(X \leq x_0) = P(X < x_0) + P(X = x_0) = P(X < x_0)$$

questa è zero

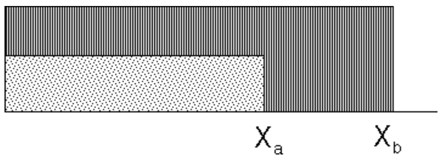
Conseguenze della definizione/2

- **La probabilità che la V.C. assuma valori maggiori di un valore dato si ricava con la probabilità complementare**

$$P[(X > x_0) \cup (X \le x_0)] = P(X > x_0) + P(X \le x_0) = 1 \Rightarrow P(X > x_0) = 1 - P(X \le x_0) = 1 - \int_{-\infty}^{x_0} f(x)dx$$

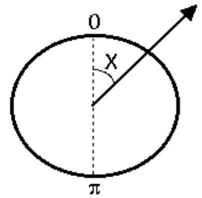
- **Se** $X_a \le X_b \Rightarrow P(X \le X_a) \le P(X \le X_b)$

Questo discende direttamente dal fatto che: $\{x | x \le X_a\} \subseteq \{x | x \le X_b\}$



Esempio

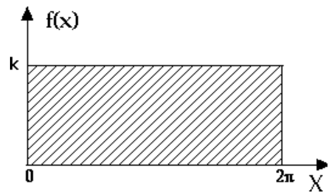
Un esperimento consiste nel girare casualmente una freccia lungo una ruota.



La variabile è l'angolo, in radianti, formato dal punto finale della freccia con il punto iniziale in "0".

Poiché nessun settore del cerchio è privilegiato ogni sottointervallo di (0,2π) ha la stessa probabilità. Questo è possibile solo con una funzione di densità costante (UNIFORME).

$$f(x) = \begin{cases} k & \text{per } 0 \le x \le 2\pi \\ 0 & \text{altrove} \end{cases}$$



La scelta di "k" segue dalla condizione che l'area ad essa sottesa è uno:

$$k * 2\pi = 1 \Rightarrow k = \frac{1}{2\pi}$$

Esercizio

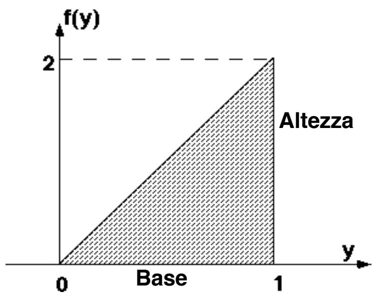
Sia Y una variabile casuale continua con funzione di densità

$$f(y) = \begin{cases} 2y & \text{per } 0 < y < 1 \\ 0 & \text{altrove} \end{cases}$$

Controlliamo che si tratti in effetti di una funzione di densità.

1) Nell'intervallo di definizione è non negativa

2) L'area sottesa alla retta è pari all'area del triangolo con base 1 ed altezza 2. Quindi: Area sottesa=1*2/2 =1



Da notare che, in base alla definizione della funzione di densità, è indifferente includere o escludere gli estremi del supporto. Si può infatti definire in modo equivalente:

$$f(y) = \begin{cases} 2y & \text{per } 0 \le y \le 1 \\ 0 & \text{altrove} \end{cases}$$

Funzione di Ripartizione

Esiste corrispondenza biunivoca tra una funzione di insieme σ-additiva: P(.) dello spazio di probabilità ed una funzione di ripartizione F

$$1.\ 0 \le F(x) \le 1,\ x \in R; \qquad 2.\ F(-\infty) = 0;\ F(\infty) = 1;$$

$$3.\ \underbrace{F(x) = F(x+)}_{\textit{continua a destra}} \qquad 4.\ F(x_1) \le F(x_2)\ \ se\ \ x_1 < x_2$$

con assegnazione della probabilità in base alla regola

$$P\{a < x \le b,\ x \in R\} = P\{(a,b]\} = P\left\{\bigcup_{i=1}^{n}(a_i, b_i]\right\} \quad a \le a_1 < b_1 \le a_2 < b_2 < \ldots < b_n \le b$$

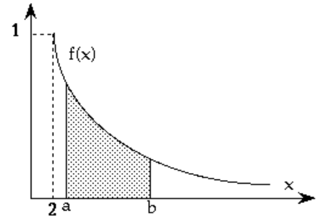
$$= \sum_{i=1}^{n}[F(b_i) - F(a_i)] = F(b) - F(a) \quad per \quad (a,b] \in B$$

Esempio

il tempo di permanenza di un prodotto sugli scaffali è descritto dalla v.c. continua X che ha funzione di densità

$$f(x) = \begin{cases} \frac{8}{x^3} & \text{per } x > 2 \\ 0 & \text{per } x \le 2 \end{cases}$$

a) Determinare la funzione di ripartizione



Questo simbolo significa: sostituisci prima x e poi 2

$$\int_2^x \frac{8}{t^3}dt = 8\int_2^x \frac{1}{t^3}dt = 8\left.\frac{t^{-3+1}}{-3+1}\right|_2^x = 8\left.\frac{t^{-2}}{-2}\right|_2^x = -\left.\frac{4}{t^2}\right|_2^x = -\frac{4}{x^2}+1$$

$$F(x) = 1 - \frac{4}{x^2}$$

b) Calcolare la probabilità che il prodotto rimanga sugli scaffali per un periodo compreso fra 2 e 5 mesi

$$P(2 \le X \le 5) = F(5) - F(2) = 1 - \frac{4}{25} - 1 + \frac{4}{4} = \frac{21}{25}$$

Esercizio

Sia "X" la V.C. che rappresenta la durata di una conversazione telefonica e si adotti come funzione di densità ESPONENZIALE

$$f(x) = e^{-x} \qquad x>0$$

Grafico della funzione di densità

La funzione di ripartizione sarà

$$F(x) = 1 - e^{-x}$$

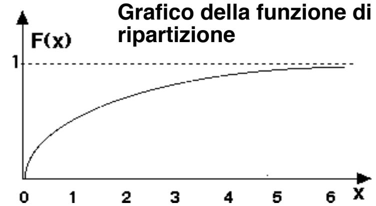


Grafico della funzione di ripartizione

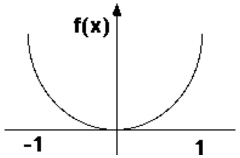
Ad esempio la probabilità che la durata sia tra 2 e 3 è:

$$P(2 \le x \le 3) = F(3) - F(2) = 1 - e^{-3} - \left(1 - e^{-2}\right)$$
$$= e^{-2} - e^{-3} = .1353 - .0498 = 0.0855$$

Esercizio

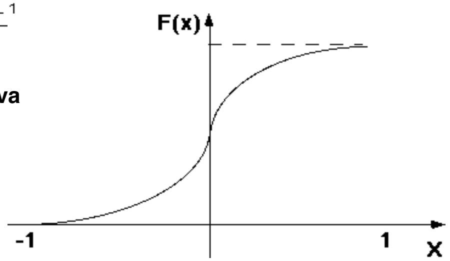
Sia X una variabile casuale continua con supporto nell'intervallo (-1,1) e con funzione di densità

$$f(x) = \begin{cases} \frac{3}{2}x^2 & \text{per } -1 \le x \le 1 \\ 0 & \text{altrove} \end{cases}$$



Costruiamo la funzione di ripartizione:

$$F(x) = \int_{-1}^{x} \frac{3}{2}t^2 dt = \frac{3}{2}\int_{-1}^{x} t^2 dt = \frac{3}{2} \cdot \left.\frac{t^3}{3}\right|_{-1}^{x} = \frac{x^3}{2} - \left(-\frac{1}{2}\right) = \frac{x^3+1}{2}$$



a) calcoliamo la probabilità che X sia negativa

$$P(X < 0) = F(0) = \frac{0^3+1}{2} = 0.5$$

b) Calcoliamo la probabilità che X ricada nell'intervallo (-0.5, 0.5)

$$P(-0.5 < X < 0.5) = F(0.5) - F(-0.5) = \frac{0.125+1}{2} - \frac{-0.125+1}{2} = 0.125$$

Var. cas. Assolutamente continue

Una funzione reale F(.) definita in R è assolutamente continua se

$$x_1 < x_2 \Rightarrow F(x_2) - F(x_1) = \int_{x_1}^{x_2} h(t)dt$$

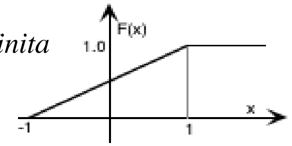
Cioè se F(.) è l'integrale definito della sua funzione di densità:

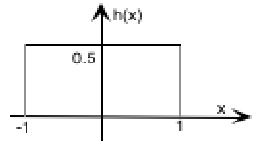
$$F(x) = \int_{-\infty}^{x} h(t)dt$$

Controesempio

$$F(x) = \begin{cases} 0 & se \;\; x \le -1 \\ \frac{x+1}{2} & se \;\; -1 < x \le 1 \\ 1 & se \;\; x > 1 \end{cases} \Rightarrow h(x) = F'(x) = \begin{cases} 0 & se \;\; x < -1, \; x > 1 \\ \frac{1}{2} & se \;\; -1 < x < 1 \end{cases}$$

La derivata non è definita Negli estremi





Tipologia della funzione di ripartizione

La funzione di ripartizione è continua su R - oppure è una funzione a gradini - o è un misto

$$F_c(x) = \alpha F_{ac} + (1-\alpha)F_s; \quad 0 \le \alpha \le 1$$

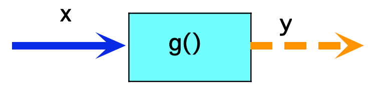
La prima componente è una funzione assolutamente continua (ne parleremo tra poco), la seconda è una funzione singolare.

Questo tipo è ostico dato che pur essendo continua ha derivata nulla in ogni punto.

Tali funzioni, tuttavia non hanno rilevanza pratica; possiamo perciò ignorarle e trattare direttamente con quelle assolutamente continue

Trasformazioni delle v.c. continue

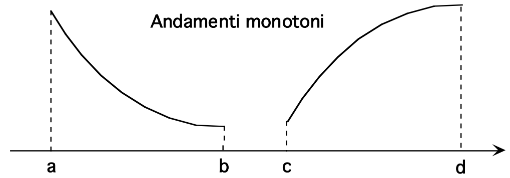
La trasformazione Y=g(X) è un sistema in cui la variabile casuale X agisce da input trasformato in output Y dal meccanismo g(.).



Resta da stabilire se è una variabile casuale , e, in caso affermativo, quale ne siano la funzione di ripartizione e di densità.

A questo fine è necessario che la trasformazione g(.) sia monotona crescente o decrescente

La g(.) non deve modificare la linea di tendenza e non deve rimanere costante nell'intervallo.



Trasformazioni /2

La monotonia stretta implica che l'inversa di g(.) esiste ed è unica: X=g^{-1}(Y).

Se g(.) è strettamente crescente l'evento {g(X)≤y) equivale a {X≤g^{-1}(y)}

$$F_y = p(Y \le y) = p[g(X) \le y] = p\left[X \le g^{-1}(y)\right] = F_x\left[g^{-1}(y)\right]$$

Se g(.) è strettamente decrescente l'evento {g(X)≤y) equivale a {X>g^{-1}(y)}

$$F_y = p(Y \le y) = p[g(X) \le y] = p\left[X > g^{-1}(y)\right] = 1 - F_x\left[g^{-1}(y)\right]$$

Pertanto, la F() della trasformata coincide con quella della variabile originaria, valutata però in un intervallo che termina o inizia nel punto immagine g^{-1}(y).

Esempi

1)

$$F(x) = \begin{cases} 0 & se \quad x \le 0 \\ x & se \quad 0 < x < 1; \\ 1 & se \quad x \ge 1 \end{cases} \qquad Y = 2x + 4 \Rightarrow \left[\begin{array}{l} x = 0 \rightarrow y = 4 \\ x = 1 \rightarrow y = 6 \end{array}\right. \Rightarrow F(y) = \begin{cases} 0 & se \quad x \le 4 \\ \dfrac{y-4}{2} & se \quad 4 < x < 6 \\ 1 & se \ge 6 \end{cases}$$

2)

$$F_x(x) = \begin{cases} 0 & se \quad x \le -1 \\ \dfrac{x+1}{2} & se \quad -1 < x < 1; \\ 1 \ \ge & se \quad y \ge 1 \end{cases} \quad \Rightarrow \quad y = x^2; \quad 0 \le y \le 1$$

gli eventi di interesse possono essere più di uno per ogni "x" in input.

$$F_y(y) = P(Y \le y) = P\left(X^2 \le y\right) = P\left(-\sqrt{y} \le X \le \sqrt{y}\right) = F_x\left(\sqrt{y}\right) - F_x\left(-\sqrt{y}\right) = \frac{1+\sqrt{y}}{2} - \frac{1-\sqrt{y}}{2} = \sqrt{y}$$

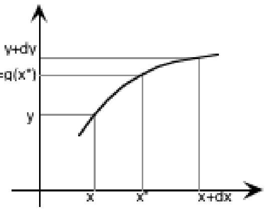
$$F_y(y) = \begin{cases} 0 & se \quad y \le 0 \\ \sqrt{y} & se \quad 0 < y < 1; \\ 1 & y \ge 1 \end{cases}$$

Trasformazioni e densità

X è una variabile casuale continua con densità h(.)
y=g(x) è una funzione strettamente monotona, differenziabile e tale da non essere mai nulla e con g'(x)>0.

