Facoltà di Ingegneria
Università degli Studi di Firenze
Dipartimento di Elettronica e Telecomunicazioni

Statistica e Probabilità per l’Ingegneria

- Variabili aleatorie - (3^ parte)

3. Distribuzioni di Probabilità continue

Ing. Andrea Zanobini
Dipartimento di Elettronica e Telecomunicazioni

Distribuzione Normale o di Gauss (1777-1855)

La distribuzione è 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 (vedere § 5.3).
La distribuzione normale è importante in statistica per tre motivi fondamentali:

1 – diversi fenomeni continui seguono, almeno approssimativamente, una distribuzione normale;

2 – la distribuzione normale può essere utilizzata per approssimare numerose distribuzioni di probabilità discrete;

3 – la distribuzione normale è alla base dell'inferenza statistica, in virtù del teorema del limite centrale, che sarà discusso nel capitolo 6.

Definizione 1
La **densità di probabilità normale**, o **distribuzione normale** o **di Gauss**, è definita dalla funzione

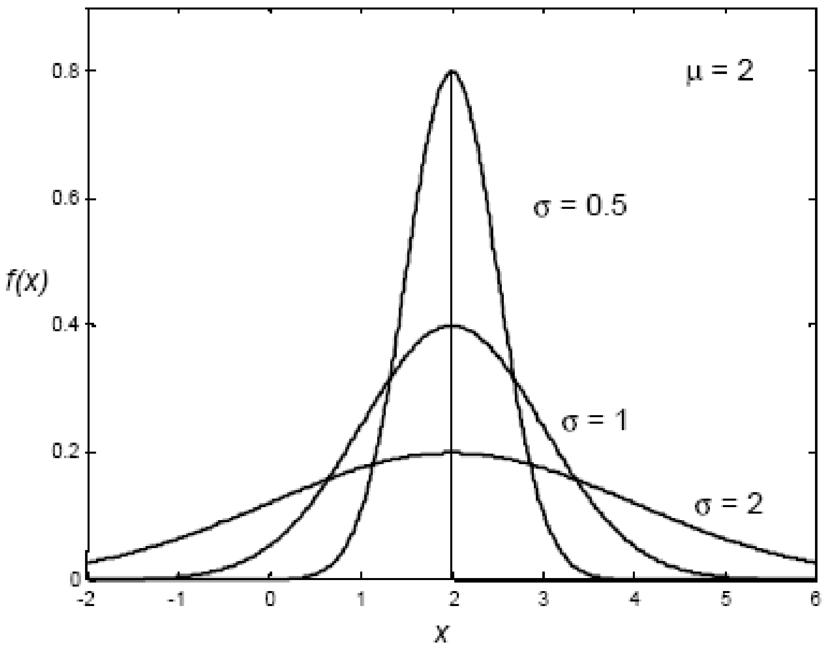
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty \qquad (5.1)$$

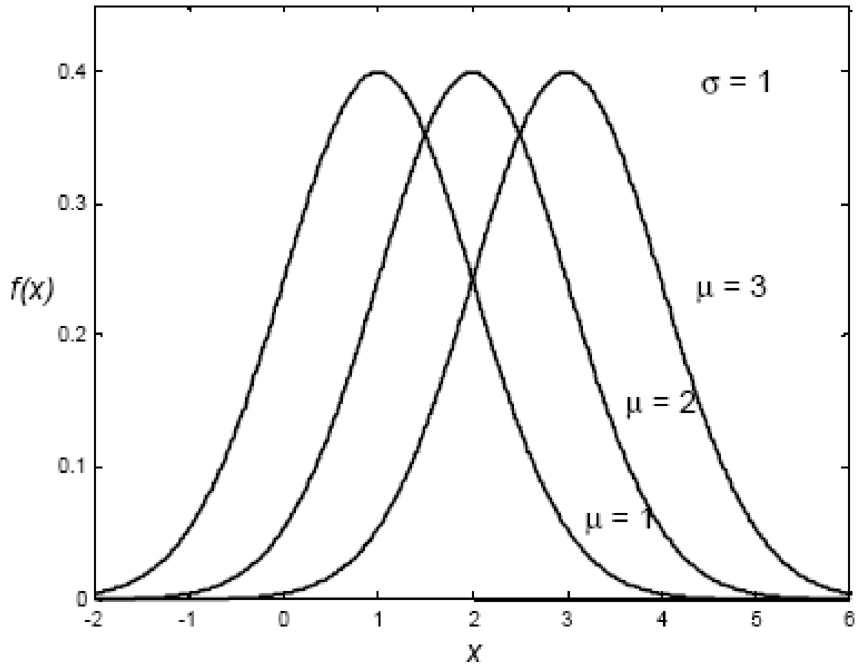
di parametri μ e σ , con $\sigma > 0$.

Si dimostra che μ e σ sono rispettivamente il valor medio e lo scarto quadratico medio della variabile aleatoria X distribuita secondo la distribuzione normale.

Le caratteristiche più importanti della distribuzione normale sono le seguenti.
La funzione $f(x)$ è definita su tutto l'asse reale e assume valori sempre positivi; è simmetrica rispetto alla retta $x = \mu$, cioè rispetto al valor medio della distribuzione. La moda e la mediana coincidono con il valor medio.

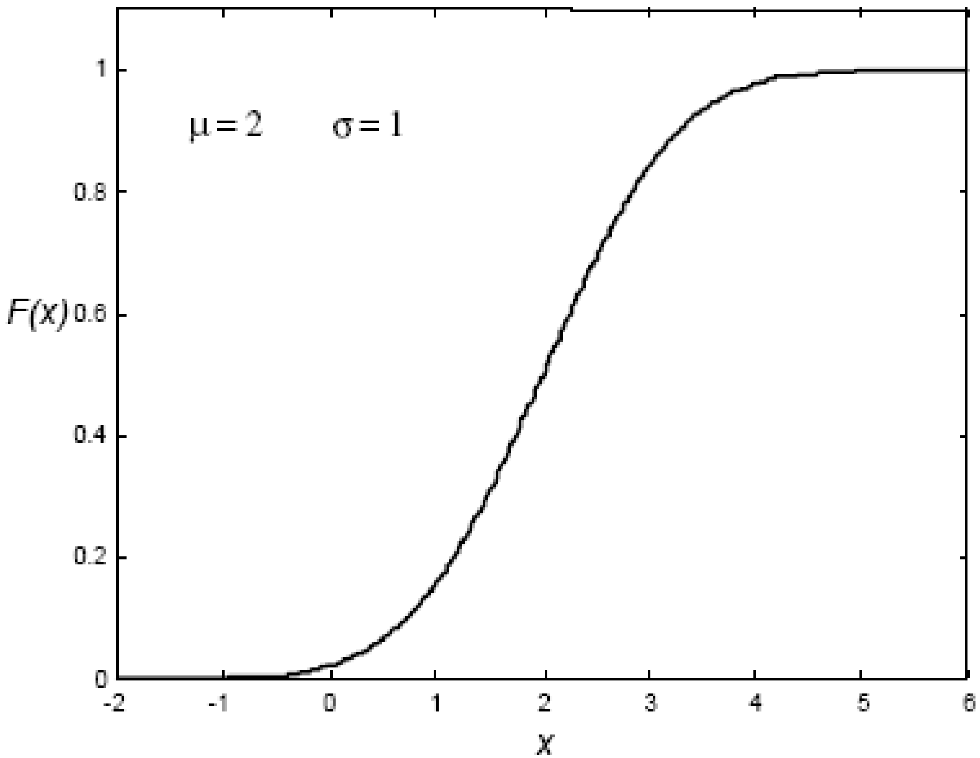
Il valore massimo della funzione viene assunto nel punto di ascissa μ ed è $y_{max} = \dfrac{1}{\sigma\sqrt{2\pi}}$; questo valore è perciò inversamente proporzionale a σ.
Lo scarto quadratico medio σ è uguale alla distanza dei punti di flesso da μ, ossia i punti di flesso hanno ascissa rispettivamente $\mu - \sigma$ e $\mu + \sigma$.





