VARIABILI ALEATORIE CONTINUE

VARIABILI CONTINUE E DISTRIBUZIONE NORMALE

Prof. Rosario Lo Franco – Lezione 7

Riferimenti: [1] Sheldon M. Ross, *Introduzione alla statistica*, Apogeo Editore;
[2] Maria Garetto, *Statistica*, Università di Torino

Variabili aleatorie continue

Densità di probabilità

Se *X* è una variabile aleatoria continua, la probabilità che *X* assuma un certo valore *x* fissato è in generale zero. Nel caso di una variabile aleatoria continua ha senso calcolare la probabilità che X sia compresa fra *a* e *b*, dove *a* e *b* sono costanti, con *a* ≤ *b*.

$$f(x) \geq 0 \qquad \forall x \in \mathbf{R} \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

Si definisce poi la probabilità che X sia compresa fra a e b

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$

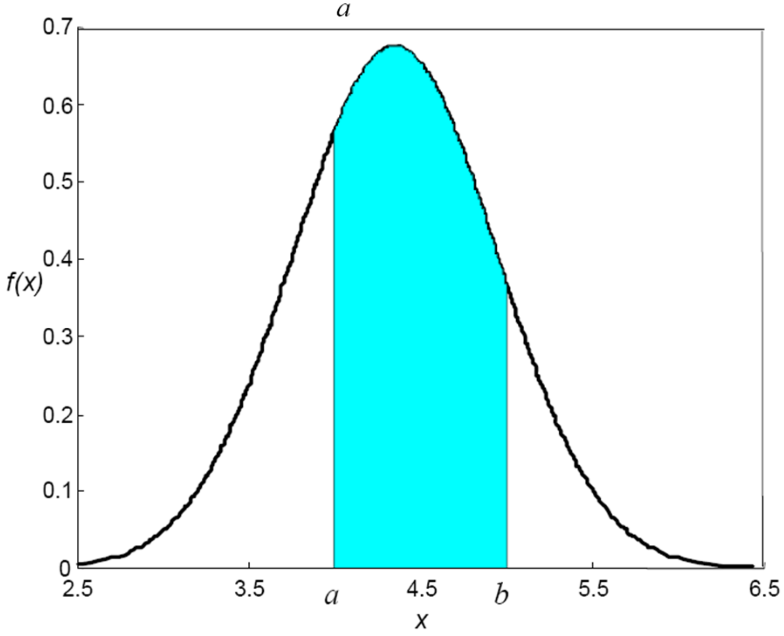
La funzione *f(x)* è detta densità di probabilità.

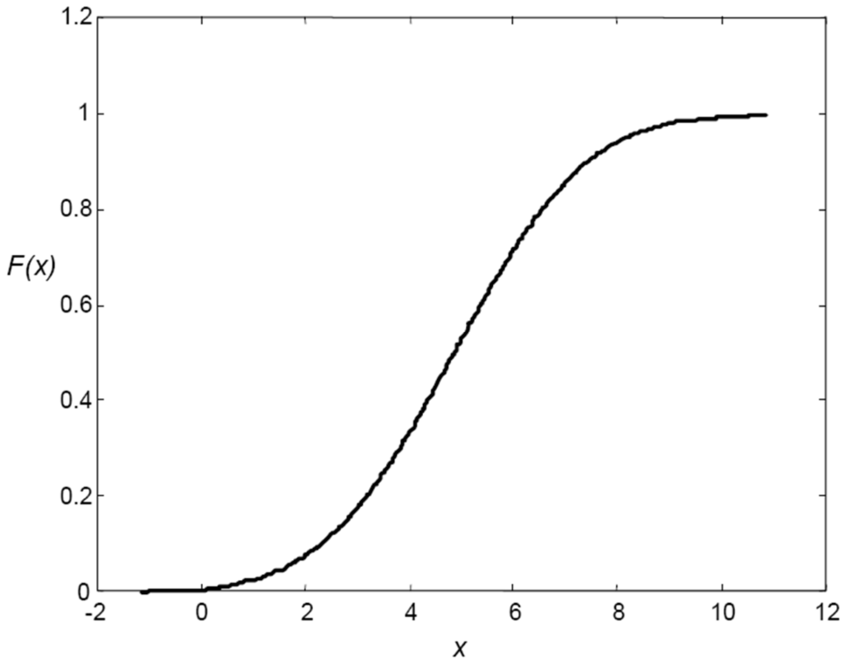
Variabili aleatorie continue

Si definisce **funzione di distribuzione** o **funzione di ripartizione** della variabile aleatoria continua X la funzione

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

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$





Variabili aleatorie continue

Se *X* è una variabile aleatoria continua, la probabilità che essa assuma un valore fissato è sempre zero

$$P(X = x) = 0 \qquad \forall x \in \mathbf{R}.$$

Nel continuo l'espressione "evento di probabilità nulla" non è sinonimo di "evento impossibile", come invece accade nel discreto.