Dato che y*=g(x*) si ha y≤y*≤y+dy se e solo se x≤x*≤x+dx:

$$\int_{y}^{y+dy} h_y(t)dt = \int_{x}^{x+dy} h_x(t)dt \Rightarrow h_y(y)dy \cong h_x(x)dx$$



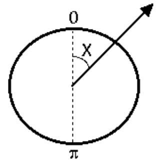
al tendere di "dx" a zero si ha

$$h_y(y) = \frac{1}{g'(x)} h_x(x)\Big|_{x=g^{-1}y} \qquad g(-\infty) < y < g(\infty)$$

Per superare la limitazione della derivata monotona crescente ed basterà ricorrere al valore assoluto

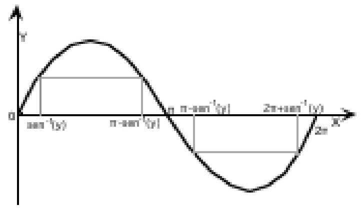
$$h_y(y) = \frac{1}{|g'(x)|} h_x(x)\Big|_{x=g^{-1}(y)} = h_x\left[g^{-1}(y)\right]\left|\frac{dg^{-1}(x)}{dx}\right|$$

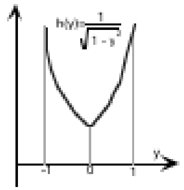
Esempio



Sia X la misura dell'angolo formato dalla freccia ruotante su di un disco con la posizione d'avvio.

il modello di densità che può descrivere l'angolo è h(x)=1/(2π) per 0≤x≤2π. Supponiamo ora di essere interessati al seno formato dall'angolo: Y=sen(X)



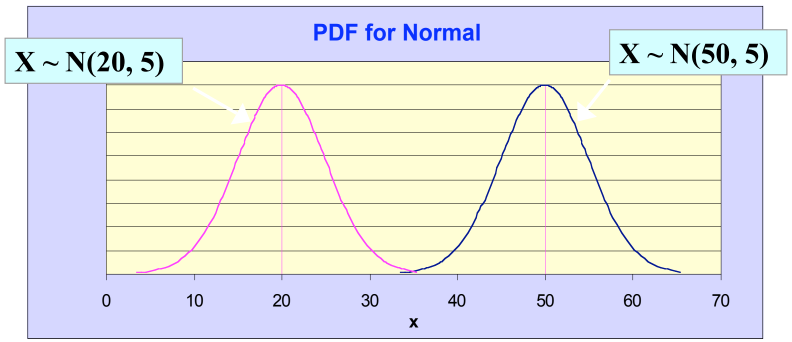


$$h_y(y) = h\left[arcosen(y)\right]\left|\frac{d\ \ arcosen(y)}{dy}\right| + h\left[\pi - arcosen(y)\right]\left|\frac{d\left[\pi - arcosen(y)\right]}{dy}\right| = \frac{1}{\pi\sqrt{1-y^2}};\quad 0 < y < 1;$$

$$h_y(y) = h\left[arcosen(y)\right]\left|\frac{d\left[\pi - arcosen(y)\right]}{dy}\right| + h\left[\pi - arcosen(y)\right]\left|\frac{d\left[2\pi + arcosen(y)\right]}{dy}\right| = \frac{1}{\pi\sqrt{1-y^2}};\quad -1 < y < 0$$

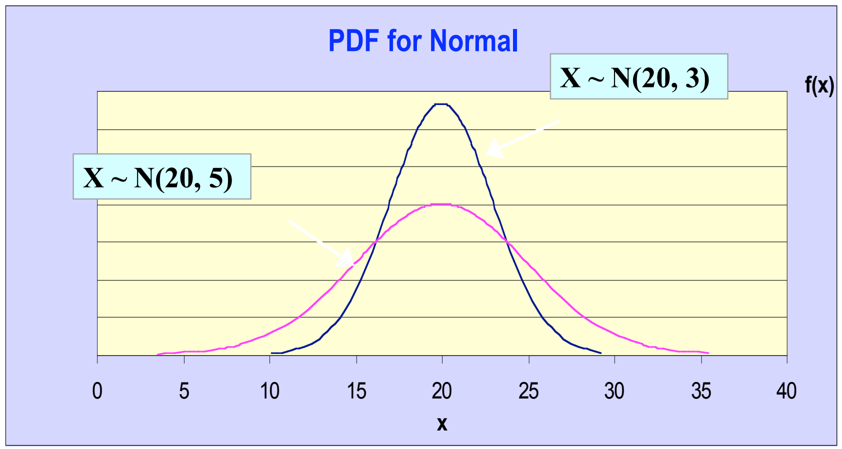
Parametri delle distribuzioni

1.	**Posizione: specificano la tendenza centrale dei valori nel modello od anche il punto di massimo addensamento**



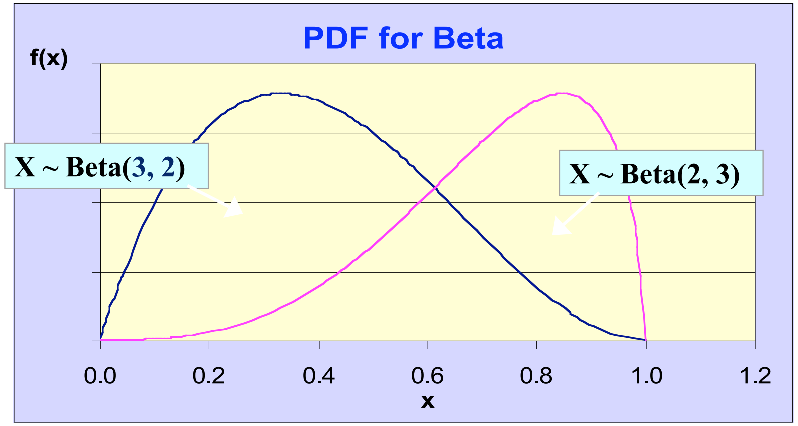
Parametri delle distribuzioni/2

2.	**Scala: determinano l'unità di misura della variabile (espansione o contrazione della scala delle ascisse)**



Parametri delle distribuzioni/3

3. Forma: Controllo la forma che assume il modello di probabilità. In particolare, l'asimmetria, l'appuntimento al centro, la pesantezza delle code



Valore atteso delle v.c. continue

La definizione del valore atteso delle v.c. continue, ricalca quella delle v.c. discrete:

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

Area sottesa alla curva formata dal prodotto della variabile per la densità

Sia X una variabile casuale con funzione di densità:

$$f(X) = \begin{cases} \dfrac{2X^2 - X}{34.5} & \text{per } 1 < X < 4 \\ 0 & \text{altrove} \end{cases}$$

Il valore atteso è

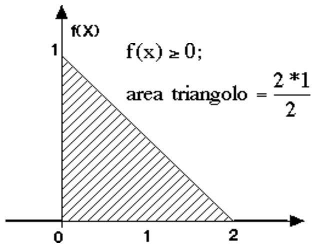
$$\mu = \int_1^4 x\left(\frac{2x^2 - x}{34.5}\right)dx = \frac{1}{34.5}\int_1^4 \left(2x^3 - x^2\right)dx = \frac{1}{34.5}\left(\frac{x^4}{2} - \frac{x^3}{3}\Bigg|_1^4\right) = 3.087$$

Esempio

Consideriamo la v.c. continua avente funzione di densità

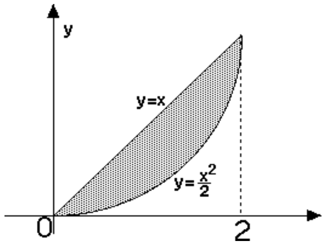
$$f(x) = \begin{cases} 1 - \dfrac{x}{2} & \text{per } 0 < x < 2 \\ 0 & \text{altrove} \end{cases}$$

$$\mu = E(X) = \int_0^2 x\left(1 - \frac{x}{2}\right)dx = \int_0^2 \left(x - \frac{x^2}{2}\right)dx$$



$f(x) \geq 0;$

$\text{area triangolo} = \dfrac{2 * 1}{2} = 1$



$$\text{Area sotto la retta } y = x \Rightarrow \frac{x * x}{2} = \frac{x^2}{2}$$

$$\text{Area sotto la parabola } y = x^2 \Rightarrow \frac{x^3}{3}$$

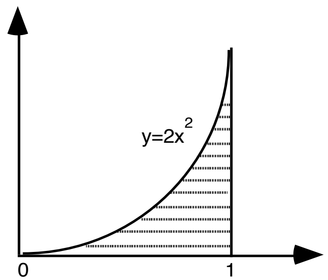
$$\text{Area sottesa} = \frac{2^2}{2} - \frac{\frac{2^3}{3}}{2} = 2 - \frac{8}{6} = \frac{2}{3}$$

Esercizio

Supponiamo che la "X" sia esprimibile dalla funzione di densità:

$$h(x) = \begin{cases} 2x & per \;\; 0 \leq x \leq 1 \\ 0 & altrove \end{cases}$$

$$\mu = \int_0^1 x 2x dx = \int_0^1 2x^2 dx = 2\frac{x^3}{3}\Bigg|_0^1 = 2\left(\frac{1^3}{3} - \frac{0^3}{3}\right) = \frac{2}{3}$$



La media aritmetica è pari all'area sottesa alla curva.

Proprietà del valore atteso

1. $E(c) = c \quad se \ \ c \ \ è \ \ una \ costante$
2. $E[cg(x)] = cE[g(x)]$
3. $E(aX + bY) = aE(X) + bE(Y)$
4. $E(X) = \int_0^{\infty}[1 - F(x)]dx - \int_{-\infty}^{0} F(x)dx$
5. $Se \ \ F_X(x) \le F_Y(x) \Rightarrow E(X) \ge E(Y)$
6. $E(X - a)^2 = minimo \ \ se \ \ a = E(X)$

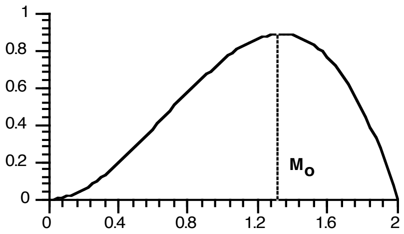
La moda

La moda -se esiste- è il valore più probabile ovvero il punto in cui la densità di probabilità raggiunge il suo massimo globale.

Se la h(x) è dotata di derivata prima per ogni punto nell'intervallo aperto [a,b], la presenza di un massimo in X_0 implicherà che $h'(X_0)=0$.

$$h(x) = \frac{3x^2(2-x)}{4}; \quad 0 < \ x \ < 2;$$

$$h'(x) = 3x - \frac{9}{4}x^2 = 0 \Rightarrow x = 0 \ \ o \ \ x = \frac{4}{3}$$



La densità aumenta, a partire dallo zero, fino ad arrivare a x=4/3. Da qui in poi la densità comincia a diminuire fino a ritornare allo zero per cui M_o= 4/3

La moda /2

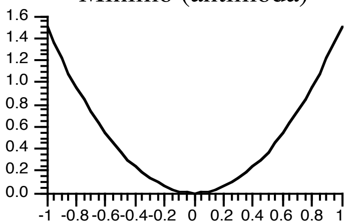
L'azzeramento della derivata prima non è necessario né sufficiente

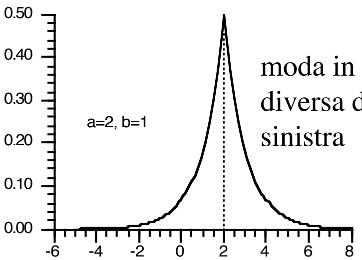
ESEMPI

$$h(x) = \frac{3}{2}x^2; \quad -1 < \ x \ < 1;$$

$$h'(x) = 3x = 0 \Rightarrow x = 0$$

Minimo (antimoda)





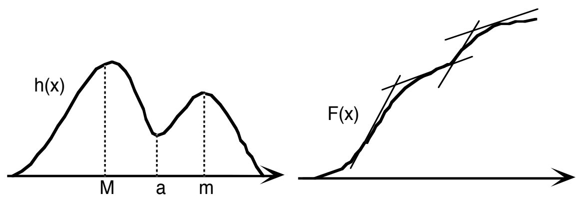
moda in un punto con diversa derivata destra e sinistra

$$h(x) = \frac{e^{-\left|\frac{x-a}{b}\right|}}{2b}; \ \ b > 0$$

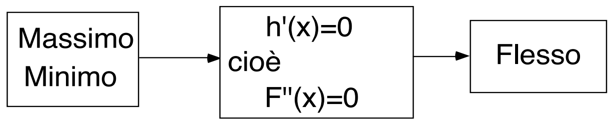
Moda nella funzione di ripartizione

La pendenza della funzione di ripartizione coincide con la densità di frequenza e la tangente alla curva di densità descrive anche la derivata seconda della funzione di ripartizione.

Densità multimodale



$$F''(x) = h'(x)$$



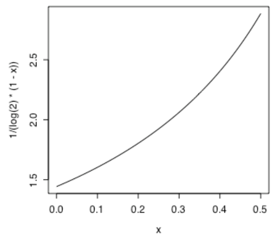
La mediana

La mediana della variabile casuale continua è la soluzione di:

$$0.5 = F\left(M_e\right) = \int_{-\infty}^{M_e} h(x)dx \quad ovvero \quad M_e = F^{-1}(0.5)$$

ESEMPIO:

Consideriamo una funzione di densità crescente in [0, 0.5] che esprime la percentuale di astensionismo potenziale in un *referendum*.