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

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \qquad -\infty < x < \infty \qquad (5.2)$$



Distribuzione normale standardizzata

Z, detta **variabile standardizzata**, ponendo

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

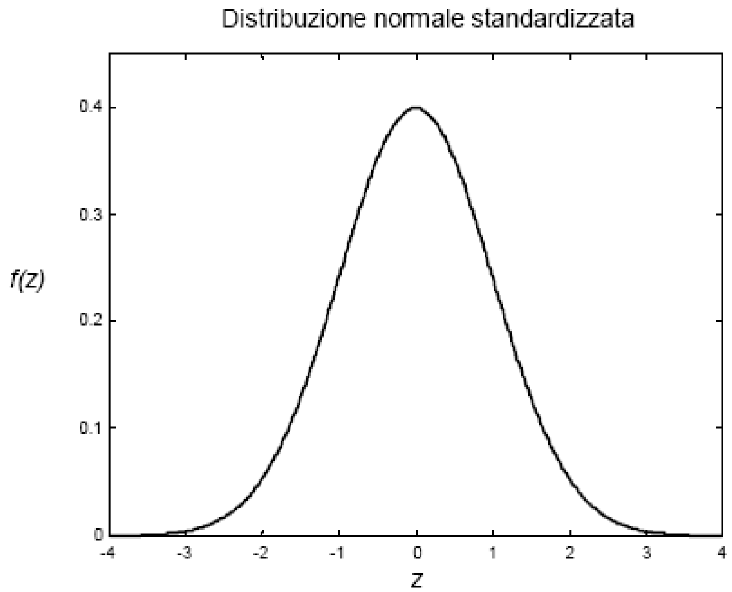
La trasformazione operata fa in modo che la media di Z sia 0 e la varianza 1.

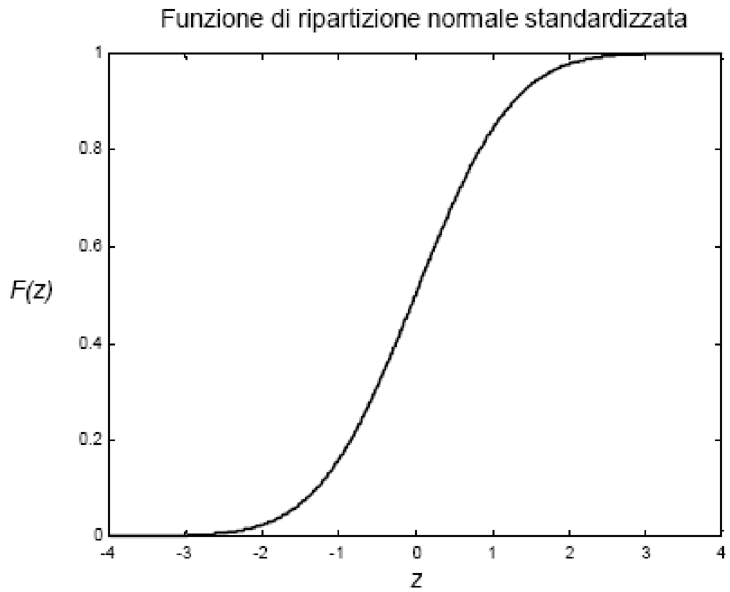
La **distribuzione di probabilità della variabile normale standardizzata** Z è data da

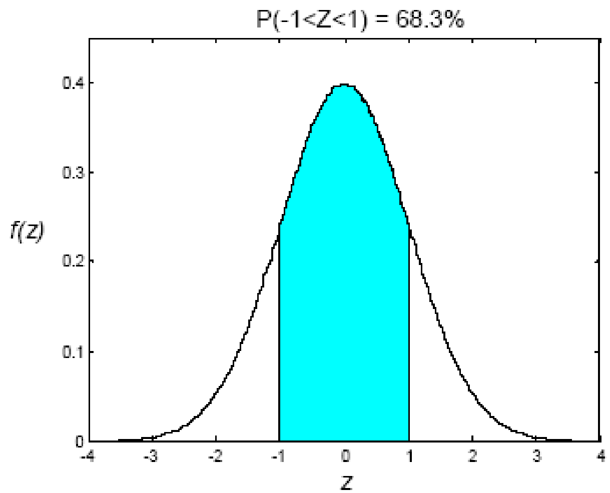
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad -\infty < z < \infty \qquad (5.3)$$

La **funzione di distribuzione** o **di ripartizione della variabile normale standardizzata** Z è data da

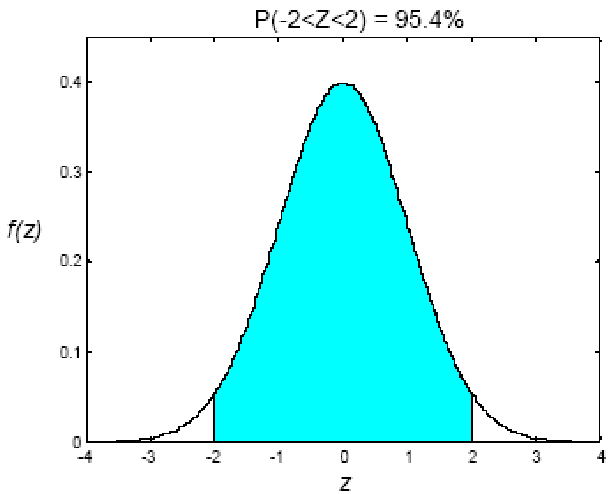
$$F(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt \qquad -\infty < z < \infty \qquad (5.4)$$



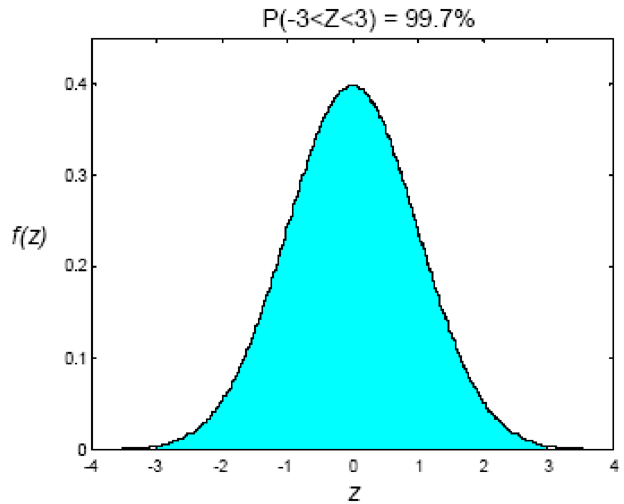




$$P(-1 \le Z \le 1) = 0.6827 \cong 68.3\%$$



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



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

Alcune applicazioni della distribuzione normale

1 – Curva degli errori casuali nella misurazione di una grandezza fisica.