Se *X* è una variabile aleatoria continua, allora

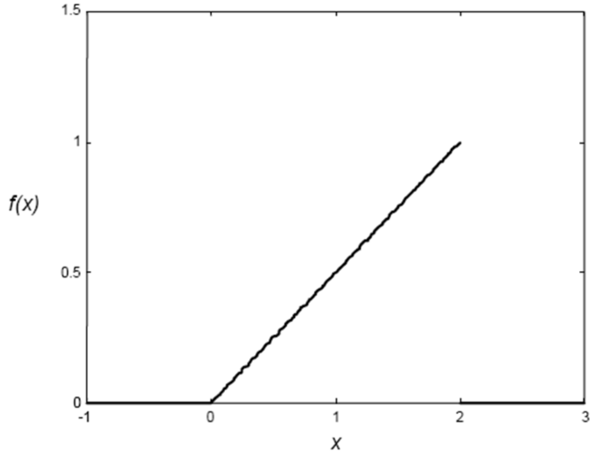
$$P(X \le a) = P(X < a)$$
$$P(X \ge a) = P(X > a)$$
$$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b)$$

Variabili aleatorie discrete e continue possono essere confrontate attraverso le rispettive funzioni di distribuzione (o di ripartizione).

Variabili aleatorie continue

Esempio

$$f(x)=\begin{cases}\dfrac{1}{2}x & 0\le x\le 2\\ 0 & \text{altrimenti}\end{cases}$$



Si può verificare che $f(x)$ è una densità di probabilità

$$f(x)\ge 0 \qquad \forall x\in\mathbf{R} \qquad \int_{-\infty}^{\infty} f(x)dx=\int_0^2 \frac{1}{2}x\,dx=\frac{1}{2}\frac{x^2}{2}\Big|_0^2=1$$

Troviamo la funzione di distribuzione

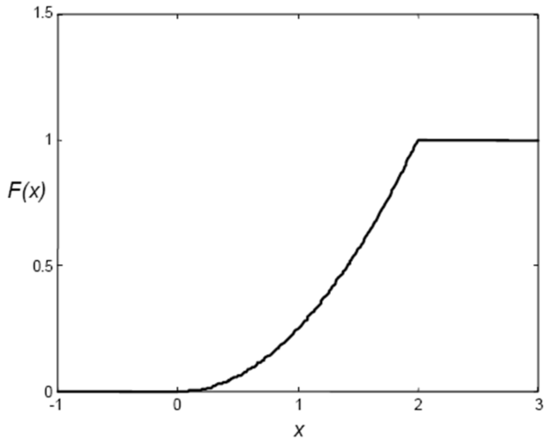
Per $x<0$ $\qquad F(x)=0$

Per $0\le x\le 2$

$$F(x)=\int_{-\infty}^{x} f(t)\,dt=\int_0^x f(t)\,dt=\int_0^x \frac{1}{2}t\,dt=\frac{1}{4}x^2$$

Per $x>2$

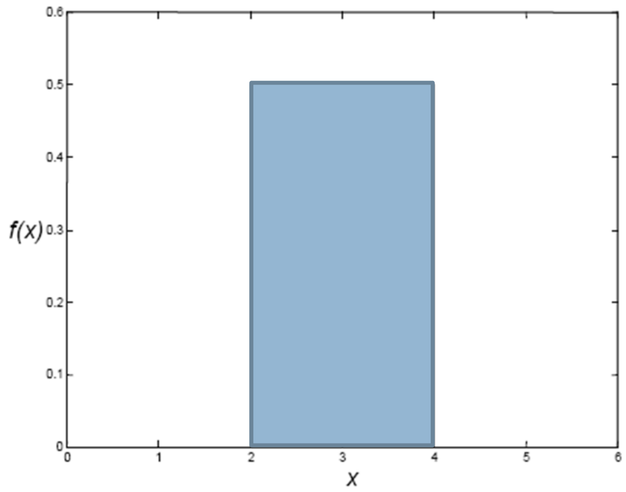
$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_0^2 \frac{1}{2}tdt=1$$



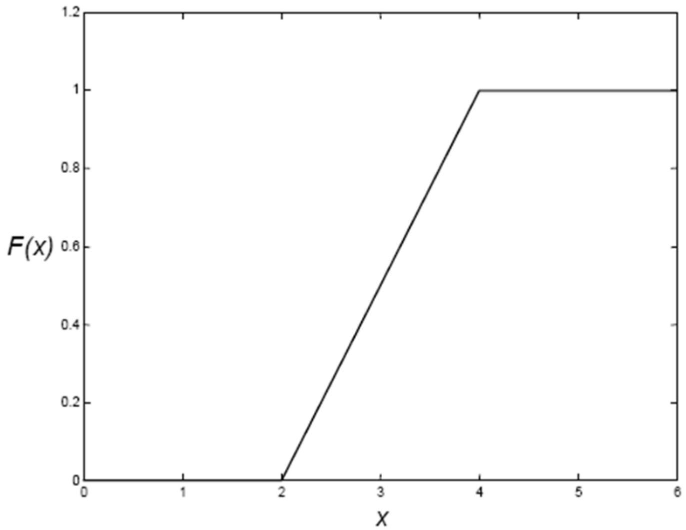
Variabili aleatorie continue

Distribuzione uniforme

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



$$F(x) = P(X \le x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases}$$



Variabili aleatorie continue

Valor medio e varianza

Sia X una variabile aleatoria continua avente densità di probabilità $f(x)$.

