

La distribuzione Gaussiana

23 maggio 2017

Consideriamo la funzione più importante e frequente che si incontra in statistica: la distribuzione Normale, chiamata molto più comunemente distribuzione Gaussiana

1 Distribuzione Normale come limite della Binomiale

Data una distribuzione binomiale $Bin(k | n, p)$, se $n \rightarrow \infty$ e $p \rightarrow 0$ allora la distribuzione diventa una distribuzione Poissoniana $Poisson(k | \mu = n \cdot p)$.

Se invece $n \rightarrow \infty$ e p ha valore finito, la distribuzione che si ottiene al limite è definita dal teorema di De Moivre-Laplace (che è un caso particolare del teorema centrale).

Teorema 1.1 (De Moivre - Laplace) Sia $Bin(k | n, p) = \binom{n}{k} p^k q^{n-k}$ una distribuzione binomiale con $|\frac{k-np}{2npq}| \leq A = \text{cost.}$ Allora:

$$Bin(k | n, p) \simeq \frac{1}{\sqrt{2\pi \cdot npq}} \cdot e^{-\frac{(k-np)^2}{2npq}} (1 + O(\frac{1}{\sqrt{n}})). \quad (1)$$

Accenno di dimostrazione Il risultato si ottiene come caso particolare del Teorema Limite Centrale (dimostrato più avanti)

Il teorema (1.1) implica che se $\mu = np$ e $\sigma = \sqrt{npq}$, usando la variabile continua x al posto di quella discreta k

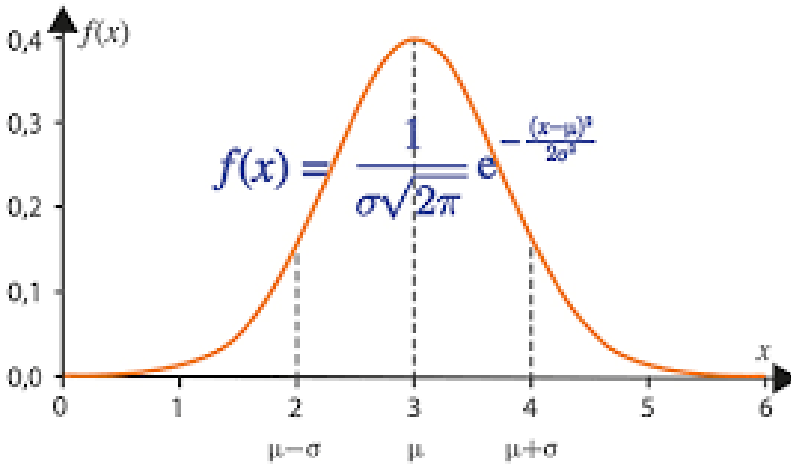
$$P_X(x_1 \leq X \leq x_2) = \lim_{n \rightarrow \infty} \sum_{x_1}^{x_2} \frac{1}{\sqrt{2\pi \cdot npq}} \cdot e^{-\frac{(x-np)^2}{2npq}} = \int_{x_1}^{x_2} \underbrace{\frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}}_{f(x)} dx.$$

La $f(x)$ integrata è allora una funzione di densità di probabilità continua che definiamo come distribuzione normale o Gaussiana

$$\mathcal{N}(x | \mu, \sigma) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (2)$$

dove i parametri μ, σ^2 sono rispettivamente la media e la varianza della distribuzione. La PDF della Normale è riportata in figura 1

Si noti che essendo la Gaussiana funzione simmetrica e unimodale (e che raggiunge il suo valore massimo al centro), *mediana e moda coincidono con la media μ*



Esempio di curva gaussiana

Figura 1: Distribuzione La forma a campana tipica della densità Gaussiana $\mathcal{N}(x | \mu, \sigma)$ (nell'esempio $\mu = 3$)

2 Normalizzazione e standardizzazione delle variabili

Per dimostrare che la distribuzione $\mathcal{N}(x | \mu, \sigma)$ gode della proprietà di essere normalizzata, $\int_{-\infty}^{+\infty} \mathcal{N}(x | \mu, \sigma) dx = 1$, standardizzeremo le variabile aleatoria X .