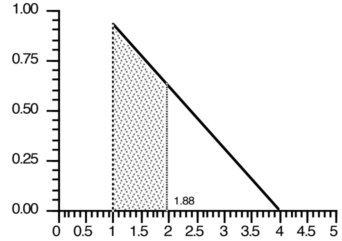


.

$$h(x) = \frac{1}{Ln(2)(1-x)}; \quad 0 < x < 0.5; \qquad F(x) = -\frac{Ln(1-x)}{Ln(2)} \Rightarrow M_e = 1 - 0.5^{0.5} = 0.2929$$

Esempio

La densità lineare fornisce un quadro teorico per un fenomeno che si presenta con densità di frequenza in progressione aritmetica inversa rispetto alle modalità.



$$h(x) = \left(\frac{4-x}{3}\right); \quad 1 < \ x \ < 4$$

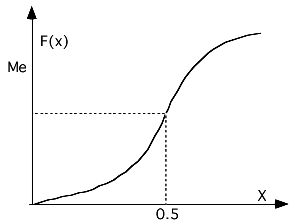
$$F(x) = \left(\frac{4-x}{3}\right)^2 = 0.5 \Rightarrow M_e = 4 - \frac{3}{\sqrt{2}} = 1.88$$

La mediana è quell'ascissa per cui passa la retta verticale che divide l'area di frequenza in due parti esattamente uguali

Esercizio

Un modello di densità molto adoperato nello studio della distribuzione dei redditi è quello di Pareto che ha funzione di ripartizione:

$$F(x) = 1 - x^{-\theta}; \ con \ \ x > \theta > 1$$



La mediana si ottiene con dei comodi passaggi algebrici:

$$0.5 = 1 - \left(M_e\right)^{-\theta} \Rightarrow \left(M_e\right)^{-\theta} = 0.5 \Rightarrow -\theta \ \ Ln\left(M_e\right) = -Ln(2)$$

$$\Rightarrow \ Ln\left(M_e\right) = \left(\frac{1}{\theta}\right) Ln(2) \Rightarrow M_e = e^{\frac{Ln(2)}{\theta}}$$

Se ad esempio θ=2 si ha M_e=1.41

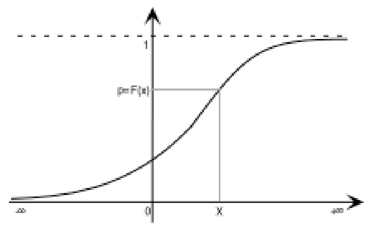
I quantili

Altre caratteristiche interessanti della variabile casuale

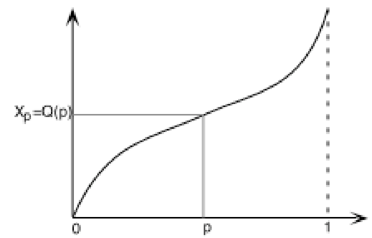
$$X_p \quad tale \quad F\left(X_p\right) = p \ \ ovvero \quad X_p = Q\left(p\right) \ \ dove \quad Q() = F^{-1}$$

La funzione quantile è l'inversa della funzione di ripartizione

E' un'altra descrizione della variabilie casuale



Distribution function



Quantile function

Esempio

Il modello di Pareto ha molte appicazioni in statistica (analisi della distribzione dei redditi)

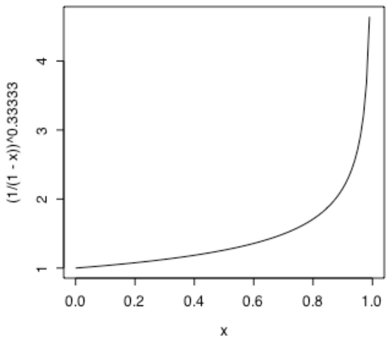
$$F(x) = 1 - x^{-\theta};\ con\ \ x > 1,\ \theta > 1$$

$$Q(p) = \left[\frac{1}{1-p}\right]^{\frac{1}{\theta}}$$

$$X_{0.75} = Q(0.75) = \left[\frac{1}{1-0.75}\right]^{\frac{1}{3}} = 1.5874$$

$$X_{0.25} = Q(0.25) = \left[\frac{1}{1-0.25}\right]^{\frac{1}{3}} = 1.10064$$

Funzione quantile



Esercizio

Sia X una variabile casuale con funzione de densità: $f(X) = \begin{cases} \dfrac{2X^2 - X}{34.5} & \text{per } 1 < X < 4 \\ 0 & \text{altrove} \end{cases}$

Il valore atteso è $\mu = \int_1^4 x\left(\frac{2x^2 - x}{34.5}\right)dx = \frac{1}{34.5}\int_1^4 \left(2x^3 - x^2\right)dx = \frac{1}{34.5}\left(\frac{x^4}{2} - \frac{x^3}{3}\bigg|_1^4\right) = 3.087$

Per il calcolo della varianza determiniamo il momento 2°

$$\sigma^2 = E\left(X^2\right) - \mu^2$$

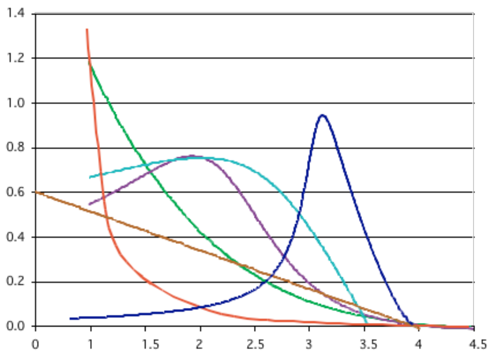
$$E\left(X^2\right) = \int_1^4 x^2\left(\frac{2x^2 - x}{34.5}\right)dx = \frac{1}{34.5}\int_1^4 \left(2x^4 - x^3\right)dx = \frac{1}{34.5}\left(\frac{2x^5}{5} - \frac{x^4}{4}\bigg|_1^4\right) = 10.013$$

$$\sigma^2 = 10.013 - 3.087^2 = 0.4834$$

Distribuzione Wakeby

Alocuni modelli sono formulati direttamente nel piano [p,Q(p)]

$$Q(p,\lambda) = \lambda_1 - \left(\lambda_2 q^{\lambda_4} + \lambda_3 q^{\lambda_5}\right),\quad 0 \le p \le 1,\ q = 1 - p$$



λ_1 parametro di centralità

λ_2, λ_3 parametri lineari di scala e forma

λ_4, λ_5 parametri di forma

Valore atteso di funzioni di v.c. continue

Sia y=g(x) una trasformazione invertibile in (a,b) della v.c. x con densità f(x)

y'=g'(x) esiste ed è sempre positiva in (a,b)

Oppure

y'=g'(x) ed è sempre negativa in (a,b)

In questo caso la densità della y è data da

$$f_Y(y) = \begin{cases} f_X\left[g^{-1}(y)\right]\left|\dfrac{dg^{-1}(y)}{dy}\right| & per\quad \alpha < y < \beta \\ 0 & altrove \end{cases}\qquad con\ \begin{cases}\alpha = \min\{g(a), g(b)\} \\ \beta = \max\{g(a), g(b)\}\end{cases}$$

Quale sarà il valore atteso di y? E' possibile conoscerlo partendo da quello della x?

Valore atteso di funzioni di v.c. continue/2

Per y=g(x) si ha

$$\mu_y = E\left[g(x)\right] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

Il cambio di variabile comporta

$$\int_{-\infty}^{\infty} g(x) f_X(x)dx = \int_{\alpha}^{\beta} y f_X\left[g^{-1}(y)\right]\left|\frac{dg^{-1}(y)}{dy}\right|dy$$

Il fattore a destra della y è la densità di questa variabile casuale. Quindi

$$\mu_y = \int_{\alpha}^{\beta} y f_Y(y)dy$$

Esempio

$$h(x) = \begin{cases} 2x & per \quad 0 \le x \le 1 \\ 0 & altrove \end{cases}; \quad \mu = \int_0^1 2x^2 dx = 2\left(\frac{x^3}{3}\right)\Bigg|_0^1 = 2\left(\frac{1}{3} - 0\right) = \frac{2}{3}$$

Calcolo del valore atteso della trasformata y=5x+3

$$y = 5x + 3; \quad x = g(y) = \frac{y-3}{5}; \; 3 < y < 8; \; \left|g'(y)\right| = \frac{1}{5}; \; h(y) = \frac{2}{5}(y-3)$$

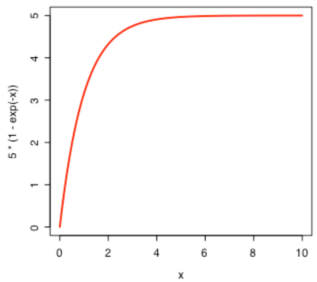
$$\mu_y = \int_3^8 \frac{y2}{25}(y-3)dy = \frac{2}{25}\frac{y^3}{3} - \frac{3y}{2}\Bigg|_3^8 = \frac{19}{3} = 5\frac{2}{3} + 3$$

Esercizio

Le entrate dall'uso di un impianto di generazione elettrica in novembre dipendono, tra l'altro, della caduta di pioggia in quel mese.

Ipotizziamo che la funzione di guadagno sia g(x)=5(1-e^{-x}) milioni dove "x" rappresenta i centimetri di pioggia caduti durante il mese.

Se la funzione di densità della X è h(x)=e^{-x} per x>0, quale sarà il guadagno atteso?



$$E\left[g(x)\right] = \int_0^{\infty} 5\left(1 - e^{-x}\right)e^{-x}dx = 5\int_0^{\infty}\left(e^{-x} - e^{-2x}\right)dx$$

$$= 5\left[1 - \frac{1}{2}\int_0^{\infty} 2e^{-2x}dx\right] = 5\left[1 - \frac{1}{2}\right] = 2.5$$

I momenti delle v.c. continue

Sia I momenti riassumono gli aspetti principali del modello di una variabile casuale: il valore atteso, lo scarto quadratico medio, l'asimmetria.

Tale capacità specificativa si è dimostrata utile sia con le distribuzioni di frequenza che per le distribuzioni di probabilità delle variabili casuali discrete.

Un ruolo analogo è loro riservato con le continue.

$$dall'origine: \quad \mu_\alpha = \int_{-\infty}^{\infty} x^\alpha h(x)dx;$$

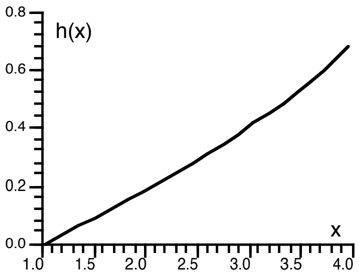
$$dalla\ media: \quad \mu'_\alpha = \int_{-\infty}^{\infty} (x-\mu)^\alpha h(x)dx;$$

$$assoluti: \quad \upsilon_\alpha = \int_{-\infty}^{\infty} |x|^\alpha h(x)dx;$$

Esempio

$$h(x)=\frac{6\left(x-\sqrt{x}\right)}{17},\quad 1\le x\le 4;\quad \mu_1=\int_1^4 x\frac{6\left(x-\sqrt{x}\right)}{17}dx=2.8;$$

$$\mu_2=\int_1^4 x^2\frac{6\left(x-\sqrt{x}\right)}{17}dx=9.693;\quad \sigma=\sqrt{9.693-2.8^2}=1.36$$

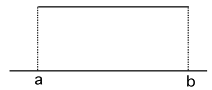


Scarto quadratico medio

$$\sigma=\sqrt{\int_{-\infty}^{\infty}(x-\mu)^2h(x)dx}=\sqrt{\int_{-\infty}^{\infty}x^2h(x)dx-\mu^2}$$

Esempio:

densità uniforme: $h(x)=\frac{1}{b-a};\qquad a\le x\le b$

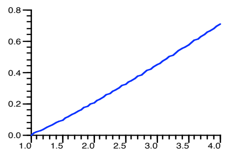


$$\int_a^b\frac{x^2}{(b-a)}dx=\left.\frac{x^3}{3(b-a)}\right|_a^b=\frac{b^3-a^3}{3(b-a)};\quad \int_a^b\frac{x}{(b-a)}dx=\left.\frac{x^2}{2(b-a)}\right|_a^b=\frac{b^2-a^2}{2(b-a)}$$

$$=\frac{(b+a)(b-a)}{2(b-a)}=\frac{b+a}{2}$$

$$\sigma^2=\frac{b^3-a^3}{3(b-a)}-\left[\frac{b+a}{2}\right]^2=\frac{(b-a)^3}{12(b-a)}-\frac{(b-a)^2}{12}\Rightarrow\sigma=\frac{b-a}{\sqrt{12}}$$

Esercizio



Un fenomeno ha densità di frequenza pari allo scarto tra livello raggiunto e la sua radice quadra

$$h(x)=\frac{6\left(x-\sqrt{x}\right)}{17};\quad 1\le x\le 4$$

$$\mu_1=\frac{6}{17}\int_1^4 x\left(x-\sqrt{x}\right)dx=\frac{6}{17}\int_1^4 x^2-x^{3/2}dx=\frac{6}{17}\left[\int_1^4 x^2dx-\int_1^4 x^{3/2}dx\right]=\frac{6}{17}\left[\left.\frac{x^3}{3}\right|_1^4-\left.\frac{x^{5/2}}{5/2}\right|_1^4\right]$$

$$=\frac{6}{17}\left[\left(\frac{64}{3}-\frac{1}{3}\right)-\left(\frac{32}{5/2}-\frac{1}{5/2}\right)\right]=\frac{6}{17}\left[\frac{61}{3}-\frac{62}{5}\right]=2.8$$

$$\mu_2=\frac{6}{17}\int_1^4 x^2\left(x-\sqrt{x}\right)dx=\frac{6}{17}\int_1^4 x^3-x^{5/2}dx=\frac{6}{17}\left[\left.\frac{x^4}{4}\right|_1^4-\left.\frac{x^{7/2}}{7/2}\right|_1^4\right]$$

$$=\frac{6}{17}\left[\left(64-\frac{1}{4}\right)-\left(128-\frac{1}{7/2}\right)\right]=\frac{6}{17}\left[\frac{61}{3}-\frac{62}{5}\right]=9.693\qquad \sigma=\sqrt{9.693-2.8^2}=1.36$$