2 – Distribuzione di una caratteristica quantitativa di una popolazione, che presenta oscillazioni casuali attorno a una media.

3 – Dimensione effettiva di oggetti prodotti in serie, che si cerca di produrre in modo identico.

Uso delle tavole della distribuzione normale

Uso delle tavole → anche MINITAB !!

$$F(z) = P(Z \le z)$$

Proprietà

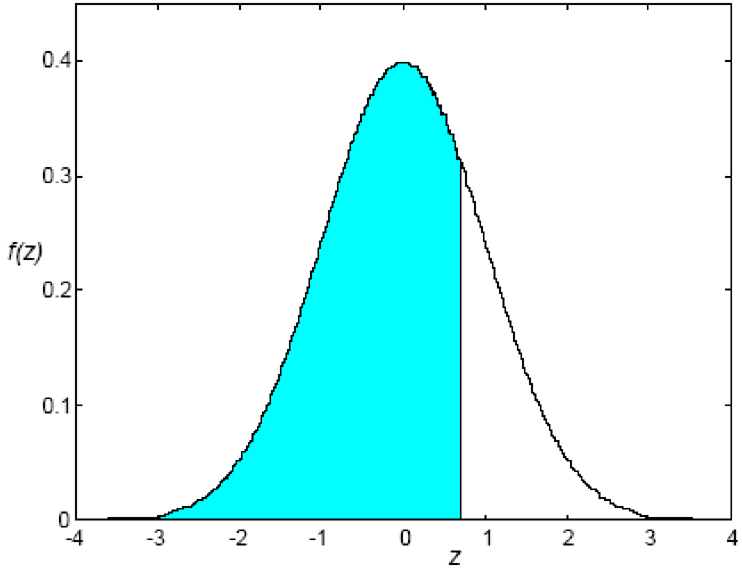
$$P(-\infty < Z < \infty) = 1$$

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

$$P(Z \le -z) = F(-z) = 1 - F(z)$$

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

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



Esempio 1

Calcolare, usando la tavola della distribuzione normale standardizzata, la probabilità che una variabile aleatoria Z avente la distribuzione normale standardizzata assuma valori tali che[1]

a – $0.87 \le Z \le 1.28$

b – $-0.34 \le Z \le 0.62$

c – $Z \ge 0.85$

d – $Z \ge -0.65$

$$\boxed{\begin{aligned} &P(0.87 \le Z \le 1.28) = P(Z \le 1.28) - P(Z \le 0.87) = \\ &\quad = F(1.28) - F(0.87) = 0.8997 - 0.8078 = 0.0919 \\ &P(-0.34 \le Z \le 0.62) = F(0.62) - F(-0.34) = \\ &\quad = 0.7324 - [1 - F(0.34)] = 0.7324 - 1 + 0.6331 = 0.3655 \\ &P(Z \ge 0.85) = 1 - P(Z \le 0.85) = 1 - F(0.85) = 1 - 0.8023 = 0.1977 \\ &P(Z \ge -0.65) = P(Z \le 0.65) = F(0.65) = 0.7422 \end{aligned}}$$

$|Z| > 1.2$

$$\boxed{\begin{aligned} &P(|Z| > 1.2) = P(Z > 1.2) + P(Z < -1.2) = 1 - F(1.2) + F(-1.2) = \\ &\quad = 1 - F(1.2) + 1 - F(1.2) = 2 - 2 \cdot F(1.2) = 2 - 2 \cdot 0.8849 = 0.2302 \end{aligned}}$$

Esempio 3

Sia X una variabile aleatoria avente distribuzione normale, con $\mu = 4.35$ e $\sigma = 0.59$; trovare la probabilità $P(4 \le X \le 5)$ (figura 13).

Con il cambiamento di variabile $Z = \dfrac{X-\mu}{\sigma}$ si passa alla variabile standardizzata

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

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

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

Esempio 4

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

$$\mu = 174 \qquad \sigma = 15$$

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

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

Esempio 5

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.

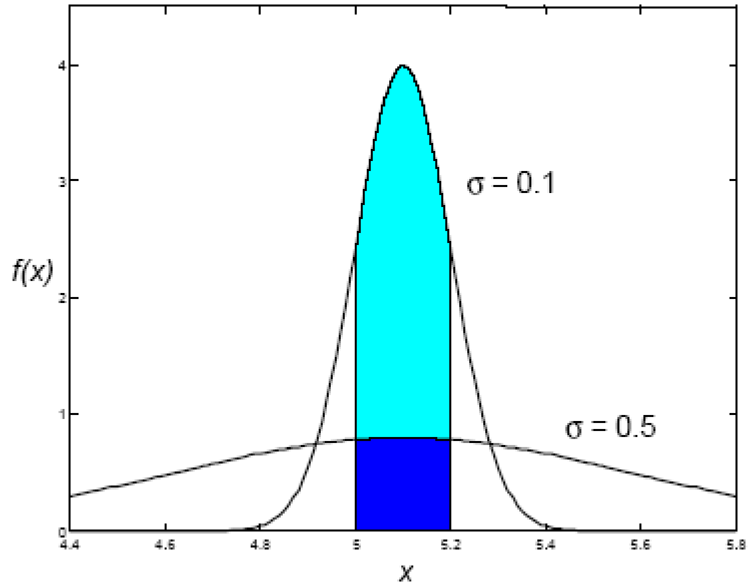
b – Calcolare la stessa probabilità, supponendo che lo scarto quadratico medio sia $\sigma = 0.5$ cm.

$\mu = 5.1 \qquad \sigma = 0.1$

$$X = 5.0 \Rightarrow Z = \frac{5.0 - 5.1}{0.1} = -1$$

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

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



$\mu = 5.1 \qquad \sigma = 0.5$

$$X = 5.0 \Rightarrow Z = \frac{5.0 - 5.1}{0.5} = -0.2$$

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

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

Fare da soli gli esempi:

6, 7, 8, 9, 10, 11 !!

Esempio 12

La variabile aleatoria X ha la distribuzione normale con valor medio $\mu = 19$ e varianza $\sigma^2 = 49$; determinare il valore x_α tale che

a – $P(X > x_\alpha) = 0.20 = 20\%$;

b – $P(X < x_\alpha) = 0.90 = 90\%$.

a – Passando alla variabile normale standardizzata si ha

$$P(X > x_\alpha) = P\left(Z > z_\alpha = \frac{x_\alpha - 19}{7}\right) = 0.20 = 20\%$$

Sulla tavola 4 si trova

$$z_\alpha = \frac{x_\alpha - 19}{7} = 0.842$$

$$x_\alpha = 19 + 7 \cdot 0.842 = 24.894 \cong 24.9$$

b – La condizione richiesta significa che il 90% dell'area sottesa dalla curva normale è a destra di x_α, quindi il 10% è a sinistra. Passando alla variabile normale standardizzata si ha

$$P(X < x_\alpha) = P\left(Z < z_\alpha = \frac{x_\alpha - 19}{7}\right) = 0.90 = 90\%$$

$$P\left(Z > z_\alpha = \frac{x_\alpha - 19}{7}\right) = 0.10 = 10\%$$

Sulla tavola 4 si trova

$$z_\alpha = \frac{x_\alpha - 19}{7} = 1.282$$

$$x_\alpha = 19 + 7 \cdot 1.282 = 27.97$$

Esempio 13

Una macchina viene usata per tagliare assi di legno; la lunghezza media è di 2m, ma il 10% degli assi tagliati hanno una lunghezza inferiore a 1.95m.
Assumendo che le lunghezze degli assi tagliati abbiano una distribuzione normale, determinare la percentuale di assi più lunghi di 2.10m.