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \sigma^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

Esempio

Distribuzione uniforme

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{altrimenti} \end{cases} \qquad \mu = \int_a^b \frac{x}{b-a} dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{a+b}{2}$$

$$\sigma^2 = \int_a^b \frac{x^2}{b-a} dx - \mu^2 = \left.\frac{x^3}{3(b-a)}\right|_a^b - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

Distribuzione normale o di Gauss

- Anche nota come **legge degli errori**, in quanto essa descrive in particolare la distribuzione degli errori casuali relativi a successive misure di una quantità fisica

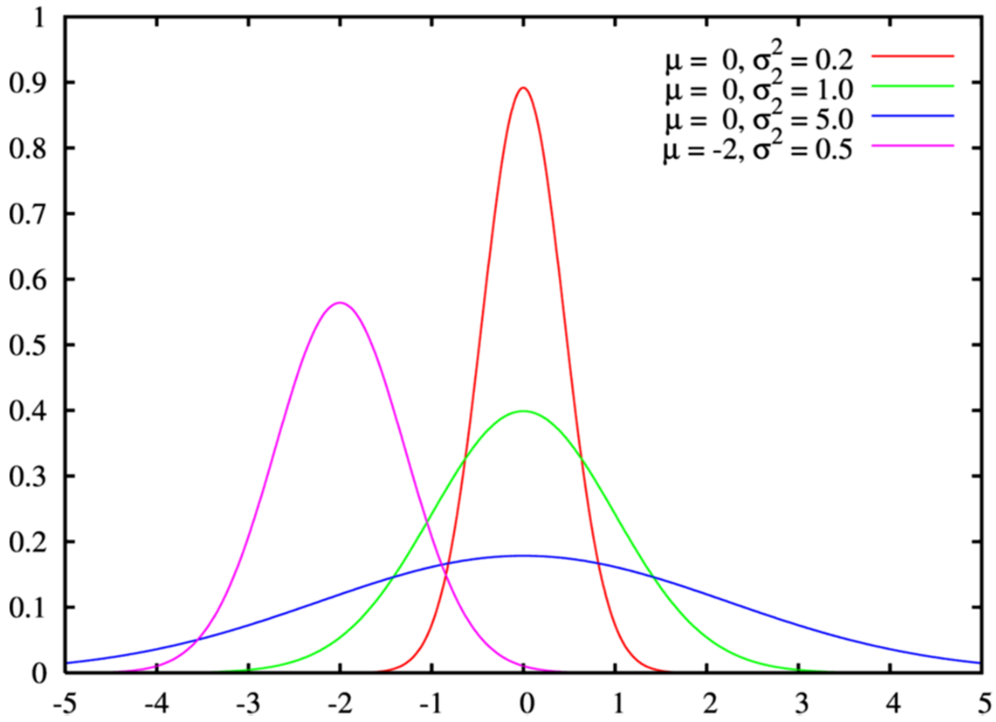
- La distribuzione normale è importante in statistica per tre motivi fondamentali:

i) diversi fenomeni continui seguono, almeno approssimativamente, una distribuzione normale;
ii) la distribuzione normale può essere utilizzata per approssimare numerose distribuzioni di probabilità discrete;
iii) la distribuzione normale è alla base dell'inferenza statistica, in virtù del teorema del limite centrale.

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

μ e σ sono rispettivamente il valor medio e lo scarto quadratico medio

Distribuzione normale o di Gauss



Il valore di tutta l'area sottesa dalla curva f(x) è uguale a 1.

Il valore massimo della funzione viene assunto nel punto di ascissa μ ed è $y_{max} = \dfrac{1}{\sigma\sqrt{2\pi}}$

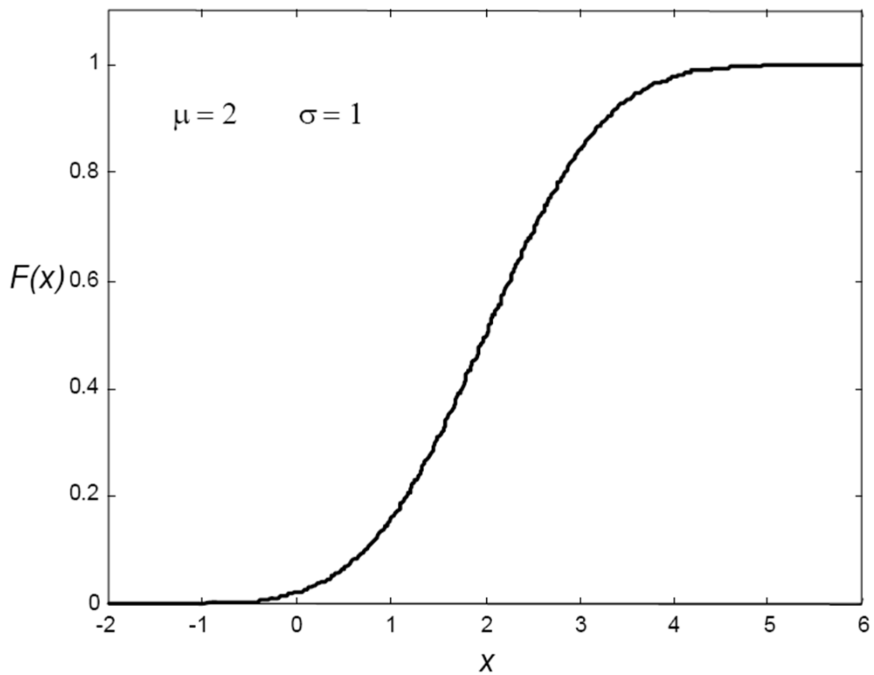
Lo scarto quadratico medio σ è uguale alla distanza dei punti di flesso da μ, ossia i punti di flesso hanno ascissa rispettivamente $\mu - \sigma$ e $\mu + \sigma$.