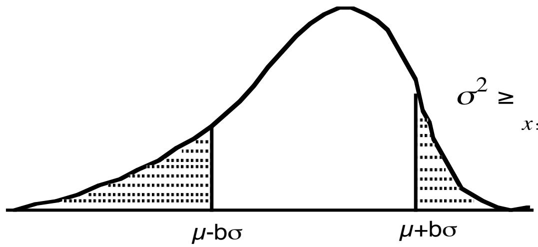
La disuguaglianza di Tchebycheff

L'idea è di approssimare, in base a "μ" e "σ" la probabilità associabile ad intervalli del tipo:|X-μ|≥bσ

$$\sigma^2=\sum_{i=1}^{k}\left(X_i-\mu\right)^2f_i=$$

$$\sum_{x\le\mu-b\sigma}\left(X_i-\mu\right)^2f_i+\sum_{x\ge\mu+b\sigma}\left(X_i-\mu\right)^2f_i+\cancel{\sum_{\mu-b\sigma x<\mu+b\sigma}\left(X_i-\mu\right)^2f_i}$$

E' positivo!



$$\sigma^2\ge\sum_{x\le\mu-b\sigma}\left(X_i-\mu\right)^2f_i+\sum_{x\ge\mu+b\sigma}\left(X_i-\mu\right)^2f_i$$

La disuguaglianza di Tchebycheff/2

Nelle due regioni estreme si ha: $|X-\mu| \geq b\sigma \Rightarrow (X-\mu)^2 \geq (b\sigma)^2$

Quindi: $\sigma^2 \geq \sum_{x \leq \mu - b\sigma} (b\sigma)^2 f_i + \sum_{x \geq \mu + b\sigma} (b\sigma)^2 f_i = \sigma^2 b^2 \left[\sum_{x \leq \mu - b\sigma} f_i + \sum_{x \geq \mu + b\sigma} f_i \right]$

ovvero: $\sigma^2 \geq \sigma^2 b^2 \left[\sum_{x \leq \mu - b\sigma} f_i + \sum_{x \geq \mu + b\sigma} f_i \right] \Rightarrow \left[\sum_{x \leq \mu - b\sigma} f_i + \sum_{x \geq \mu + b\sigma} f_i \right] \leq \frac{1}{b^2}$

In definitiva: $\Pr\left(|X-\mu| < b\sigma\right) \geq 1 - \frac{1}{b^2}$

La probabilità che il valore della v.c. ricada in un intervallo simmetrico intorno al valore atteso, di ampiezza basata sullo scarto quadratico medio, ha una soglia minima prestabilita

Soglie tipiche

per b<1 non sono informative

b	Soglia	b	Soglia
1.00	0.00	2.75	0.87
1.25	0.36	3.00	0.89
1.50	0.56	3.25	0.91
1.75	0.67	3.50	0.92
2.00	0.75	3.75	0.93
2.25	0.80	4.00	0.94
2.50	0.84	4.25	0.94

La disuguaglianza di Tchebycheff è utile soprattutto dal punto di vista teorico Poniamo c=bσ ovvero b=c/σ; ne consegue:

$$\Pr\left(|X-\mu| < c\right) \geq 1 - \left(\frac{\sigma}{c}\right)^2$$

Più grande è σ meno tipico e rappresentativo sarà il valore atteso;

Più piccolo è σ più corto sarà l'intervallo nel quale ricade una data massa di probabilità e, tanto meno dispersa sarà la distribuzione

La asimmetria

Per misurare l'asimmetria delle funzioni di densità si possono usare gli indici già adoperati: l'α_1 e α_2 .

$$h(x) = \begin{cases} \frac{2x^2 - x}{34.5}; & per \quad 1 \leq x \leq 4 \\ 0 & altrove \end{cases}$$

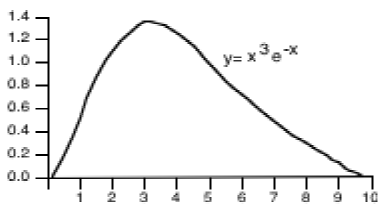
$$M_e = 3.2395, \; \mu = 3.0869, \; \sigma = 0.69547, \; \alpha_1 = -0.2194$$

$$h(x) = x^n e^{-bx} \quad x > 0; \quad \mu = \int_0^\infty x^{n+1} e^{-bx} dx = \frac{(n+1)!}{b^{n+2}};$$

$$h'(x) = x^{n-1} e^{-bx}(n - bx) \Rightarrow M_o = \frac{n}{b}$$

$$\sigma = \int_0^\infty x^{n+2} e^{-bx} dx - \mu^2 = \frac{(n+2)!}{b^{n+3}} - \left[\frac{(n+1)!}{b^{n+2}}\right]^2;$$

$$\alpha_2 = \frac{\frac{(n+1)!}{b^{n+2}} - \frac{n}{b}}{\sigma}; \; nullo \;\; se \;\; b = \left[\frac{(n+1)!}{n}\right]^{\frac{1}{n+1}}$$



La asimmetria/2

Feller (1950, pp. 78-81) propone il modello arcoseno per situazioni in cui si possono verificare molti casi nei due estremi e basse al centro quali per esempio le durate dei soggiorni turistici.

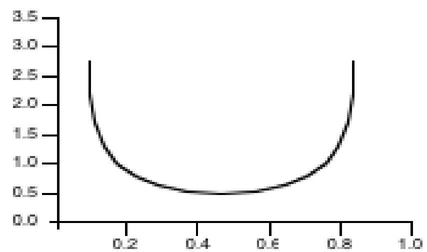
$$h(x) = \frac{1}{\pi\sqrt{x(1-x)}} \quad 0 < x < 1; \quad F(x) = \left(\frac{2}{\pi}\right) arcoseno\left(\sqrt{x}\right)$$

$$X_p = sen^2\left(\frac{\pi p}{2}\right) \Rightarrow M_e = sen^2\left(\frac{\pi}{4}\right) = \frac{1}{2}$$

$$\mu = \int_0^1 sen^2\left(\frac{\pi p}{2}\right) dp = \frac{2}{\pi}\left\{\frac{x}{4} - \frac{sen(2x)}{4}\Bigg|_0^{\frac{\pi}{2}}\right\} = \frac{1}{2}; \; S_{M_e} = \frac{1}{\pi}; \; \alpha_1 = \pi\left(\frac{1}{2} - \frac{1}{2}\right) = 0$$

Le misure di asimmetria hanno le stesse finalità e carenze che in descrittiva.

L'α_1 è nullo se la funzione di densità è simmetrica, ma il fato che sia nullo non implica, necessariamente, che la densità sia simmetrica



Esistenza (finita) dei momenti

il momento r-esismo esiste finito se e solo se

$$\lim_{x\to\infty} x^r\left[1-F(x)\right]\to 0; \qquad \lim_{x\to\infty} x^r\left[F(-x)\right]\to 0;$$

maggiore è l'esponente "r" più rapido sarà il decadimento verso le ascisse.

Se, per un certo "r", i momenti sono infiniti allora sono infiniti tutti i momenti di ordine superiore perché:

$$\upsilon_r^{s-t} \le \upsilon_t^{s-r}\upsilon_s^{r-t} \Rightarrow \upsilon_{r+1} \le \frac{\upsilon_r^2}{\upsilon_{r-1}}$$

se per una variabile la media risulta infinita sarà infinita anche la varianza e gli altri momenti a seguire.

Se la varianza esiste finita allora anche la media esiste finita.

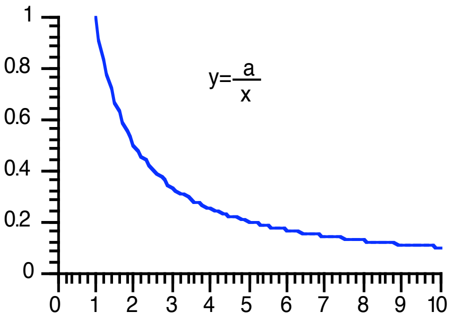
Esempio

Densità elevate per valori alti pongono a rischio l'esistenza della media (la centralità deve essere misurata con altri indici).

Modello iperbolico:

$$h(x)=\frac{a}{x^2};\ x>a$$

$$\mu=\int_a^\infty \frac{xa}{x^2}dx=\int_a^\infty \frac{a}{x}dx=Ln(x)\Big|_a^\infty=$$
$$=-Ln(a)+\lim_{x\to\infty} Ln(x)=\infty$$



Funzione generatrice dei momenti

Serve per determinare i momenti ed altre proprietà delle distribuzioni

$$M_X(t)=E\left(e^{tX}\right)=\int_{-\infty}^{\infty} e^{tx}f_X(x)dx;\quad \text{per } -\delta<t<\delta,\ \delta>0$$

L'integrale esiste sempre in t=0 dato che $M_X(0)=E(e^{0X})=E(1)=1$

L'esitenza della f.g.m. presuppone la convergenza dell'integrale in un intonro completo dell'origine

La f.g.m. può non esistere se la densità $f_X(x)$ è positiva su un intervallo non limitato

Esempi

Esponenziale

$$f_X(x)=\lambda e^{-\lambda x},\ x>0,\ \lambda>0$$

$$\int_{-\infty}^{\infty} e^{tx}f_X(x)dx\Rightarrow\int_{-\infty}^{\infty} e^{tx}\lambda e^{-\lambda x}dx=\lambda\int_{-\infty}^{\infty} e^{-x(\lambda-t)}dx=\begin{cases}\dfrac{\lambda}{\lambda-t} & se\ \ t<\lambda\\ \infty & se\ \ t\ge\lambda\end{cases}$$

Poiché $\lambda>0$ l'insieme dei valori reali di t per cgui l'integrale esiste finito contiene un'intorno dell'origine e quindi X ha una f.g.m.

Iperbolica

$$f_X(x)=x^{-2},\ \ x>1$$

$$\int_{-\infty}^{\infty} e^{tx}f_X(x)dx\Rightarrow\int_{-\infty}^{\infty}\frac{e^{tx}}{x^2}dx=\int_1^{\infty}\frac{e^{tx}}{x^2}dx\begin{cases}\ge\displaystyle\int_1^\infty\frac{1+tx}{x^2}dx\to\infty,\ t>0\\ \le\displaystyle\int_1^\infty\frac{1}{x^2}dx=1,\ t\le 0\end{cases}$$

L'insieme dei valori per cui l'integrale converge non include un intorno completo dello zero

Aspetti teorici

Se X ed Y hanno funzione di ripartizione $F_X()$ ed $F_Y()$ rispettivamente e con f.g.m. esistenti finite $M_X(t)$ e $M_Y(t)$ allora

$$\left\{M_X(t)=M_Y(t)\right\}\xleftrightarrow{def}\left\{F_X(x)=F_Y(y)\right\}$$

Se una variabile casuale è dotata di f.g.m. allora essa identifica univocamente la variabile casuale e viceversa.

La f.g.m. se esiste finita è sviluppabile in serie di Mac Laurin

$$M_X(t)=\sum_{n=0}^{\infty}\left[\frac{d^m}{dt^n}M_X(t)\right]_{t=0}\left(\frac{t^n}{n!}\right)=\sum_{n=0}^{\infty}E\left(X^n\right)\left(\frac{t^n}{n!}\right),\ n\geq 1\ intero$$

Se la f.g.m. di X esiste finita allora X possiede finiti tutti i momenti (ricavabili come derivate della f.g.m.). Il contrario non è sempre vero.

Esempio

$$M_X(t)=\frac{\lambda}{\lambda-t}\quad -\infty<t<\lambda$$

Esponenziale

Momento primo e secondo

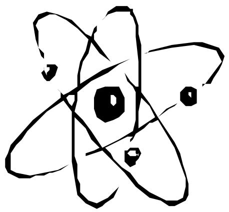
$$M'_X(t)=\frac{\lambda}{(\lambda-t)^2}\quad\Rightarrow M'_X(0)=E(X)=\frac{\lambda}{(\lambda-0)^2}=\lambda$$

$$M''_X(t)=\frac{2\lambda}{(\lambda-t)^3}\quad\Rightarrow M''_X(0)=E\left(X^2\right)=\frac{2}{\lambda^2}$$

Modelli di variabili casuali

Le probabilità (come tabella o come formula) derivano:

- **Dalle condizioni sperimentali**
- **Dall'aspetto descritto dalla variabile casuale**
- **Da informazioni aggiuntive**

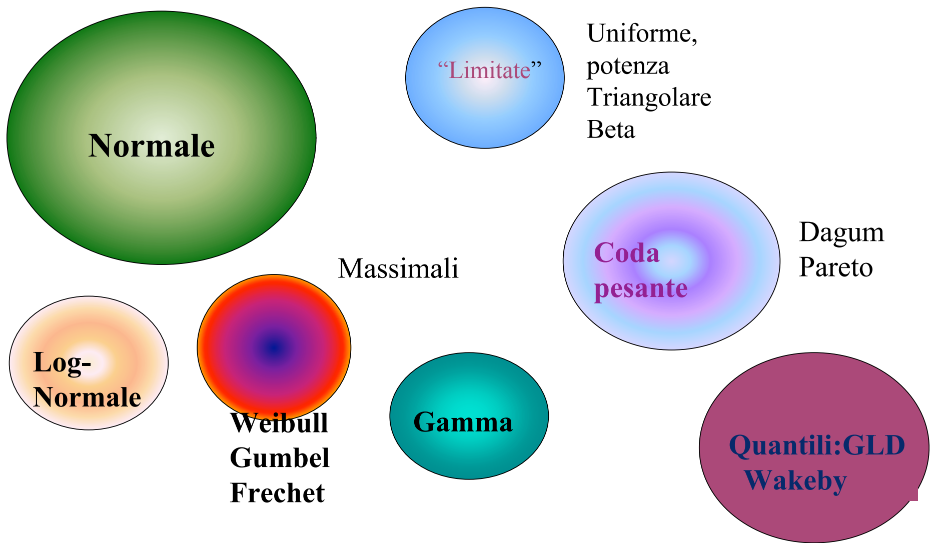


La statistica ha elaborato molti modelli sia per le variabili nominali e discrete che per le variabili continue.