Sia X la variabile aleatoria che misura la lunghezza; X è distribuita normalmente con media $\mu = 2$; inoltre si sa che

$$P(X < 1.95) = 10\% .$$

Si deve calcolare $P(X > 2.10)$ e per far questo occorre prima determinare lo scarto quadratico medio σ. Passando alla variabile aleatoria standardizzata si ha

$$P(X < 1.95) = P\left(Z < \frac{1.95 - 2.00}{\sigma}\right) = 10\%$$

$$P\left(Z < \frac{1.95 - 2.00}{\sigma}\right) = P\left(Z > -\frac{1.95 - 2.00}{\sigma}\right) = 10\%$$

Sulla tavola 4 si trova che

$$-\frac{1.95 - 2.00}{\sigma} = 1.282$$

$$0.05 = 1.282 \cdot \sigma \qquad \Rightarrow \qquad \sigma = \frac{0.05}{1.282} = 0.039$$

Calcoliamo ora $P(X > 2.10)$. Passando alla variabile aleatoria standardizzata si ha

$$X = 2.10 \;\Rightarrow\; Z = \frac{2.10 - 2.00}{0.039} \cong 2.56$$

$$P(X > 2.10) = P(Z > 2.56) = 1 - P(Z < 2.56) = 1 - 0.9948 = 0.0052 .$$

In altre parole la percentuale di assi più lunghi di 2.10m è circa dello 0.5%.

Esempio 14

La variabile aleatoria X ha distribuzione normale con media μ e varianza σ^2. E' noto che il 10% dei valori di X è maggiore di 17.24 e che il 25% dei valori è minore di 14.37. Trovare il valor medio e la varianza.

Sono note le probabilità

$$P(X > 17.24) = 10\% \qquad P(X < 14.37) = 25\%.$$

Standardizzando la variabile e usando la tabella 4 si trova

$$P(X > 17.24) = P\left(Z > \frac{17.24 - \mu}{\sigma}\right) = 10\%$$

$$\frac{17.24 - \mu}{\sigma} = 1.282.$$

$$P(X < 14.37) = P\left(Z < \frac{14.37 - \mu}{\sigma}\right) = 25\%$$

$$P\left(Z > -\frac{14.37 - \mu}{\sigma}\right) = 25\%$$

$$-\frac{14.37 - \mu}{\sigma} = 0.674.$$

Risolvendo il sistema seguente si determinano i valori di μ e σ

$$\begin{cases} \dfrac{17.24 - \mu}{\sigma} = 1.282 \\ \dfrac{14.37 - \mu}{\sigma} = -0.674 \end{cases}$$

$$\begin{cases} 17.24 - \mu = 1.282 \cdot \sigma \\ 14.37 - \mu = -0.674 \cdot \sigma \end{cases}$$

Sottraendo le due equazioni si ha

$$2.87 = 1.956 \cdot \sigma \quad \Rightarrow \quad \sigma = 1.47$$

$$\mu = 17.24 - 1.282 \cdot 1.47 = 15.4$$

Fare da soli l'esempio 15 !!

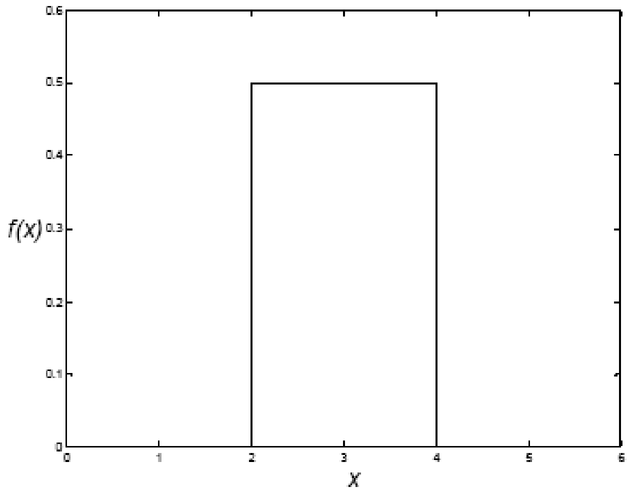
Distribuzione uniforme

Dati due numeri reali a e b , con $a < b$, si dice che la variabile aleatoria X ha **distribuzione uniforme** con parametri a e b, se la sua densità di probabilità è

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{altrimenti} \end{cases} \qquad (5.13)$$

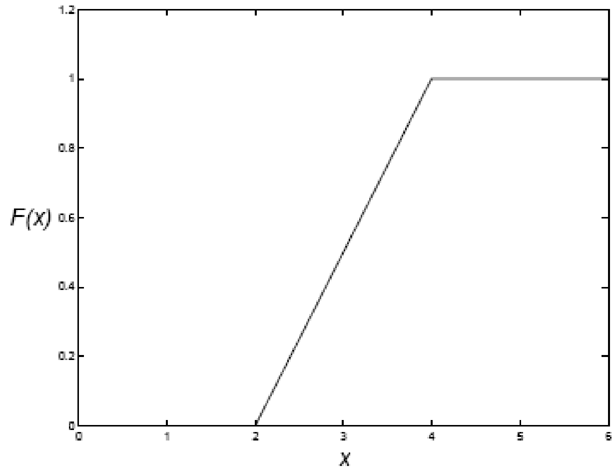
La **funzione di distribuzione uniforme** ha la seguente espressione

$$F(x) = P(X \le x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases} \qquad (5.14)$$



i grafici di $f(x)$ e $F(x)$

nel caso $a = 2$, $b = 4$



Parametri della Distribuzione Uniforme

Proprietà 2
Il **valor medio** e la **varianza** della distribuzione uniforme continua sono dati da

$$\mu = \frac{a+b}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12} \tag{5.15}$$

Infatti si ha

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

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

Esempio 24

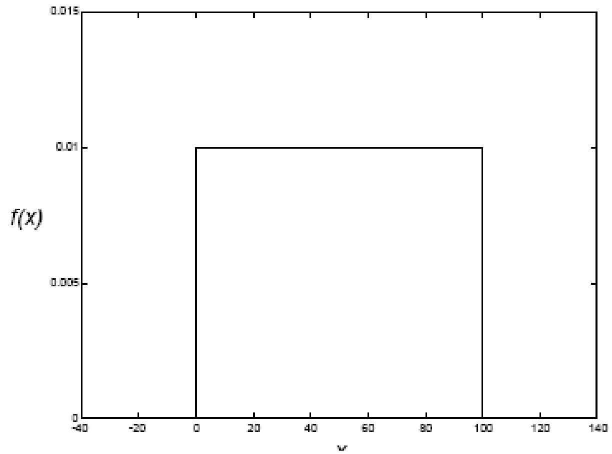
Una variabile aleatoria X è distribuita uniformemente nell'intervallo (0,100).

a – Calcolare la probabilità $P(20 < X < 60)$;

b – calcolare la media μ e la varianza σ^2 e trovare la probabilità $P(|X - \mu| < \sigma)$.