Distribuzione normale o di Gauss

La **funzione di distribuzione** o **funzione di ripartizione normale** è data da

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

il grafico della funzione di distribuzione $F(x)$ per $\mu = 2$ e $\sigma = 1$



Distribuzione normale standard

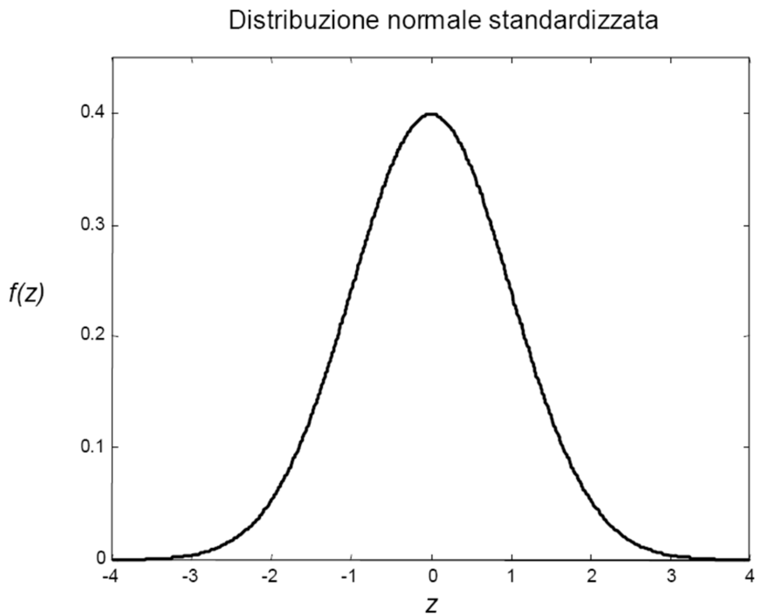
variabile standardizzata

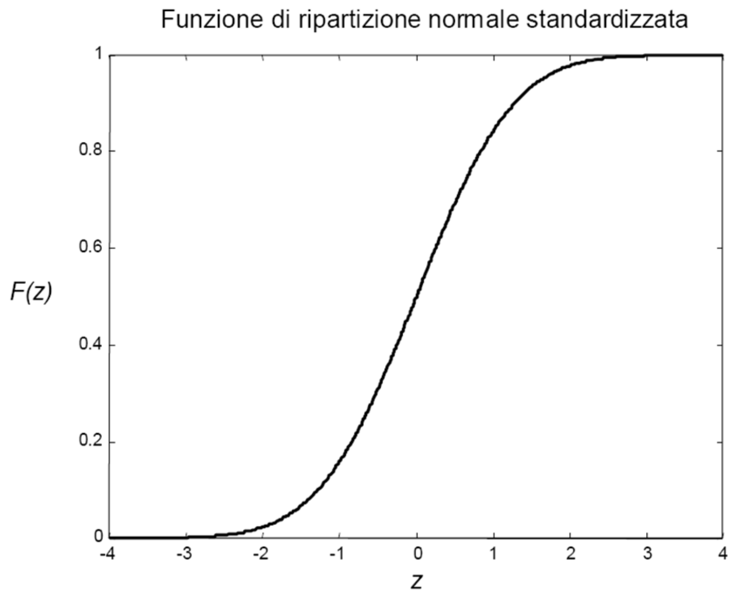
$$\boxed{Z = \frac{X-\mu}{\sigma}}$$

La media di *Z* è 0 e la varianza è 1.

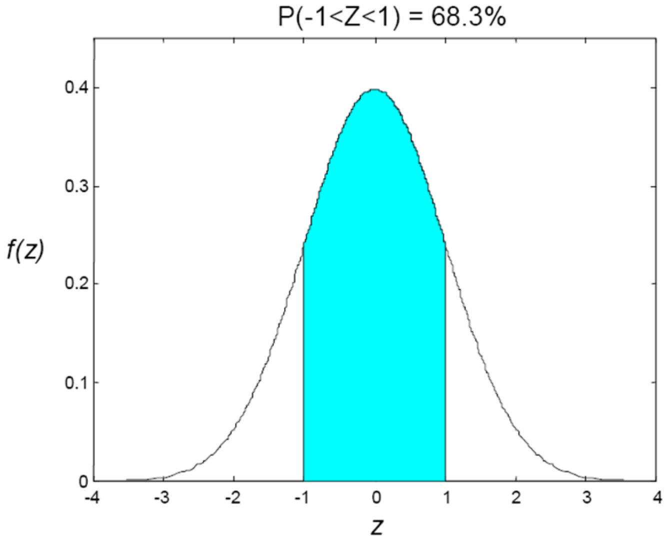
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad -\infty < z < \infty$$

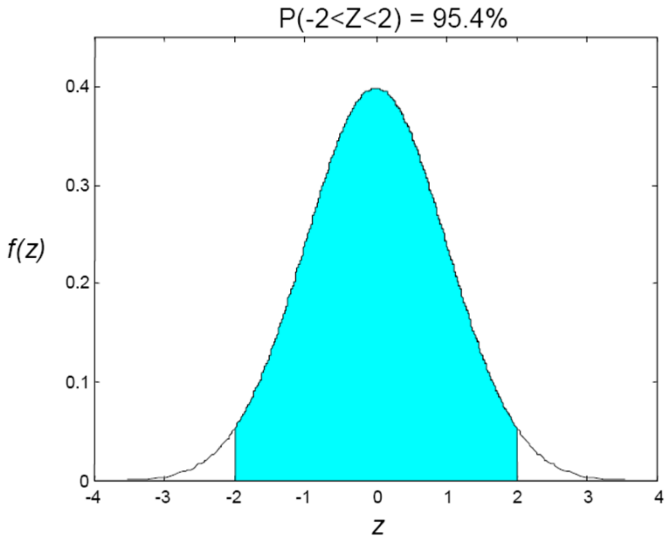
$$F(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$