Molti possono essere studiati e approfonditi nel libro di testo (apprendimento libero)

Alcuni meritano una presentazione nel corso

Breve lista di modelli probabilistici



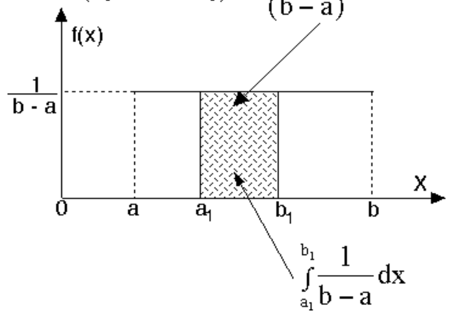
La v.c. Uniforme (Rettamgolare)

Le caratteristica principale è che le sue realizzazioni sono equiprobabili

Si applica nelle situazioni in cui il fenomeno:

- **Assume valori in un intervallo limitato [a,b]**
- **La probabilità di ogni sottointervallo di [a, b] è proporzionale all'ampiezza del sottointervallo**

$$P(a_1 \le X \le b_1) = \frac{(b_1 - a_1)}{(b-a)}$$



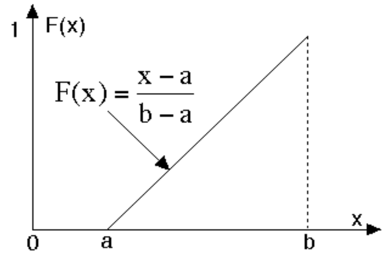
La funzione di densità è:

$$f(X) = \begin{cases} \frac{1}{b-a} & \text{per} \quad a \le X \le b \\ 0 & \text{altrove} \end{cases}$$

Uniforme/2

La funzione di ripartizione è pure semplice:

$$F(x) = P(a \le X \le x) = \int_a^x \frac{1}{b-a} dt = \frac{(x-a)}{(b-a)}$$



$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} * \frac{x^2}{2}\Big|_a^b$$

$$= \frac{1}{b-a} * \left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{1}{b-a} * \left[\frac{(b-a)*(b+a)}{2}\right] = \frac{b+a}{2}$$ **= valore atteso della distribuzione**

Nella v.c. UNIFORME il valore atteso coincide con il valore centrale del suo supporto.

Uniforme/3

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

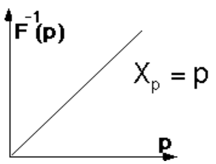
Quantili: $X_p = a + p * (b-a)$

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

Esempio:
Supponiamo che la v.c. X abbia la funzione di densità $f(x) = 1 \quad \text{per} \quad 0 < x < 1$

Allora:

$$E(X) = \frac{0+1}{2} = 0.5; \quad \sigma^2(X) = \frac{(1-0)^2}{12} = 0.083$$



Esercizio

Gli autobus dell'università passano ogni ora fra le 8.30 e le 13.30.

1) Calcolare la probabilità che una persona, capitando a caso durante tale periodo, debba aspettare almeno un quarto d'ora;
2) Calcolare il tempo medio di attesa ed il suo SQM

La v.c. "X" è "tempo mancante al prossimo bus". Il modello è uniforme dato che "capitare a caso" significa che può capitare in uno qualsiasi dei 60 minuti. Quindi:

$$P(x \ge 15) = 1 - P(X < 15) = 1 - F(15)$$
$$= 1 - \frac{15-0}{60-0} = 1 - 0.25 = 0.75$$

$$E(X) = \frac{0+60}{2} = 30 \ \text{(mezzora)}$$

$$\sigma(X) = \sqrt{\frac{(60-0)^2}{12}} = 17.32 \ \text{minuti}$$

Dalla uniforme alle altre

Se X ha funzione di ripartizione F(.) la trasformazione y=F(x) produce una variabile casuale uniforme nell'intervallo [0,1].

$$\Pr(Y \le y) = \Pr\left[F(X) \le y\right] = \Pr\left\{F^{-1}\left[F(X)\right] \le F^{-1}(y)\right\}$$
$$= \Pr\left[X \le F^{-1}(y)\right] = F\left[F^{-1}(y)\right] = y$$

In caso di ripartizione invertibile si otterrà

$$F_y(y) = \begin{cases} 0 & se \;\; y \le 0 \\ y & se \;\; 0 < y < 1 \\ 1 & se \;\; y \ge 1 \end{cases} \;\Rightarrow\; h(y) = \begin{cases} 1 & se \;\; 0 \le x \le 1 \\ 0 & altrove \end{cases}$$

cioè il modello di densità uniforme sull'intervallo unitario.

Generazione di variabili casuali

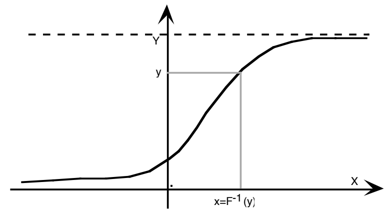
L'aspetto più interessante di questa trasformazione è però la sua inversa.

Partiamo dalla uniforme [0,1] Y e definiamo una nuova variabile casuale con la trasformazione X=Q(Y) dove Q(.) è la funzione di quantile.

$$\Pr(X \le x) = \Pr\left[Q(Y) \le x\right] = \Pr\left[Y \le F(x)\right] = F(x)$$

Per ottenere una variabile casuale che abbia la funzione di ripartizione F(.) basta disporre della sua inversa.

Si usa poi la uniforme



Esempi

Se X è uniforme sull'intervallo [0,1], la trasformazione per arrivare alla ripartizione di Gumbel è:

$$F(x) = 1 - e^{-e^{-\left(\frac{x-a}{b}\right)}}; \;\; x > 0; \quad p = F(x) \Rightarrow x_p = a + bLn\left[-Ln(p)\right]; \quad 0 < p < 1$$

Per ottenere una variabile casuale che abbia la funzione di ripartizione Weibull

$$F(x) = 1 - e^{-\left[\frac{x-\mu}{\sigma}\right]^{\gamma}} \Rightarrow X = \mu + \sigma\left[-Ln(U)\right]^{\frac{1}{\gamma}}$$

Se si ha una osservazione casuale sulla uniforme allora si può ottenere una osserfazione casuale sulla Weibull o sulla Gumbel con opportune trasformazioni

La variabile casuale esponenziale

Questa v.c. si origina dalla Poisson ed è in risposta alla domanda:

Se una sequenza di eventi si verifica nel tempo secondo il modello di Poisson alla media di λ eventi per unità di tempo, quanto tempo bisogna aspettare perchè si verifichi il primo?

Riconsideriamo la Poisson per l'arrivo di clienti ad uno sportello bancario

$$f(X) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad \textbf{dove} \quad \begin{cases} X & \text{numero di clienti nell' unità di tempo} \\ E(X) = \lambda t & \text{media di arrivi} \\ \sigma^2(X) = \lambda t & \text{varianza degli arrivi} \end{cases}$$

Indichiamo con "T" il tempo trascorso fino all'arrivo del primo cliente ovvero quello che intercorre tra un cliente ed un altro.

"T" è una v.c. continua che assume valori tra zero e infinito

Esponenziale/2

Analizziamo ora l'evento: il primo arrivo avviene dopo un tempo "t", cioè (T>t)

Tale evento si realizza se non ci sono occorrenze nell'arco di tempo "t", cioè se la v.c. di Poisson è x=0.

Il converso è anche vero: il numero di arrivi durante l'arco di tempo "t" è zero (x=0) se il primo arrivo si verifica a T>t. Quindi

$$P(T > t) = P(X = 0) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

Poiché $P(T \leq t) + P(T > t) = 1$ **si ha inoltre**

$$P(T \leq t) = 1 - e^{-\lambda t}$$

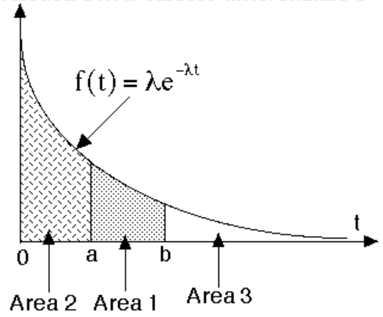
Questa è la funzione di ripartizione della v.c. esponenziale

$$F(t) = 1 - e^{-\lambda t} \quad \text{per} \quad t \geq 0$$

Esponenziale/3

La funzione di densità della esponenziale si ricava con il calcolo differenziale e risulta essere

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



$$F(t) = 1 - e^{-\lambda t} \qquad X_p = \frac{1}{\lambda} * Ln\left(\frac{1}{1-p}\right)$$

$$\text{Area } 1 = P(a < T < b) = \int_a^b \lambda e^{\lambda t} dt = F(b) - F(a) = \left(1 - e^{-\lambda b}\right) - \left(1 - e^{-\lambda a}\right) = e^{-\lambda a} - e^{\lambda b}$$

$$\text{Area } 2 = P(T \leq a) = F(a) = 1 - e^{-\lambda a} \qquad \text{Area } 3 = P(T > b) = 1 - F(b) = e^{-\lambda b}$$

Esempio

I clienti di un ipermercato arrivano alle entrate secondo un modello di Poisson ad una media di λ=4 al minuto e cioè

$$P(X = x) = \frac{4^x e^{-4}}{x!}; \quad x = 0, 1, 2, \ldots,$$

X=numero di clienti in un dato minuto

Quanto tempo occorre aspettare dopo l'apertura prima che entri il primo cliente?

Quanto si deve aspettare tra un cliente e l'altro?

$$P(X = x) = 4e^{-4x} \quad X > 0$$

Ad esempio la probabilità di aspettare meno di mezzo minuto è

$$F(0.5) = 1 - e^{-4 * \frac{1}{2}} = 1 - e^{-2} = 86.4\%$$

Esercizio

Si supponga che un rappresentante effettui le consegne secondo una legge di Poisson al ritmo di 4 consegne ogni ora. Cioè

$$P(\text{Rappresentante arriva in t ore}) = \frac{(4t)^x e^{-4t}}{x!} \quad x = 0, 1, 2, \ldots$$

con una media di λt=4t consegne ogni t ore.

Calcolare la probabilità che un punto vendita aspetti meno di un quarto d'ora prima di vederlo arrivare

Poiché si effettuano 4 consegne all'ora si ha $\lambda = \frac{4}{60} = \frac{1}{15}$ **per cui**

$$P(\text{arriva in meno di t'}) = F(t) = 1 - e^{\frac{t}{15}}$$

Nel caso particolare:

$$P(t \leq 15) = F(15) = 1 - e^{\frac{15}{15}} = 1 - e^{-1} = 0.632$$

Altro esercizio

il monitor di una linea di televisori ha durata descritta da una v.c. esponenziale con media 10 anni.

a) calcolare la probabilità che si guasti dopo 15 anni;
b) si guasti prima del periodo di garanzia di 2 anni.