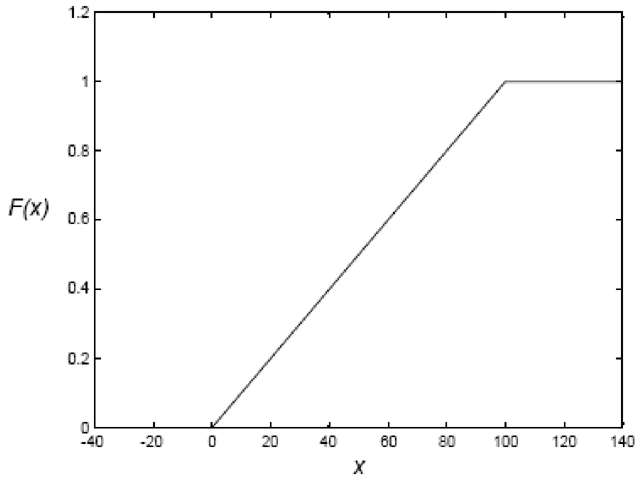
La variabile X ha la distribuzione uniforme (figura 24)

$$f(x) = \begin{cases} \dfrac{1}{100} & 0 < x < 100 \\ 0 & \text{altrimenti} \end{cases}$$



La funzione di distribuzione è (figura 25)

$$F(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x}{100} & 0 < x < 100 \\ 1 & x \ge 100 \end{cases}$$



$$P(20 < X < 60) = F(60) - F(20) = \frac{60}{100} - \frac{20}{100} = 0.4$$

$$\mu = 50$$

$$\sigma^2 = \frac{100^2}{12} \qquad \sigma = \frac{100}{\sqrt{12}} = \frac{50}{\sqrt{3}}$$

$$P\left(\left|X - 50\right| < \frac{50}{\sqrt{3}}\right) = P\left(50 - \frac{50}{\sqrt{3}} < X < 50 + \frac{50}{\sqrt{3}}\right) =$$

$$= F\left(50 - \frac{50}{\sqrt{3}}\right) - F\left(50 + \frac{50}{\sqrt{3}}\right) = \frac{50 + \frac{50}{\sqrt{3}}}{100} - \frac{50 - \frac{50}{\sqrt{3}}}{100} = \frac{1}{\sqrt{3}} \cong 0.577$$

Esempio 25

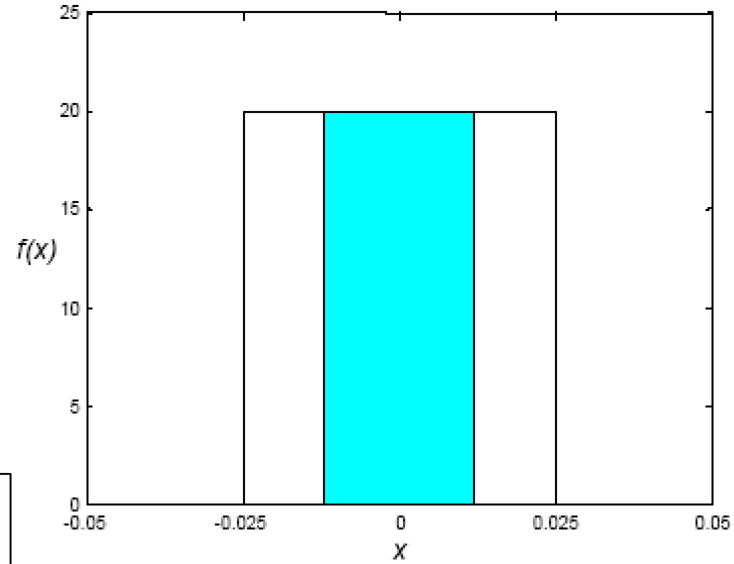
In certi esperimenti l'errore commesso nella determinazione della solubilità di una sostanza è una variabile aleatoria X avente distribuzione uniforme con $a = -0.025$ e $b = 0.025$.

Trovare la probabilità che l'errore

a – sia compreso fra 0.010 e 0.015;

b – sia compreso fra –0.012 e 0.012.

$$f(x) = \begin{cases} \dfrac{1}{0.05} = 20 & -0.025 < x < 0.025 \\ 0 & \text{altrimenti} \end{cases}$$



$$F(x) = \begin{cases} 0 & x \le 0.025 \\ \dfrac{x + 0.025}{0.05} & 0.025 < x < 0.025 \\ 1 & x \ge 0.025 \end{cases}$$

$$P(0.010 < X < 0.015) = F(0.015) - F(0.010) =$$

$$= \frac{0.015 + 0.025}{0.05} - \frac{0.010 + 0.025}{0.05} = 0.1$$

$$P(-0.12 < X < 0.012) = F(0.012) - F(-0.012) =$$

$$= \frac{0.012 + 0.025}{0.05} - \frac{-0.012 + 0.025}{0.05} = 0.48$$

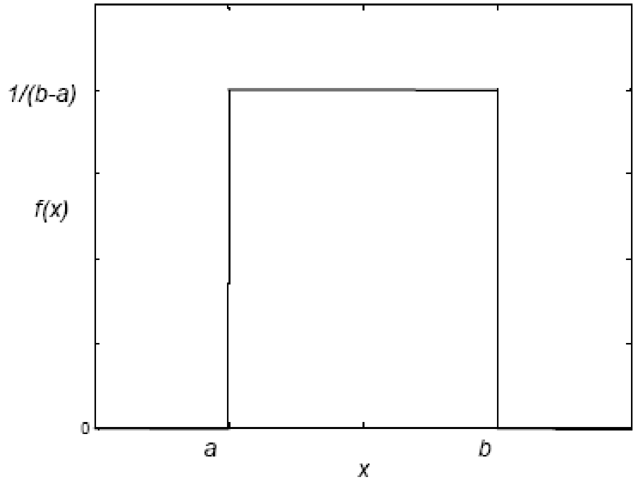
Disegnare

F(x) !!

Esempio 26

La variabile aleatoria X è distribuita uniformemente nell'intervallo (a,b); sapendo che $P(X<3)=\frac{1}{4}$ e $P(X<7)=\frac{3}{4}$, calcolare a e b.

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



Dai valori delle probabilità assegnate si deduce subito che deve essere $a<3$ e $b>7$.