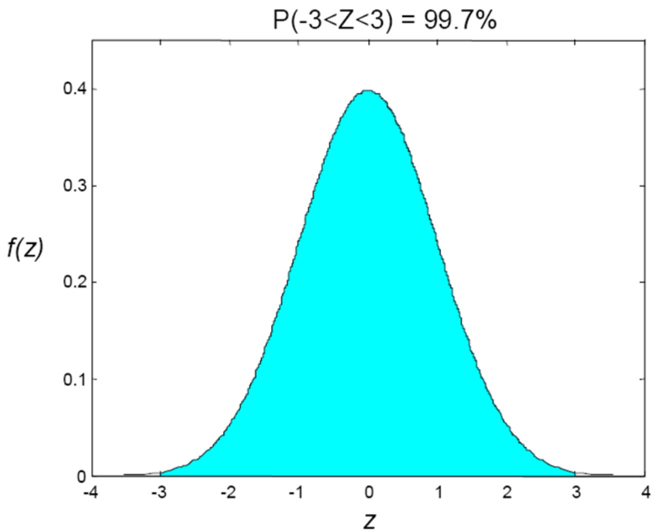




Distribuzione normale standard







$$P(-1 < Z < 1) = 0.6827 \cong 68.3\%$$

$$P(-2 < Z < 2) = 0.9544 \cong 95.4\%$$

$$P(-3 < Z < 3) = 0.9973 \cong 99.7\%$$

$$P(\mu - \sigma < X < \mu + \sigma) \cong 68.3\%$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) \cong 95.4\%$$

$$P(\mu - 3\sigma < X < \mu + 3\sigma) \cong 99.7\%$$

Tavola della funzione di ripartizione della normale standard

$$P(-\infty < Z < \infty) = 1$$

$$P(-\infty < Z < 0) = P(0 < Z < \infty) = F(0) = \frac{1}{2}$$

$$P(Z \le -z) = F(-z) = 1 - F(z)\,, \quad z > 0$$

$$P(z_1 \le Z \le z_2) = F(z_2) - F(z_1)$$

$$P(-z_1 \le Z \le 0) = P(0 \le Z \le z_1)$$

Tabella A.1 Funzione di ripartizione della distribuzione normale standard

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy = P(Z \le x)$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	$\boxed{0.9750}$	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Distribuzione normale standard

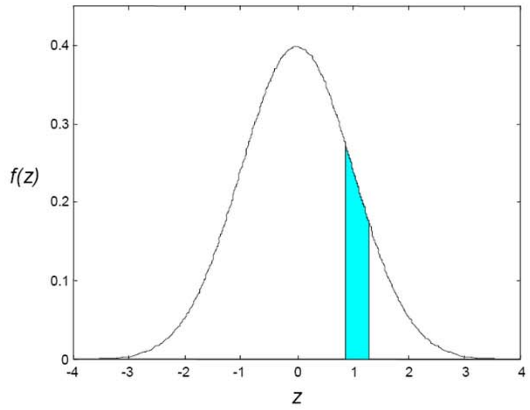
Esempio

a – $0.87 \le Z \le 1.28$

b – $-0.34 \le Z \le 0.62$

c – $|Z| > 1.2$

d – $Z \ge -0.65$

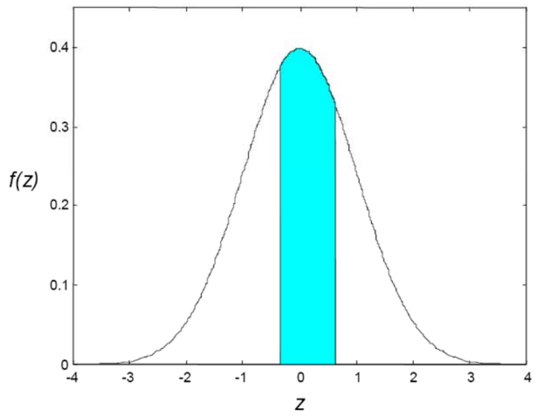


a –

$$P(0.87 \le Z \le 1.28) = P(Z \le 1.28) - P(Z \le 0.87) =$$
$$= F(1.28) - F(0.87) = 0.8997 - 0.8078 = 0.0919$$

b –

$$P(-0.34 \le Z \le 0.62) = F(0.62) - F(-0.34) =$$
$$= 0.7324 - [1 - F(0.34)] = 0.7324 - 1 + 0.6331 = 0.3655$$



c –

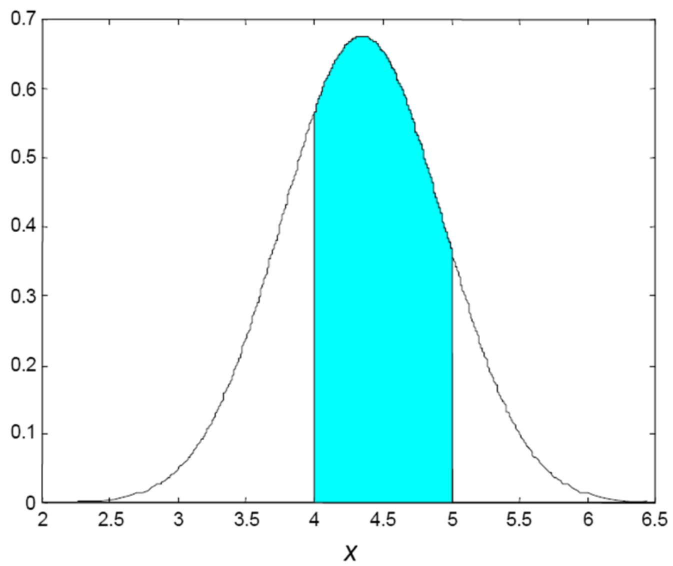
$$P(|Z| > 1.2) = P(Z > 1.2) + P(Z < -1.2) = 1 - F(1.2) + F(-1.2) =$$
$$= 1 - F(1.2) + 1 - F(1.2) = 2 - 2 \cdot F(1.2) = 2 - 2 \cdot 0.8849 = 0.2302$$