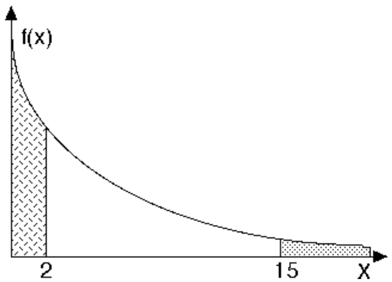
indichiamo con "X" la durata e ricaviamo il parametro "λ"

$$E(X) = 10 \Rightarrow \lambda = \frac{1}{E(X)} = \frac{1}{10} = 0.1$$

a) $P(X > 15) = e^{-0.1*15} = 0.22$

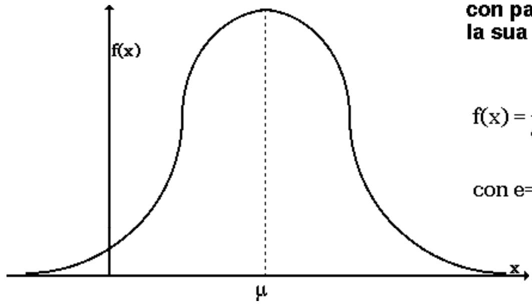
b) $P(X \le 2) = 1 - e^{-0.1*2} = 0.18$

Il 22% avrà lunga durata, ma il 18% darà dei costidi riparazione in garanzia



La v.c. normale (o gaussiana)

E' la v.c. più importante, più nota e più usata in statistica



Una variabile casuale è di tipo normale con parametri μ e σ : $X\sim N(\mu, \sigma)$ se la sua funzione di densità è

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left\{\frac{-1}{2\sigma 2}[x-\mu]^2\right\}} \quad -\infty<\mu\infty;\ \sigma>0$$

con $e=2.71828;\ \pi = 3.14159$

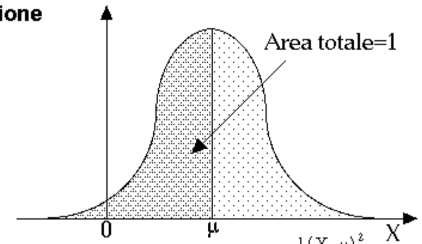
L'importanza di questa v.c. risiede negli indubbi vantaggi formali, ma anche nel fatto che moltissimi fenomeni empirici possono essere rappresentati con un modello di tipo gaussiano

Caratteristiche della Normale

- **La funzione utilizzata è in effetti una funzione di densità**

$$f(X) \ge 0; \qquad \int_{-\infty}^{+\infty} f(X)dX = 1$$

- **La densità diminuisce man mano che si allontana dal centro (L'asse delle ascisse è un asintoto della f(x))**



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

- **la funzione di densità è simmetrica intorno a μ**

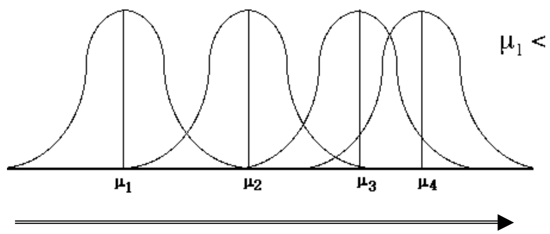
$$f(\mu + x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\mu+x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\mu-x-\mu}{\sigma}\right)^2} = f(\mu - x)$$

Ne consegue che il valore atteso coincide con la Mediana $Me = X_{50}$

Significato del parametro "μ"

La distribuzione è unimodale e l'ordinata massima si raggiunge per X=μ (La moda)

Quindi, il parametro μ rappresenta il valore più probabilie nonchè il valore atteso e quello che bipartisce il supporto dei valori.
Cambiando μ si modifica la collocazione del grafico



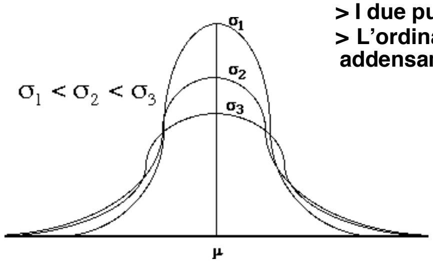
$\mu_1 < \mu_2 < \mu_3 < \mu_4$

Al variare di μ il grafico resta inalterato nella sua forma. Si modifica solo la sua localizzazione: più a destra se μ aumenta; più a sinistra se μ diminuisce

Significato del parametro "σ"

Il σ corrisponde allo scarto quadratico medio.
La curvatura del grafico della distribuzione normale cambia due volte inflessione in corrispondenza dei punti $x=\mu \pm \sigma$. Inoltre

Quindi $$f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$$



Al diminuire di σ :
> I due punti di flesso tendono ad accentrarsi;
> L'ordinata massima aumenta a causa del maggiore addensamento intorno al centro della distribuzione.

La funzione di ripartizione

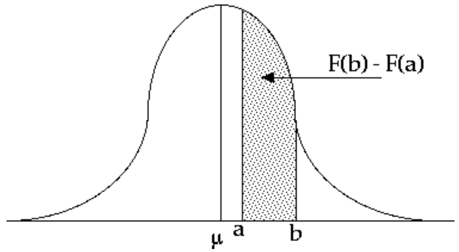
La formula della Normale è complicata e l'integrale

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

è calcolato con metodi di approssimazione numerica

La v.c. Normale è completamente specificata dai parametri μ e σ cioè noti questi non c'è altro da sapere. Ne consegue che ogni calcolo della F(x) deve essere ripetuto per ogni combinazione dei due parametri

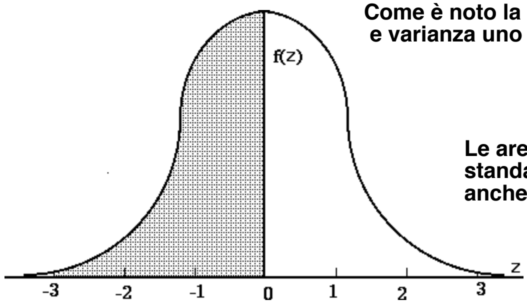
Esiste comunque una tecnica che ci permette ci calcolare le probabilità di tutte le Normali usando una sola tabella



Curva normale standardizzata

E' possibile esprimere la v.c. Normale in unità standard $z = \frac{x - \mu}{\sigma}$

La funzione di densità è ora $f(z) = \frac{1}{\sqrt{2\pi}} e^{\left\{\frac{-z^2}{2}\right\}}$ **Indicata con N~(0,1)**

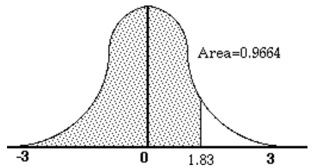


Come è noto la v.c. standardizzata ha media zero e varianza uno

Le aree sottese alla curva normale standardizzate sono state tabulate (sono anche disponibili su alcune calcolatrici).

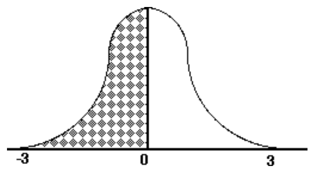
Calcolo delle aree sottese alla curva

Esempio_1: calcolare l'area compresa fra -∞ e 1.83, in simboli:$\phi(1.83)$



$\phi(1.83) = 0.9664$ $P(Z \leq 1.83)$

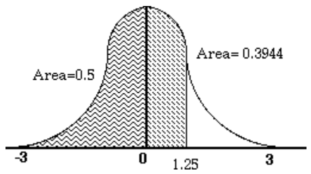
Esempio_2: calcolare l'area compresa fra -∞ e 0, in simboli: $\phi(0)$



$2\phi(0)=1$ e quindi $\phi(0)=\frac{1}{2}$ $P(Z \leq 0)$

Altri esempi

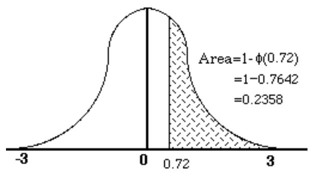
Esempio_3: calcolo dell'area tra -∞ e 1.25 $P(Z \leq 1.25)$



Alcune tavole danno invece l'area compresa tra il valore dato e lo zero. Per ottenere l'intera area occorre sommare 0.5 che è l'area a sinistra dello zero.

$\phi(1.25) = 0.5 + 0.3944 = 0.8944$

Esempio_4: calcolo dell'area a destra di 0.72 $P(Z > 0.72)$



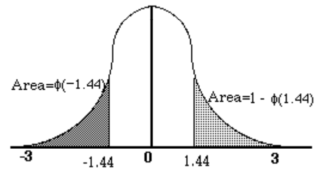
Poiché l'area totale è uguale ad uno si ha l'identità:

Area a sinistra=1- Area a destra

$$1 - P(Z \leq 0.72)$$

Esercizio

Esempio_5: calcolo dell'area per valori negativi: $\phi(-1.44)$ $P(Z \leq -1.44)$



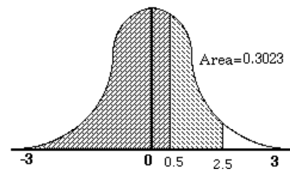
La curva normale è simmetrica per cui

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

Per calcolare l'area a sinistra di un valore negativo si usa l'area a destra del valore positivo corrispondente:

$\phi(-1.44) = 1 - \phi(1.44) = 1 - 0.9251 = 0.0749$

Esempio_6: calcolo dell'area compresa tra due valori positivi. $P(0.5 \leq Z \leq 2.5)$



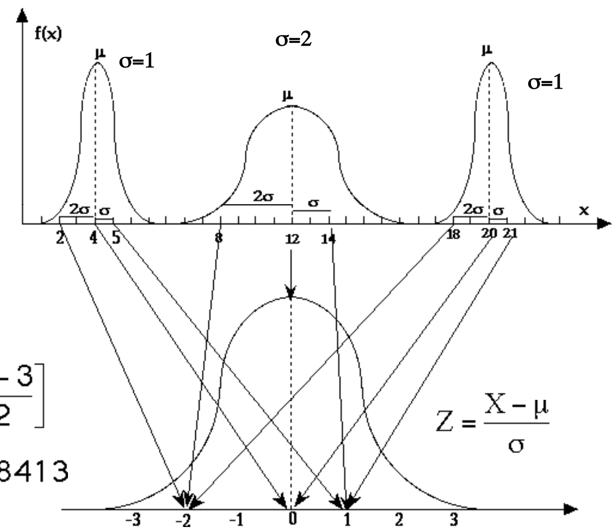
L'area compresa tra due valori qualsiasi si ottiene per differenza tra le aree a sinistra degli estremi dell'intervallo

Area=$\phi(2.5) - \phi(0.5)$ =0.9938 - 0.6915 =0.3023

Generalizzazione

Le tre diverse v.c. normali possono essere ricondotte ad una sola distribuzione con la trasformazione in unità standard.

Le aree sottese a $X\sim(\mu,\sigma)$ sono identiche a quelle della $Z\sim N(0,1)$ dopo aver trasformato la X in unità standard



$$Z = \frac{X - \mu}{\sigma}$$

$$X \sim N(3,4)$$

$$P(X < 5) = P\left[\frac{X-3}{2} \leq \frac{5-3}{2}\right]$$

$$= P(Z \leq 1) = 0.8413$$

Esempio

Sia X una v.c. Normale con media μ=16 e deviazione standard σ=5.

1) Calcolare: $P(11 < X < 21)$

$$P(11 < X < 21) = P\left(\frac{11-16}{5} < \frac{X-16}{5} < \frac{21-16}{5}\right) = P(-1 < Z < 1)$$

$$= \phi(1) - \phi(-1) = \phi(1) - [1 - \phi(1)] = 2\phi(1) - 1$$

$$= 2 * 0.8413 - 1 = 0.6826$$

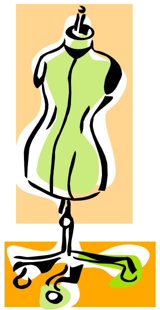
2) Calcolare: $P(X > 26)$

$$P(X > 26) = 1 - P(X \leq 26) = 1 - P\left(\frac{X-16}{5} \leq \frac{26-16}{5}\right) = 1 - P(Z \leq 2)$$

$$= 1 - 0.9772 = 0.0228$$

Esercizio

Supponiamo che le vendite di vestiti di un negozio specializzato seguano la distribuzione Normale con una media di 36 vendite al giorno ed una deviazione standard di 9 vendite al giorno.

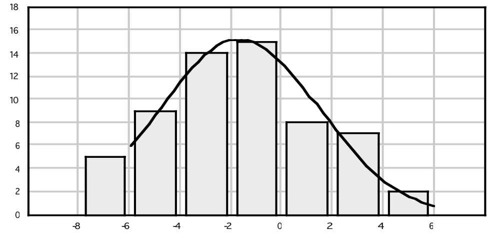


Calcolare la probabilità che in un dato giorno si vendano più di 12 vestiti

$$P(X > 12) = 1 - P(X \le 12) = 1 - P\left(\frac{X-36}{9} \le \frac{12-36}{9}\right) = 1 - P(Z \le -2.67)$$
$$= P(Z \le 2.67) = \phi(2.67) = 0.9962$$

Uso della normale/1

Si è rilevato l'errore (rispetto ad un valore prefissato) nel peso di un campione di 60 lotti.



-7.2	-4.8	-1.7	-3.1	-0.7	1.6
-6.7	-4.8	-1.6	-3.1	-0.5	2.2
-6.7	-4.8	-1.5	-2.9	-0.1	2.3
-6.6	-4.6	-1.4	-2.8	0.3	3.4
-6.5	-3.8	-1.4	-2.7	0.3	3.4
-5.6	-3.7	-1.3	-2.7	0.6	3.5
-5.5	-3.6	-1.3	-2.2	0.8	3.5
-5.3	-3.5	-1.2	-2.2	1.3	3.6
-5.1	-3.5	-1.0	-1.9	1.5	4.2
-5.0	-3.3	-0.8	-1.8	1.5	5.5

$$\hat{\mu} = -1.683, \quad s = 3.106$$

Ipotizziamo che tali valori siano validi per tutti i possibili lotti.

Se si sceglie a caso un lotto, qual è la probabilità che l'errore nel peso sia compreso tra -0.5 e +0.5

$$P(-0.5 \le x \le 0.5) = P\left(\frac{-0.5+1.683}{3.106} \le z \le \frac{0.5+1.683}{3.106}\right) = P(0.38 \le z \le 0.70)$$
$$= \Phi(0.7) - \Phi(0.38) = 0.758 - 0.648 = 0.11 \Rightarrow 11\%$$

Approssimazione di v.c. Discrete

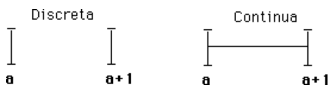
In alcuni casi e nonostante la contraddizione, le distribuzioni delle v.c. discrete possono essere approssimate dalla v.c. normale o gaussiana

Occorre però tenere conto che i valori possibili delle discrete sono assimilabili ad un insieme di numeri interi naturali: 0, 1, 2, ..., n,... laddove la Gaussiana può assumere ogni valore reale

Nel discreto si ha $P(X \le a) + P(X \ge a+1) = 1$

Nel continuo $P(X \le a) + P(X \ge a+1) \le 1$

in quanto non sono inclusi i valori dell'intervallo $a < X < a+1$



Approssimazione di v.c. discrete/2

Per superare tali problemi si usa un fattore di correzione della continuità pari a 0.5.