La probabilità $P(X<3)$ è uguale all'area del rettangolo di base $3-a$ e altezza $\frac{1}{b-a}$; analogamente la probabilità $P(X<7)$ è uguale all'area del rettangolo di base $7-a$ e altezza $\frac{1}{b-a}$; si ottiene il sistema

$$\begin{cases}(3-a)\dfrac{1}{b-a}=\dfrac{1}{4}\\ (7-a)\dfrac{1}{b-a}=\dfrac{3}{4}\end{cases}$$

Risolvendo il sistema si ricava

$$a=1 \qquad b=9$$

$$f(x)=\begin{cases}\dfrac{1}{8} & 1<x<9\\ 0 & \text{altrimenti}\end{cases}$$

Distribuzione Esponenziale

- Modella il tempo trascorso prima che un evento si verifichi (*tempo di attesa*). Spesso viene usata per modellare la durata di un componente.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x > 0 \\ 0, \quad \text{otherwise} \end{cases}.$$

$$F(x) = \begin{cases} 0, x \le 0 \\ 1 - e^{-\lambda x}, x > 0 \end{cases}.$$

$$\mu_X = 1/\lambda.$$

$$\sigma_X^2 = 1/\lambda^2.$$

@ Minitab

Prima Proprietà

- Se gli eventi seguono un processo di *Poisson* di tasso λ e se T rappresenta il tempo di attesa tra un qualunque evento e il successivo allora

$$T \sim \mathrm{Exp}(\lambda).$$

$$P\{X = x\} = \frac{e^{-\lambda t}(\lambda t)^x}{x!}; \quad P\{X = 0\} = e^{-\lambda t} \text{non ci sono eventi in } (0, t] = P\{y > t\}$$

dove y indica la v.a. che indica l'intervallo di tempo al primo evento:

$$F_Y(t) = P\{y \le t\} = 1 - e^{-\lambda t} \quad \text{e ricordando che f (t) è la derivata di F(t) avremo}$$

$$f_Y(t) = \lambda e^{-\lambda t}. \quad \text{Se ne ricava dunque la FDP della distribuzione esponenziale.}$$

Seconda Proprietà - Memoryless

- La distribuzione esponenziale “*non ricorda*” quanto si è dovuto attendere in precedenza. In particolare se la durata di un componente segue una Exp allora la probabilità che un componente avente *s* anni di età duri per più di *t* anni è la stessa che un nuovo componente duri più di *t* anni.
- In altre parole questo componente non mostra gli effetti dell’età e avremo:

$$P(T > t + s \mid T > s) = P(T > t).$$

Esempio pratico
(Lack of Memory Property)

Hp1) Circuito integrato ≈ Exp(vita media pari a 2 anni)

Ts1) P(T > 3 anni)=1- P(T <3)=1-[1-exp(-0.5*3)]=exp(-1.5)=0.223

Hp2) Assumiamo che lo stesso circuito abbia un'età di 4 anni e sia ancora funzionante, avremo:

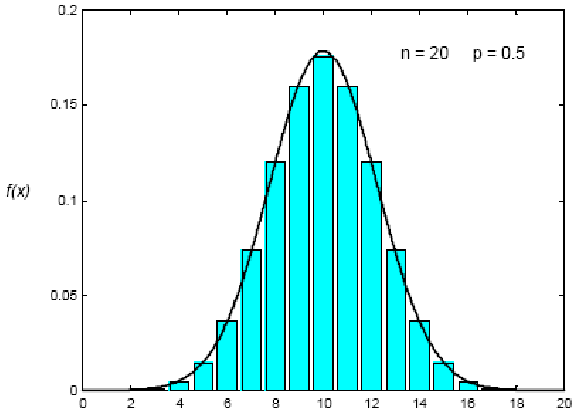
Ts2) P(che il circuito funzioni ancora per 3 anni) =

$$P(T > 7 \| T > 4) = \frac{P(T > 7, T > 4)}{P(T > 4)} = \frac{P(T > 7)}{P(T > 4)} = \frac{e^{-0.5(7)}}{e^{-0.5(4)}} = 0.223 \quad (!!)$$

Relazione tra la distribuzione binomiale e la distribuzione normale

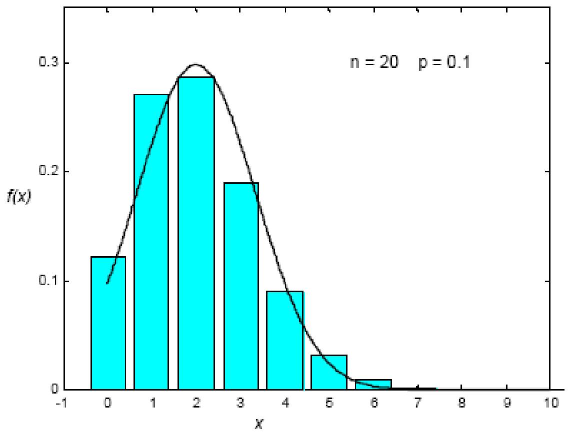
Si può dimostrare che, quando n è grande e p è vicino a 0.5, la distribuzione binomiale della variabile aleatoria X può essere approssimata da una distribuzione normale con variabile aleatoria standardizzata

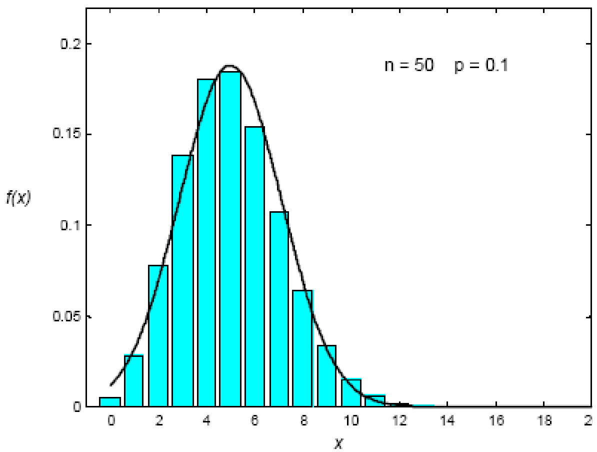
$$Z = \frac{X - np}{\sqrt{np(1-p)}} . \qquad \mu = np \qquad \sigma^2 = np(1-p) \qquad (5.11)$$

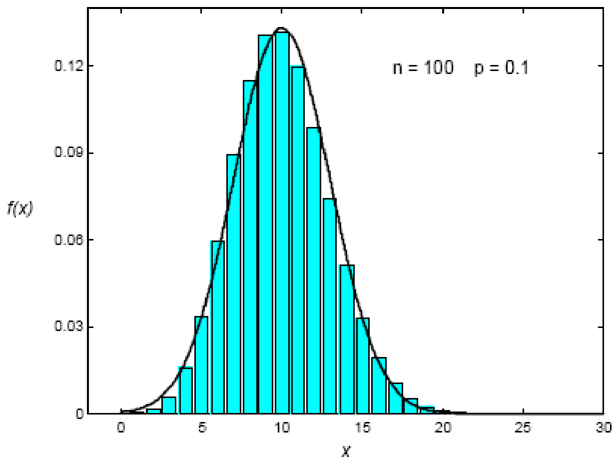


Come **regola pratica** si usa la distribuzione normale per approssimare la binomiale se si verificano entrambe le condizioni $np \geq 5$ e $n(1-p) \geq 5$.

La regola suggerita è soddisfatta se n è abbastanza grande e l'approssimazione è tanto più precisa quanto più p è prossima a 0.5.







Correzione di Continuità

Per poter usare correttamente la distribuzione normale, che è continua, per approssimare la distribuzione di una variabile aleatoria discreta occorre effettuare la **correzione di continuità**[1]: questo avviene rappresentando ogni valore intero x assunto dalla variabile aleatoria discreta con l'intervallo di estremi $x-\frac{1}{2}$ e $x+\frac{1}{2}$. Quindi, se X è una variabile aleatoria con distribuzione binomiale di parametri n e p, la probabilità $P(a \le X \le b)$ che X assuma valori compresi fra a e b, viene approssimata con il valore della probabilità che la variabile aleatoria normale con media $\mu = np$ e varianza $\sigma^2 = np(1-p)$ assuma valori compresi tra $a-\frac{1}{2}$ e $b+\frac{1}{2}$, ossia con il valore dell'area sottesa dalla curva normale tra $a-\frac{1}{2}$ e $b+\frac{1}{2}$.