Distribuzione normale standard

Esempio Sia X una variabile aleatoria avente distribuzione normale, con $\mu = 4.35$ e $\sigma = 0.59$; trovare la probabilità $P(4 \leq X \leq 5)$

$$X = 4 \;\Rightarrow\; Z = \frac{4 - 4.35}{0.59} = -0.5932 \qquad\qquad X = 5 \;\Rightarrow\; Z = \frac{5 - 4.35}{0.59} = 1.1017$$

$$P(4 \leq X \leq 5) = P(-0.59 \leq Z \leq 1.10) = F(1.10) - F(-0.59) =$$
$$= 0.8643 - 1 + F(0.59) = 0.5867$$



Distribuzione normale standard

Esempio

L'altezza di un gruppo di ragazzi è distribuita normalmente con media μ = 174 cm e scarto quadratico medio σ = 15 cm. Calcolare la probabilità che un ragazzo scelto a caso abbia una statura superiore a 190 cm.

$$X = 190 \quad \Rightarrow \quad Z = \frac{190 - 174}{15} \cong 1.07$$

$$P(Z > 1.07) = 1 - P(Z < 1.07) = 1 - F(1.07) =$$
$$= 1 - 0.8577 = 0.1423 = 14.23\%$$

Distribuzione normale standard

Esempio

Il diametro effettivo delle sfere di acciaio prodotte da una ditta può essere considerato una variabile aleatoria normale di media $\mu = 5.1$ cm e scarto quadratico medio $\sigma = 0.1$ cm.

a – Calcolare la probabilità che il diametro di una sfera scelta a caso sia compreso tra 5.0 e 5.2 cm.

b – Calcolare la stessa probabilità, supponendo che lo scarto quadratico medio sia $\sigma = 0.5$ cm.

a – $X = 5.0 \Rightarrow Z = \dfrac{5.0 - 5.1}{0.1} = -1$

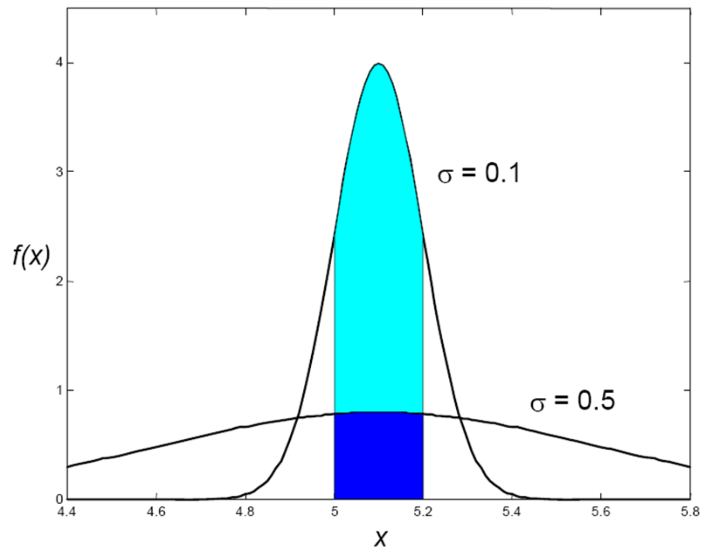
$X = 5.2 \Rightarrow Z = \dfrac{5.2 - 5.1}{0.1} = 1$

$$P(5.0 \le X \le 5.2) = P(-1 \le Z \le 1) = 2[P(Z \le 1) - 0.5] =$$
$$= 2(0.8413 - 0.5) = 0.6826 \cong 68\%$$

b – $X = 5.0 \Rightarrow Z = \dfrac{5.0 - 5.1}{0.5} = -0.2$

$X = 5.2 \Rightarrow Z = \dfrac{5.2 - 5.1}{0.5} = 0.2$

$$P(5.0 \le X \le 5.2) = P(-0.2 \le Z \le 0.2) = 2[P(Z \le 0.2) - 0.5] =$$
$$= 2(0.5793 - 0.5) = 0.1586 \cong 16\%$$