Se "X" è discreta e deve essere approssimata con una continua "Y" si ha:

$$P(X = a) \cong P(a - 0.5 \le Y \le a + 0.5)$$

Inoltre, per recuperare la differenza tra includere o no l'estremo nelle discrete, si pone

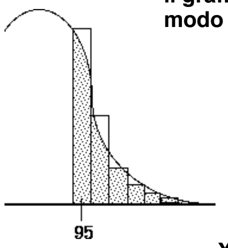
Sono convenzioni

$P(X < a) = P(Y \le a - 0.5);$ $\quad P(X \le a) = P(Y \le a + 0.5)$

$P(X > a) = P(Y > a + 0.5);$ $\quad P(X \ge a) = P(Y \ge a - 0.5)$

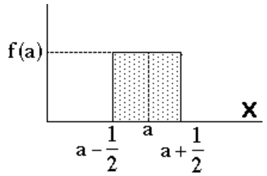
Esempio

Le risposte esatte "X" in un test attitudinale hanno μ=80 e σ=10. Calcolare, con la gaussiana "Y", la frazione di candidati con punteggio inferiore a 95



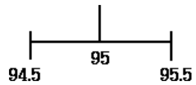
il grafico della f(x) assume che gli istogrammi abbiano base "1" in modo che l'area sia uguale alla probabilità.

$$P(X = a) = f(a)*1$$



La normale si usa per approssimare le aree degli istogrammi

X ≥ 95 significa in realtà ≥ 94.5
X ≤ 95 significa in realtà ≤ 95.5



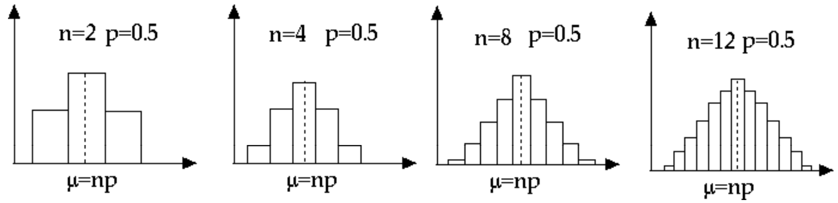
Calcolo relativo all'esempio:

$$P(X < 95) \cong P(Y \le 94.5) = P\left(\frac{Y-80}{10} \le \frac{94.5-80}{10}\right) = \phi(1.45) = 0.9265$$

Approssimazione della Binomiale

La Normale è una buona approssimazione della Binomiale se

- **"n", il numero delle prove, è molto grande**
- **"p", la probabilità di successo, è molto vicino a 0.5**



Come si vede, all'aumentare di "n" la distribuzione assume una forma sempre più campanulare giustificando l'uso della normale.

Regola pratica

la procedura per l'approssimazione della binomiale con la normale ha 3 passi:

1. **Si calcolano** $\mu = n*p; \quad \sigma = \sqrt{n*p*(1-p)}$

2. **Si applica il fattore di correzione della continuità per convertire la v.c. discreta in continua**

$$Z = \frac{(X-0.5)-\mu}{\sigma}; \quad \text{oppure} \quad Z = \frac{(X+0.5)-\mu}{\sigma}$$

3. **Si usano le tavole della normale standardizzata per determinare le probabilità della binomiale**

$$P(X=a) \cong \phi\left(\frac{a+0.5-\mu}{\sigma}\right) - \phi\left(\frac{a-0.5-\mu}{\sigma}\right)$$

$$P(X<b) \cong \phi\left(\frac{b-0.5-\mu}{\sigma}\right)$$

$$P(X \le b) \cong \phi\left(\frac{b+0.5-\mu}{\sigma}\right)$$

Esempio

Le nuove imprese hanno il 50% di probabilità di fallire al primo anno. Usando la formula della binomiale, la probabilità che 40 (su n=100) imprese siano ancora in vita dopo il primo esercizio è

$$P(X = 40) = \frac{100!}{40!60!}(0.5)^{40}(0.5)^{60} = 0.0108$$

Usiamo l'approssimazione normale

1. $\mu = n*p = 100*0.5 = 50; \quad \sigma = \sqrt{50*0.5} = 5$

2. $P(X = 40) \cong P(39.5 \le X \le 40.5) = P\left(\frac{39.5-50}{5} \le Z \le \frac{40.5-50}{5}\right) = P(-2.1 < Z < -1.9)$

3. $\phi(-1.9) - \phi(-2.1) = 1 - \phi(1.9) - [1 - \phi(2.1)] = 0.9821 - 0.9713 = 0.0108$

Esercizio

il 20% della "Gente" è convinta di sopravvivere ad un conflitto nucleare. Si indichi con "Y" il numero di tali persone in un campione di ampiezza n=25.

a) Si adoperi l'approssimazione Normale della Binomiale per determinare: $P(6 \le Y \le 9)$
b) La si confronti con il valore ottenuto usando direttamente la Binomiale.

a) $\hat{\mu} = 25*0.20 = 5; \quad \hat{\sigma} = \sqrt{\hat{\mu}*0.8} = 2$

$$P(6 \le Y \le 9) = P(5.5 \le X \le 9.5) = P\left(\frac{5.5-5}{2} \le \frac{X-5}{2} \le \frac{9.5-5}{2}\right)$$
$$= P(0.25 \le Z \le 2.25) = \phi(2.25) - \phi(0.25) = 0.9878 - 0.5987$$
$$= 0.3891$$

b) $P(6 \le Y \le 9) = \sum_{i=0}^{9}\binom{25}{i}0.2^i 0.8^{25-i} - \sum_{i=0}^{6}\binom{25}{i}0.2^i 0.8^{25-i} = 0.9827 - 0.7800 = 0.2027$

L'asimmetria della binomiale (dovuta al parametro "p" lontano da 0.5) ha comportato una notevole differenza nell'approssimazione della probabilità

Approssimazione normale della Poisson

Anche la Poisson è approssimabile dalla normale ponendo $\mu = \lambda; \quad \sigma = \sqrt{\lambda}$

però si deve avere una media elevata, diciamo, $\lambda \ge 10$

ESEMPIO

il numero delle vendite settimanali di un prodotto è ben approssimato dalla Poisson con una media di 2 vendite al giorno. Calcolare la probabilità che in una settimana si vendano meno di 23 prodotti

Dato che in media si vendono 2 unità al giorno, in media, si venderanno 14 prodotti in una settimana, per cui

$$P(0 \le X \le 23) = \sum_{i=0}^{23}\frac{(14)^i e^{-14}}{i!} = 0.99067$$

Approssimazione Normale con $\mu = \lambda = 14; \quad \sigma = \sqrt{14} = 3.7417$

$$P(0 \le X \le 23) \cong \phi\left(\frac{23+0.5-14}{3.7417}\right) - \phi\left(\frac{0-0.5-14}{3.7417}\right) = \phi(2.54) - \phi(-3.88) =$$
$$= 0.9944$$

La curtosi

La discussione è limitata alle curve unimodali e simmetriche di variabili continue

Curtosi è la deformazione cui deve essere sottoposta una curva normale, di pari media e pari deviazione standard, per sovrapporsi alla curva data

- **Curva Leptocurtica** — Curva che presenta, rispetto alla normale, ordinate più alte al centro (la sommità è più acuminata) e code spesse e allungate
- **Curva Platicurtica** — Curva che ha ordinate più alte nei valori intermedi (la sommità è più appiattita) e code sottili e brevi.
- **Curva Mesocurtica** — Curva con sommità e code simili a quelle della normale

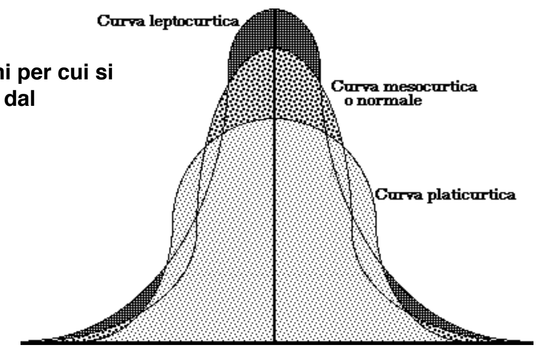
Considerazioni sulla curtosi

La curtosi è considerata una disnormalità.

Come nell'asimmetria è importante accertare il grado di allontanamento dal modello normale.

E' anche importante capire le ragioni per cui si verifica l'eventuale scostamento dal modello di riferimento

La curtosi è un concetto composito: le code e la sommità andrebbero studiate separatamente.

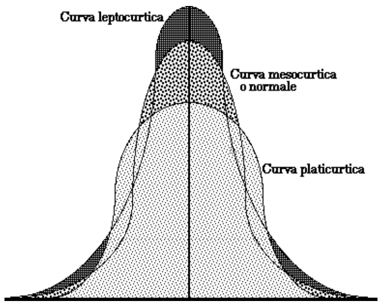


Misura della curtosi

La curtosi è misurata con il momento quarto della standardizzata

$$\gamma_2 = \int_{-\infty}^{\infty} \left[\frac{x-\mu}{\sigma}\right]^2 f(x)dx - 3$$

Poiché per la normale standardizzata si ha γ_2=0.



- **Per distribuzioni platicurtiche l'indice $\gamma_2<0$**
- **Per distribuzioni leptocurtiche $\gamma_2>0$**
- **Per distribuzioni mesocurtiche $\gamma_2=0$**

L'indice è legato alla asimmetria dalla relazione:

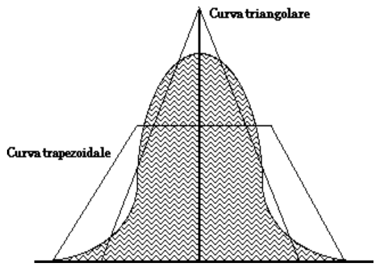
$$\gamma_2 - (\gamma_1)^2 \geq 1$$

che può essere utile nel controllo dei calcoli

Difficoltà nel concetto di curtosi

Il problema con la curtosi è che la disnormalità è dovuta a due fattori distinti:

GRADO DI APPUNTIMENTO AL CENTRO **SPESSORE DELLE CODE**



Per raggiungere lo stesso grado di appuntimento della curva triangolare, la normale dovrà essere tirata verso l'alto e assottigliata nei fianchi senza modificare le code.

Per portare le code della normale allo spessore della curva trapezoidale, si possono allargare i fianchi senza intervenire sulla sommità.

L'effetto congiunto di questi due fattori porta a valori di γ_2 incoerenti con l'aspetto della distribuzione di frequenza

L'indice di Groeneveld e Meeden

Per ovviare alla mancanza di univocità tra i valori di γ_2 e la morfologia delle curve ci sono diversi indici

In particolare c'è quello suggerito dai suddetti autori

$$GM = \frac{(X_{0.85} - Q_{0.75}) - (Q_{0.75} - X_{0.65})}{(X_{0.85} - Q_{0.75}) + (Q_{0.75} - X_{0.65})} = \frac{X_{0.85} + X_{0.65} - 2Q_{0.75}}{X_{0.85} - X_{0.65}}$$

L' indice varia nell'intervallo [-1,1] ed è invariante rispetto a trasformazioni lineari.

Valori grandi e positivi di GM saranno associati a curve leptocurtiche laddove curve platicurtiche daranno luogo a valori positivi ma piccoli (per la normale standardizzata si ha GM=0.073).

Il valore nullo di GM si ha solo per distribuzioni uniformi ed un GM negativo si avrà per curve a forma di U

Trasformazioni per la normalità

Le indagini statistiche coinvolgono solitamente moltissime variabili e non tutte possono essere indagate per gli aspetti finora considerati: centralità, variabilità e forma.

In genere ci si limita ai primi due aspetti e per il terzo si assume che la forma della poligonale delle frequenze sia simile alla curva normale. In moltissimi casi tale assunzione regge. Altre volte è meno valida: ad esempio nelle distribuzioni ad "U", a "J"

Le trasformazioni per la normalità sono delle modifiche applicate alle modalità originarie tali che la loro poligonale è più simile alla normale.