Nel caso particolare in cui $a = b$, la probabilità binomiale $P(X = a)$ viene approssimata con il valore della probabilità $P\left(a-\frac{1}{2} \le X \le a+\frac{1}{2}\right)$ calcolata con la distribuzione normale.

Esempio 16

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

Per calcolare la probabilità $P(X=40)$ usando la distribuzione normale, occorre effettuare la correzione di continuità e calcolare la probabilità

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

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

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

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

Usando le tavole della distribuzione normale si trova

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

Questa approssimazione è molto buona, perché il valore di n è sufficientemente grande e il valore di p è 0.5.

Esempio 17

Trovare la probabilità che, in 10 lanci di una moneta, testa si presenti un numero di volte compreso fra 3 e 6, usando

a – la distribuzione binomiale;

b – la distribuzione normale per approssimare la distribuzione binomiale.

a – Sia X la variabile aleatoria binomiale. $n = 10$ $p = \frac{1}{2}$

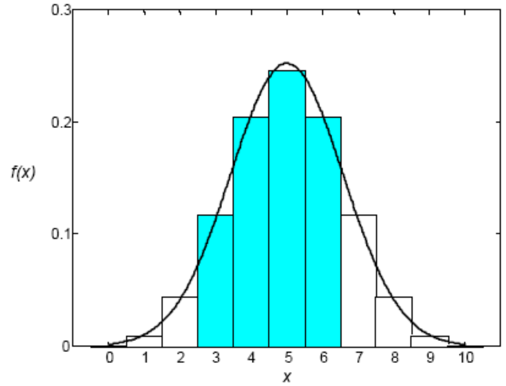
$$P(3 \le X \le 6) = P(X \le 6) - P(X \le 2) = 0.8281 - 0.0547 = 0.7734$$

b – Se si considera la variabile X come continua, si deve fare la correzione di continuità e calcolare la probabilità $P(2.5 \le X \le 6.5)$; standardizzando la variabile con la (5.11) si ha

$$\mu = np = 10 \cdot \frac{1}{2} = 5 \qquad \sigma^2 = np(1-p) = 10 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2.5$$

$$X = 2.5 \quad \Rightarrow \quad Z = \frac{2.5-5}{\sqrt{2.5}} = -1.58$$

$$X = 6.5 \quad \Rightarrow \quad Z = \frac{6.5-5}{\sqrt{2.5}} = 0.95$$



Usando le tavole della distribuzione normale si trova

$$P(-1.58 < Z < 0.95) = P(Z < 0.95) - P(Z < -1.58) =$$
$$= 0.8289 - [1 - P(Z < 1.58)] = 0.8289 - 1 + 0.9429 = 0.7718$$
$$P(3 \le X \le 6) \cong 0.7718$$

Esempio 18

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.

Usare l'approssimazione della distribuzione binomiale con la normale.

a – In questo caso si cerca la probabilità che il numero di teste sia compreso fra 240 e 260, ossia, con la correzione di continuità, la probabilità

$$P(239.5 < X < 260.5)$$

Effettuando il passaggio alla variabile standardizzata si ha

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

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

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

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

Usando le tavole della distribuzione normale si trova

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

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

b – In questo caso si cerca la probabilità che il numero di teste sia compreso fra 220 e 280, ossia, con la correzione di continuità, la probabilità

$$P(219.5 < X < 280.5) .$$

Effettuando il passaggio alla variabile standardizzata si ha

$$\mu = np = 250 \qquad \sigma = 11.18$$

$$X = 219.5 \quad \Rightarrow \quad Z = \frac{219.5 - 250}{11.18} = -2.73$$

$$X = 280.5 \quad \Rightarrow \quad Z = \frac{280.5 - 250}{11.18} = 2.73$$

Usando le tavole della distribuzione normale si trova

$$P(-2.73 < Z < 2.73) = 2 \cdot 0.9968 - 1 = 0.9936 \cong 99.4\%$$

$$P(220 \le X \le 280) \cong 0.9936$$

Esempio 19

Un dado viene lanciato 120 volte. Calcolare la probabilità che il numero 3 si presenti al più 15 volte.

La faccia con il numero 3 ha la probabilità $p = \frac{1}{6}$ di presentarsi. La probabilità che il numero 3 si presenti un numero di volte compreso fra 0 e 15, con la distribuzione binomiale è

$$n = 120 \qquad p = \frac{1}{6}$$

$$P(0 \le X \le 15) = P(X = 0) + P(X = 1) + \ldots + P(X = 15)$$

Il lavoro necessario per il calcolo dei 16 addendi presenti nella somma è eccessivo ed è preferibile usare l'approssimazione con la normale; si ottiene una buona approssimazione, dato che

$$np = 120 \cdot \frac{1}{6} = 20 \qquad \text{e} \qquad n(1-p) = 120 \cdot \frac{5}{6} = 100 \, .$$

Effettuando la correzione di continuità e standardizzando la variabile si trova

$$\mu = np = 120 \cdot \frac{1}{6} = 20 \qquad \sigma = \sqrt{np(1-p)} = \sqrt{120 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 4.08$$

$$X = -0.5 \quad \Rightarrow \quad Z = \frac{-0.5 - 20}{4.08} = -5.02$$

$$X = 15.5 \quad \Rightarrow \quad Z = \frac{15.5 - 20}{4.08} = -1.10$$

$$P(-5.02 < Z < -1.10) = P(1.10 < Z < 5.02) =$$

$$= 0.9999997 - 0.8643 = 0.1357$$

$$P(0 \le X \le 15) \cong 0.1357$$

Esempio 20

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 – Si deve calcolare la probabilità $P(X \leq 15)$.

Usando l'approssimazione con la normale ed effettuando la correzione di continuità, si ha

$$np = 100 \cdot 0.2 = 20 \qquad n(1-p) = 100 \cdot 0.8 = 80$$

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

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

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

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

b – Si deve calcolare $P(X = 15)$.

Usando l'approssimazione con la normale ed effettuando la correzione di continuità, si ha

$$X = 14.5 \quad \Rightarrow \quad Z = \frac{14.5 - 20}{4} = -1.38$$

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

$$P(-1.38 < Z < -1.13) = P(Z < 1.38) - P(Z < 1.13) =$$
$$= 0.9162 - 0.8708 = 0.0454$$
$$P(X = 15) \cong 0.0454$$

Per confronto si può effettuare con un software statistico il calcolo delle probabilità con la distribuzione binomiale e si trovano i valori

$$P(X \le 15) = 0.1285$$
$$P(X = 15) = 0.0481$$

Esempio 21

I voti di un questionario vanno da 1 a 10, a seconda del numero di risposte a 10 domande. Il voto medio è $\mu = 6.7$ e lo scarto quadratico medio è $\sigma = 1.2$. Supponendo che i voti siano distribuiti normalmente determinare

a – la percentuale di studenti che ha ottenuto il voto 6;

b – il voto minimo del miglior 10% del gruppo di studenti;

c – il voto massimo del peggior 10% del gruppo di studenti.

a – Effettuando la correzione di continuità, calcoliamo con la distribuzione normale la probabilità $P(5.5 < X < 6.5)$.