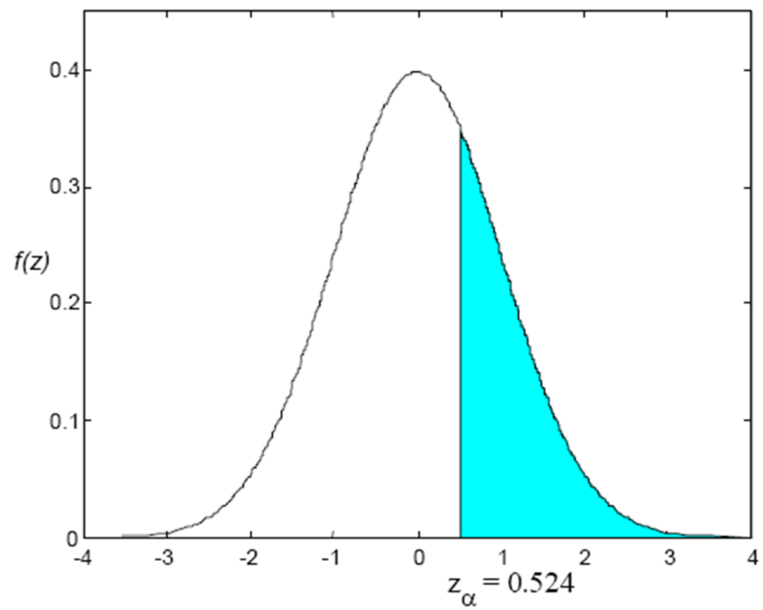


Distribuzione normale standard
Percentili

$$P(Z > z_\alpha) = \alpha$$

z_α è il $100(1-\alpha)$-mo percentile e rappresenta il valore per cui la probabilità che Z sia maggiore di z_α è uguale ad α.

Ad esempio, il valore di z_α per il quale il 30% dei valori di Z cade a destra di z_α è z_α = 0.524.



Distribuzione normale standard
Percentili

Esempio

La variabile aleatoria Z ha la distribuzione normale standardizzata. Determinare il valore di z_α per cui

$$P(Z < z_\alpha) = 0.9953 \Rightarrow P(Z < 2.6) = 0.9953 \Rightarrow z_\alpha = 2.6$$

$$P(Z > z_\alpha) = 0.2743 \Rightarrow P(Z < z_\alpha) = 1 - P(Z > z_\alpha) = 1 - 0.2743 = 0.7257 \Rightarrow z_\alpha = 0.6$$

$$P(|Z| < z_\alpha) = 0.5762 \Rightarrow \begin{aligned} &P(|Z| < z_\alpha) = P(-z_\alpha < Z < z_\alpha) = 2 \cdot P(0 < Z < z_\alpha) = \\ &= 2 \cdot [P(Z < z_\alpha) - 0.5] = 2 \cdot P(Z < z_\alpha) - 1 = 0.5762 \end{aligned}$$

$$\Rightarrow P(Z < z_\alpha) = \frac{1 + 0.5762}{2} = 0.7881 \Rightarrow z_\alpha = 0.8$$

Relazione tra distribuzione binomiale e distribuzione normale

Sia X la variabile aleatoria che fornisce il numero di successi in n prove e p la probabilità di successo in una singola prova: **quando il numero *n* delle prove è grande**, il calcolo con la distribuzione binomiale è molto lungo. In tal caso **è possibile utilizzare la distribuzione normale per approssimare la distribuzione binomiale**.

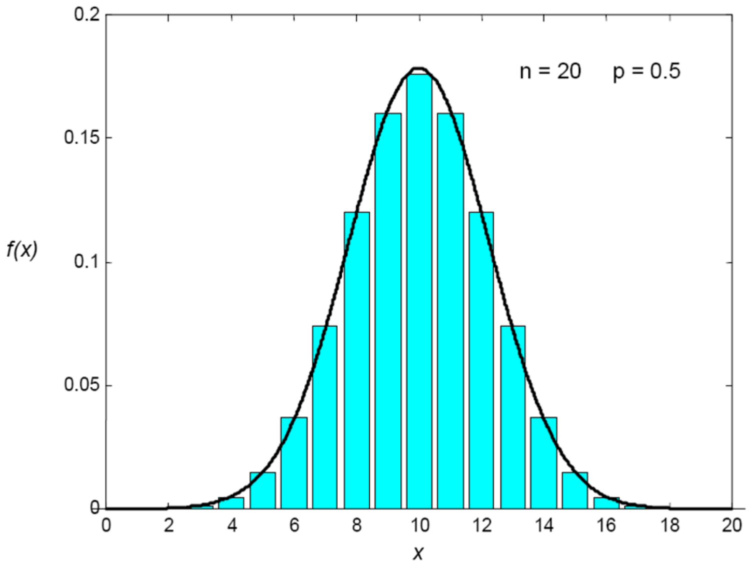
Come regola pratica si usa la distribuzione normale per approssimare la binomiale se si verificano entrambe le condizioni $np \geq 5$ e $n(1-p) \geq 5$.
La regola suggerita è soddisfatta se n è abbastanza grande e l'approssimazione è tanto più precisa quanto più p è vicina a 0.5.

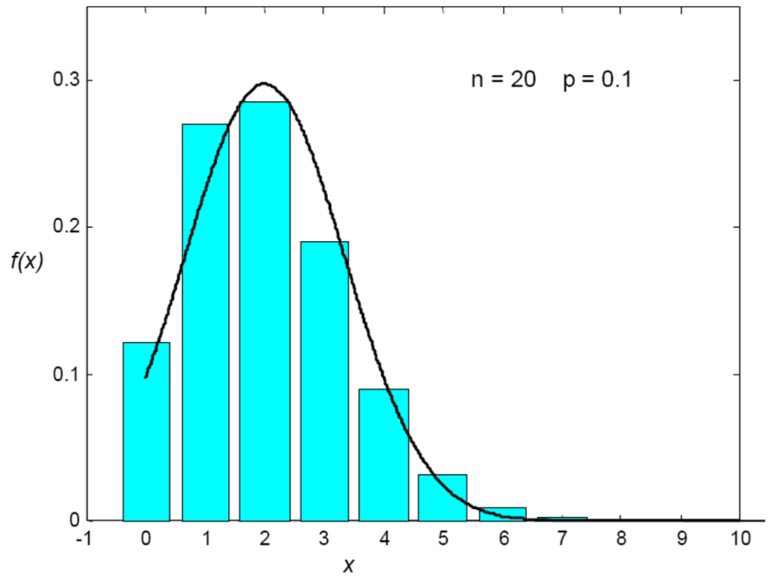
$$\mu = np \qquad \sigma^2 = np(1-p) \quad \Rightarrow \quad Z = \frac{X - np}{\sqrt{np(1-p)}}$$