La procedura di scelta della trasformazione può essere automatizzata

La trasformazione di Box-Cox

Fra le molte trasformazioni possibili esamineremo la più famosa

$$Y_i(\alpha) = \begin{cases} \dfrac{X_i^{\alpha} - 1}{\alpha} & per \;\; \alpha \neq 0 \\ Ln(X) & per \;\; \alpha = 0 \end{cases}$$

$\alpha = 1,\ trasformazione \;\; lineare : \; Y_i = X_i - 1;$

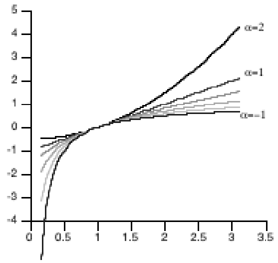
$\alpha = 0.5,\ radice : \; Y_i = 2\left(\sqrt{X_i} - 1\right);$

$\alpha = 2,\ trasformazione \;\; quadratica : \; Y_i = \dfrac{X_i^2 - 1}{2};$

$\alpha = -1,\ reciproca : \; Y_i = 1 - \dfrac{1}{X_i};$

- **Di solito il parametro α è incognito e per stimarlo occorre fare ricorso ad una routine di calcolo.**
- **Un modo meno rigoroso, ma abbastanza efficace è quello di calcolare i due coefficienti γ_1 e γ_2 (od anche gli analoghi indici basati sui percentili) per alcuni valori predefiniti di α**

{-2.5, -2.0, - 1.5, -1, -0.5, 0, 0.5, 1.0, 1.5, 2.0 2.5}



QQ_plot

Il grafico è ottenuto riportando in un sistema di assi cartesiano I valori osservati (standardizzati ed ordinati) ed i quantili della Normale corrispondenti (da cui: grafico quantile-quantile)

In dettaglio

$$\left[Z(p_i),\ \frac{x_{(i)}}{\hat{\sigma}_i} \right]; \qquad i = 1,2,\ \ldots, n$$

Il pedice tra parentesi indica valori ordinati in senso ascendente.

$Z(p_i)$ sono i quantili della normale standardizzata corrispondenti alla frequenza cumulata p_i.

Se il grafico somiglia molto ad una retta che passa per l'origine e con inclinazione unitaria (la bisettrice del quadrante per intenderci), l'approssimazione normale è considerata buona.

Posizioni grafiche

La scelta dei quantili cui far corrispondere $X_{(i)}$ stimato è controversa.

Scartata la soluzione $p_i = i/n$ che è inadatta per un supporto infinito (raggiunge subito l'unità per i=n), ci si è orientati sulla seguente formula generale

$$p_i(\alpha,\beta) = \frac{i - \alpha}{n + 1 - (\alpha + \beta)}$$

Dove α e β sono tali da assicurare valori crescenti compresi tra zero ed uno.

I valori di questi parametri cambiano da distribuzione a distribuzione.

Spesso si sceglie α=β

Scelta delle posizioni grafiche

Name	α	β	$p_i(\alpha,\beta)$
0) *Hazen*	0.5	0.5	$\left(\frac{i-0.5}{n}\right)$
1) *Weibull*	0	0	$\left(\frac{i}{n+1}\right)$
2) *Bl* om	0.375	0.375	$\left(\frac{i-0.375}{n+0.25}\right)$
3) *Landwehr*	0.35	0.65	$\left(\frac{i-0.35}{n}\right)$
4) *Tukey*	$0.\underline{3333}$	$0.\underline{3333}$	$\left(\frac{i-0.\underline{3333}}{n+0.\underline{3333}}\right)$
5) *Cunnane*	0.4	0.4	$\left(\frac{i-0.4}{n+0.2}\right)$
6) *Benard*	0.3	0.3	$\left(\frac{i-0.3}{n+0.4}\right)$
7) *Filliben*	0.3175	0.3175	$\left(\frac{i-0.3175}{n+0.65}\right)$
8) *Gringorten*	0.44	0.44	$\left(\frac{i-0.44}{n+0.12}\right)$
9) *Larsen*	0.567	0.567	$\left(\frac{i-0.3175}{n-0.134}\right)$

La formula di Weibull gode di una certa popolarità.

Quella che si consiglia è la Landwehr

$p_i(\alpha,\beta)$ è una approssimazione dei quantili della distribuzione uniforme nell'intervallo unitario.

Grazie al teorema dell'inversione possiamo ottenere i quantili di moltissime distribuzioni a partire proprio da $p_i(\alpha,\beta)$

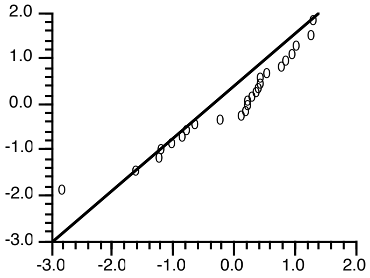
Esempio

Riprendiamo i dati dell'esempio

Si può notare un certo sbilanciamento al centro per i valori sotto la retta nonchè scarso adattamento nelle code.

L'ipotesi di normalità è dubbia per questi dati

5.9	7.1	5.5	1.5	4.6
-2.1	3.1	-0.5	4.8	-6.6
0.1	4.8	5.3	5.5	0.8
5.4	-0.7	5.0	6.9	7.5
8.6	8.8	7.8	1.0	4.4

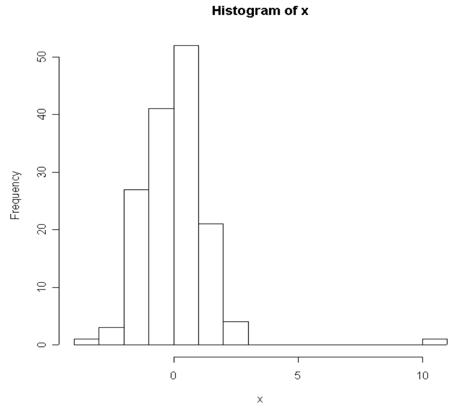


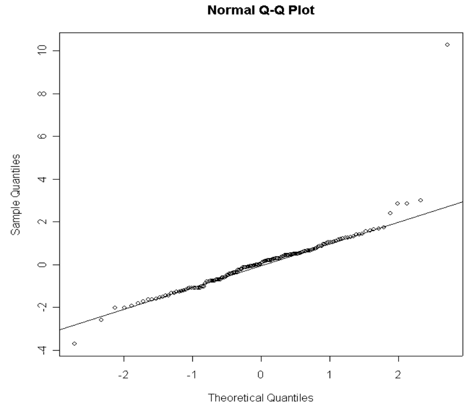
$p_i=(3i-1)/(3n+1)$

Il grafico Q-Q è un semplice ed utile ausilio per accertare l'aderenza alla Normale ovvero a qualsiasi distribuzione di cui siano noti i quantili.

Casi di disnormalità: valori remoti

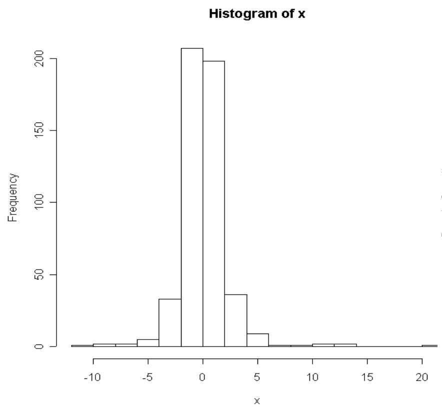
I casi di osservazioni estreme (alla luce del campione osservato) sono subito evidenti nel grafico QQ

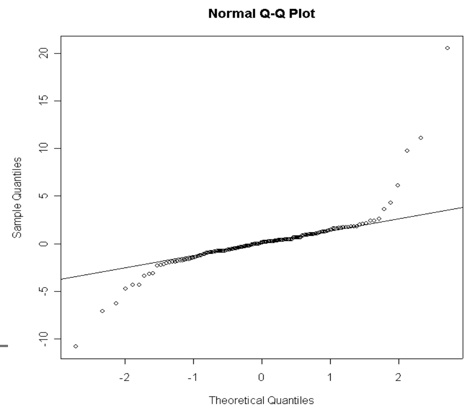




Casi di disnormalità: code pesanti

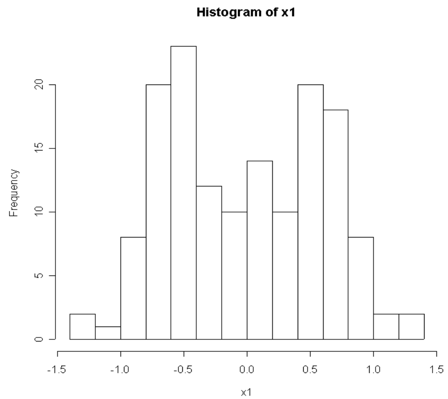
Le code pesanti si mostrano con la presenza di varie osservazioni per valori troppo bassi e/o troppo alti rispetto a ciò che ci si aspetta in un fenomeno normale

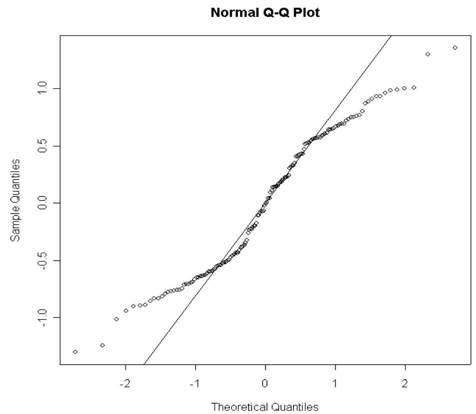




Casi di disnormalità: code smorzate

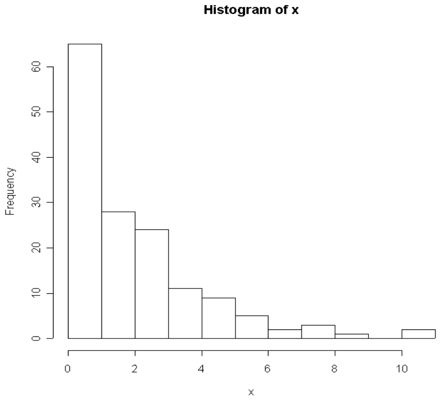
Le code smorzate si mostrano con l'assenza di osservazioni per valori bassi e/o alti laddove un fenomeno normale produrrebbe diversi valori sia pure con frequenza decrescente

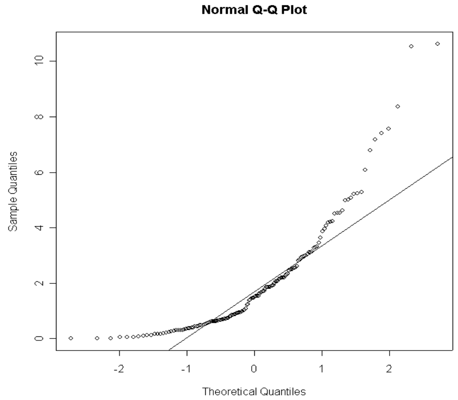




Casi di disnormalità: asimmetria

L'asimmetria della distribuzione si riscontra nella presenza di una forte curvatura del diagramma QQ.

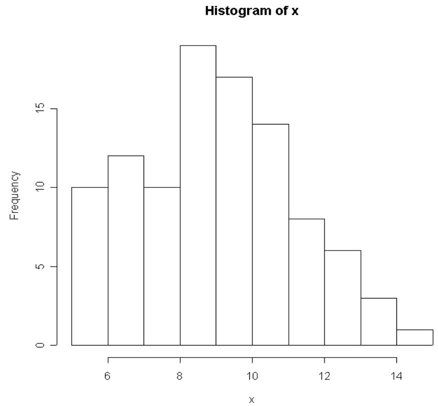


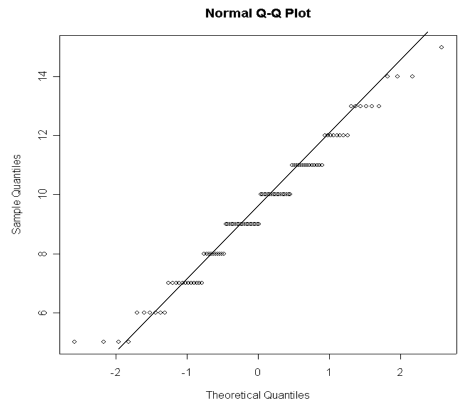


Casi di disnormalità: livellamenti

Il verificarsi di valori ripetuti (a causa di unità di misura troppo grezze) si traduce in una deviazione dalla normalità.

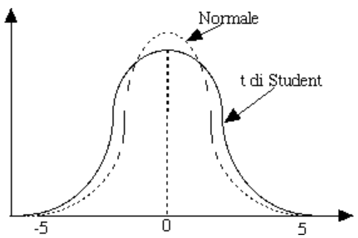
Anche la presenza di gap (vuoti) tra i valori è segno di disnormalità





La t di Student

Questo modello è molto simile a quello gaussiano, ma ha code più "pesanti" (ordinate estreme più alte)



$$-\infty < t_n < \infty, \quad n > 2$$

$$E(t_n) = 0, \quad Var(t_n) = \frac{n}{n-2}$$

La varianza è superiore all'unità, ma si avvicina ad uno all'aumentare di "n"

L'elemento caratterizzante della t di Student sono "i gradi di libertà" n che è il parametro della t di Student.

Per ogni grado di libertà esiste una t di Student, sebbene queste diventino poco distinguibili per n≥60.

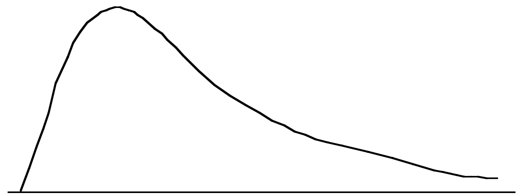
Questa v.c. è stata analizzata da W.S. Gosset, nel 1906, che firmò l'articolo con lo pseudonimo "STUDENT" ed è da allora nota come "La t di Student"

χ^2 (chi-quadrato)

Questo modello è definito per valori non negativi e presenta una marcata asimmetria positiva

Anche in questo caso l'elemento caratterizzante sono "i gradi di libertà" cioè "g"

$$E\left(\chi^2\right) = g, \qquad Var\left(\chi^2\right) = 2g$$



Per "g" superiore a 30 la distribuzione del χ^2 si avvicina a quella normale

La distribuzione del χ^2 si incontra nello studio dell'adattamento e della variabilità

Questa è importante nelle analisi cliniche e nel controllo della qualità