Standardizzando la variabile con la (5.5) si ha

$$\mu = 6.7 \qquad \sigma = 1.2$$

$$X = 5.5 \quad \Rightarrow \quad Z = \frac{5.5 - 6.7}{1.2} = -1.0$$

$$X = 6.5 \quad \Rightarrow \quad Z = \frac{6.5 - 6.7}{1.2} \cong -0.17$$

Usando le tavole della distribuzione normale si trova

$$P(5.5 < X < 6.5) = P(-1.0 < Z < -0.17) = P(0.17 < Z < 1.0) =$$
$$= F(1.0) - F(0.17) = 0.8413 - 0.5675 = 0.2738 \cong 27.4\%$$

b – Sia x_1 il voto minimo richiesto e z_1 il voto corrispondente in unità standardizzate.

Dalla figura 21 (pagina seguente), si vede che l'area a destra di z_1 è il 10% dell'area totale. Dalle tavole dei quantili per la distribuzione normale si ricava

$$z_1 = 1.282$$

Dalla relazione (5.5) si ottiene

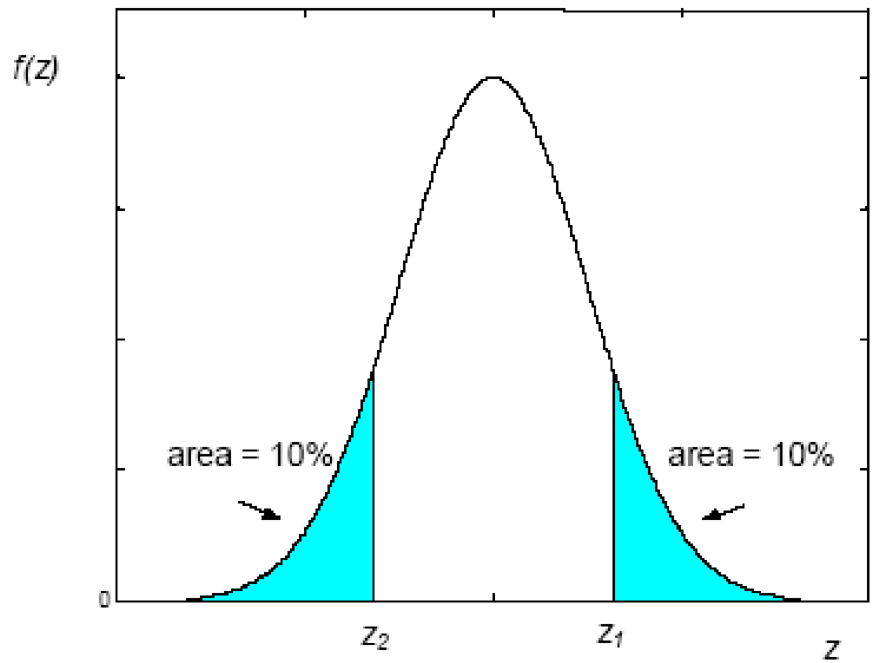
$$z_1 = \frac{x_1 - 6.7}{1.2} = 1.282 \qquad \Rightarrow \qquad x_1 = 1.2 \cdot 1.282 + 6.7 \cong 8.24$$

Il voto minimo del miglior 10% degli studenti è 8 (l'intero più prossimo a x_1)

c – Il punto z_2 è il simmetrico di z_1 rispetto all'origine, ossia $z_1 = -1.282$; quindi

$$z_2 = \frac{x_2 - 6.7}{1.2} = -1.282 \qquad \Rightarrow \qquad x_2 = -1.2 \cdot 1.282 + 6.7 \cong 5.16$$

Il voto massimo del peggior 10% degli studenti è perciò 5 (l'intero più prossimo a x_2).



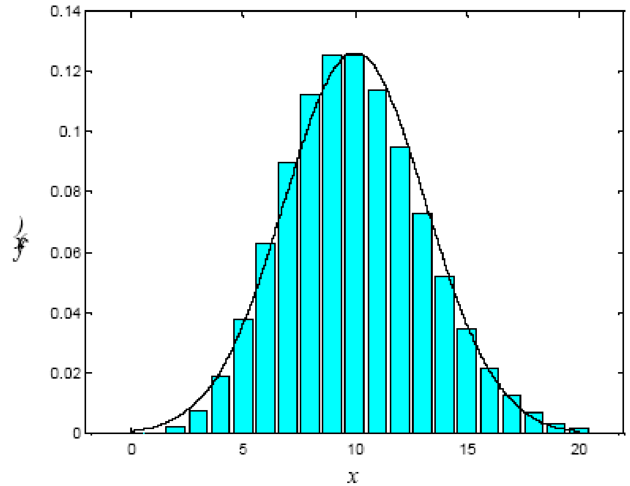
Relazione tra la distribuzione normale e la distribuzione di Poisson

Si dimostra che se X è una variabile aleatoria avente la distribuzione di Poisson, con media $\mu = \lambda$ e varianza $\sigma^2 = \lambda$, allora al crescere λ la distribuzione della variabile X può essere approssimata da una distribuzione normale con variabile aleatoria standardizzata

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \tag{5.12}$$

Come per la binomiale, anche per la distribuzione di Poisson, trattandosi di una distribuzione discreta, occorre fare la correzione di continuità.
L'approssimazione è sufficientemente buona per $\lambda \geq 10$.

Nella figura 22, per illustrare l'approssimazione fra la distribuzione di Poisson e la normale, sono riportati il grafico della distribuzione di Poisson per $\lambda = 10$ e il grafico della distribuzione normale avente valor medio $\mu = \lambda = 10$ e scarto quadratico medio $\sigma = \sqrt{\lambda} = \sqrt{10}$.



Esempio 22

La variabile aleatoria X ha distribuzione di Poisson con media $\lambda = 50$. Calcolare la probabilità $P(X < 40)$ usando l'approssimazione con la normale.

Si deve calcolare

$$P(X < 40) = P(X \leq 39)$$

Usando la distribuzione normale con la correzione di continuità si trova

$$X = 39.5 \quad \Rightarrow \quad Z = \frac{39.5 - 50}{\sqrt{50}} \cong -1.48$$

$$P(Z < -1.48) = 1 - P(Z > 1.48) = 1 - 0.9306 = 0.0694$$

$$P(X < 40) \cong 0.0694$$

Effettuando con un software statistico il calcolo delle probabilità con la distribuzione di Poisson, si trova il valore

$$P(X \leq 39) = 0.0646 \,.$$

Esempio 23

Il numero di incidenti d'auto che si verificano in un giorno ad un incrocio è una variabile aleatoria con distribuzione di Poisson e media 1.4; calcolare la probabilità che accadano più di 50 incidenti in un periodo di 4 settimane.

Il numero di incidenti che si verificano in 28 giorni è una variabile X con media $\lambda = 1.4 \cdot 28 = 39.2$. Si ha

$$P(X > 50) = 1 - P(X \leq 50)$$

$$Z = \frac{50.5 - 39.2}{\sqrt{39.2}} \cong 1.80$$

$$P(X > 50) \cong 1 - P(Z < 1.80) = 1 - 0.9645 = 0.0355$$