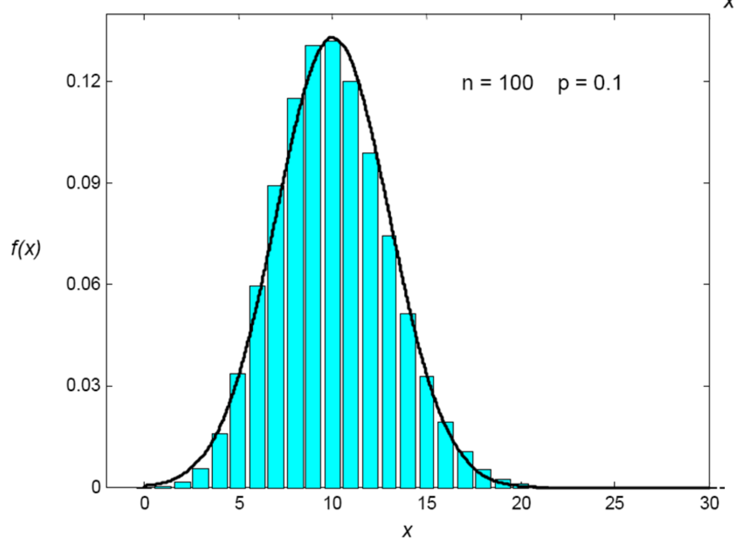
Correzione di continuità: ogni valore intero x assunto dalla variabile aleatoria discreta si rappresenta con l'intervallo di estremi $x - ½$ e $x + ½$.

Quindi: $P(a \leq X \leq b) \Rightarrow P(a - ½ \leq X \leq b + ½)$; $P(X = a) \Rightarrow P(a - ½ \leq X \leq a + ½)$

Relazione tra distribuzione binomiale e distribuzione normale







Relazione tra distribuzione binomiale e distribuzione normale

Esempio

Trovare la probabilità che in 100 lanci di una moneta, testa si presenti 40 volte, usando la distribuzione normale per approssimare la distribuzione binomiale.

$$P\left(40 - \frac{1}{2} \leq X \leq 40 + \frac{1}{2}\right) = P(39.5 \leq X \leq 40.5)$$

$$\mu = np = 100 \cdot \frac{1}{2} = 50 \qquad \sigma^2 = np(1-p) = 100 \cdot \frac{1}{2} \cdot \frac{1}{2} = 25$$

$$X = 39.5 \quad \Rightarrow \quad Z = \frac{39.5 - 50}{\sqrt{25}} = -2.1$$

$$X = 40.5 \quad \Rightarrow \quad Z = \frac{40.5 - 50}{\sqrt{25}} = -1.9$$

$$P(-2.1 < Z < -1.9) = P(1.9 < Z < 2.1) =$$
$$= P(Z < 2.1) - P(Z < 1.9) = 0.9821 - 0.9713 = 0.0108$$
$$P(X = 40) \cong 0.0108$$

Relazione tra distribuzione binomiale e distribuzione normale

Esempio

Si effettuano 500 lanci di una moneta; calcolare la probabilità che il numero di teste non differisca da 250

a – per più di 10;

b – per più di 30.

$$n = 500 \qquad p = \frac{1}{2} \qquad \mu = np = 250$$

a – $P(239.5 < X < 260.5)$

$$\sigma = \sqrt{np(1-p)} = \sqrt{500 \frac{1}{2}\frac{1}{2}} = 11.18$$

$$X = 239.5 \quad \Rightarrow \quad Z = \frac{239.5 - 250}{11.18} = -0.94$$

$$X = 260.5 \quad \Rightarrow \quad Z = \frac{260.5 - 250}{11.18} = 0.94$$

$$P(-0.94 < Z < 0.94) = P(Z < 0.94) - [1 - P(Z < 0.94)] =$$
$$= 2 \cdot 0.8264 - 1 = 0.6528 \cong 65.3\%$$

$$P(240 \le X \le 260) \cong 0.6528$$

Relazione tra distribuzione binomiale e distribuzione normale

Esempio

Il 20% dei chip di memoria prodotti da un'azienda di componenti elettronici è difettoso; calcolare la probabilità che in un campione di 100 chip scelto a caso per un controllo

a – al più 15 siano difettosi;

b – esattamente 15 siano difettosi.

a – $P(X \leq 15)$ $\quad np = 100 \cdot 0.2 = 20 \quad n(1-p) = 100 \cdot 0.8 = 80$

$$\mu = np = 100 \cdot 0.2 = 20 \qquad \sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.2 \cdot 0.8} = 4$$

$$X = 15.5 \quad \Rightarrow \quad Z = \frac{15.5 - 20}{4} = -1.13$$

$$P(Z < -1.13) = 1 - P(Z < 1.13) = 1 - 0.8708 = 0.1292$$

$$P(X \leq 15) \cong 0.1292$$