La *standardizzazione* è un procedimento che riconduce una variabile aleatoria X distribuita secondo una media μ e varianza σ^2 , ad una variabile aleatoria

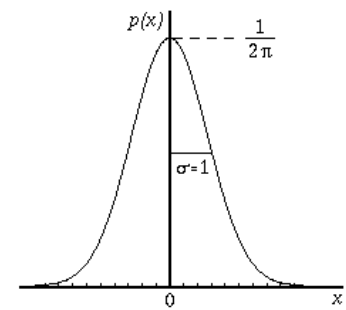
$$Z = \frac{X - \mu}{\sigma}, \quad (3)$$

con distribuzione "standard", ossia di media zero e varianza unitaria,:

$$E[Z] = 0 \quad \text{var}(Z) = E[Z^2] = 1.$$

Per la dimostrazione di normalizzazione è sufficiente standardizzare X come in (3) e usare l'integrale di Gauss nella forma posta in Equazione (13):

Distribuzione normale "standardizzata"

Figura 2: Distribuzione Normale standardizzata dove $\mu = 0$ e $\sigma = 1$

$$\begin{aligned}
& \int_{-\infty}^{+\infty} N(x \mid \mu, \sigma) dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}z^2} \sigma dz \\
&= \int_{-\infty}^{+\infty} \underbrace{\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}}_{\text{Forma Normale Standard}} dz \\
&= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz \\
&= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} \\
&= 1
\end{aligned}$$

La funzione di densità di probabilità che abbiamo ottenuto dopo aver standardizzato le variabili, cioè

$$f(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} \quad (4)$$

è detta *distribuzione normale standard o ridotta*, ed è indicata come $\mathcal{N}(z \mid \mu = 0, \sigma = 1)$ o semplicemente $\mathcal{N}(0, 1)$.

Calcoliamo ora la speranza matematica usando la standardizzazione (3) da cui $x = \sigma z + \mu \Rightarrow \sigma dz = dx$ e procedendo per sostituzione di variabile:

$$\begin{aligned}
E[X] &= \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma^2} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{+\infty} (\sigma z + \mu) \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}z^2} \sigma dz \\
&= \frac{\sigma}{\sqrt{2\pi}} \cdot \underbrace{\int_{-\infty}^{+\infty} z \cdot e^{-\frac{1}{2}z^2} dz}_0 + \frac{\mu}{\sqrt{2\pi}} \cdot \underbrace{\int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz}_{\sqrt{2\pi}} \\
&= \frac{\mu\sqrt{2\pi}}{\sqrt{2\pi}} \\
&= \mu
\end{aligned}$$

Procedendo analogamente per calcolare la varianza si ottiene $\text{var}(X) = \sigma^2$, e ovviamente $\sqrt{\text{var}(X)} = \sigma$ è la deviazione standard.

Si noti che la PDF della normale è simmetrica e unimodale con il massimo centrato su μ (su $x = 0$ nel caso ridotto), dunque moda e mediana coincidono con la media μ .

3 Proprietà della CDF (Cumulative Density Function) della distribuzione normale

Per definizione, la CDF e la funzione di sopravvivenza $S(z)$ si possono formalmente scrivere:

$$F_Z(z) = P_Z(Z \leq z) = \int_{-\infty}^z f(t)dt = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt, \quad (5)$$

$$S_Z(z) = P_Z(Z > z) = \int_z^{+\infty} f(t)dt = \int_z^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt = 1 - F_Z(z). \quad (6)$$

Come abbiamo già notato la $f(z)$ è una funzione pari:

$$f(z) = f(-z). \quad (7)$$

Usando la parità è possibile ricavare le seguenti proprietà di cui gode la PDF normale:

$$F_Z(-z) = P_Z(Z \leq -z) = \int_{-\infty}^{-z} f(t)dt = - \int_{\infty}^z f(-t)dt = \int_z^{\infty} f(t)dt = 1 - F_Z(z), \quad (8)$$

dove si è utilizzata la sostituzione $t \rightarrow -t \Rightarrow dt \rightarrow -dt$, e la parità di f .

Vale dunque la seguente:

$$F_Z(-z) = 1 - F_Z(z). \quad (9)$$

Inoltre possiamo ricavare le seguenti:

$$W_Z(z) = P_Z(-z \leq Z \leq +z) = F_Z(z) - F_Z(-z) = F_Z(z) - 1 + F_Z(z) = 2F_Z(z) - 1. \quad (10)$$

$$2R_Z(z) = P_Z(|Z| \geq z) = P_Z(Z \leq -z) + P_Z(Z \geq z) = 1 - F_Z(z) + 1 - F_Z(z) = 2(1 - F_Z(z)). \quad (11)$$

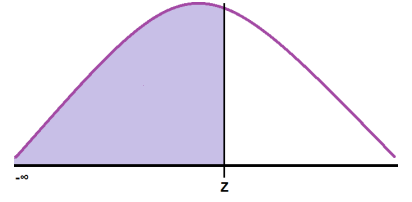


Figura 3: $P_Z(Z \leq z)$

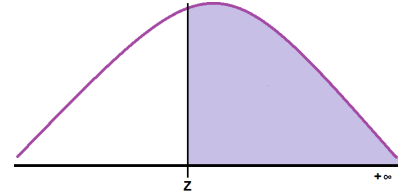


Figura 4: $P_Z(Z \geq z)$

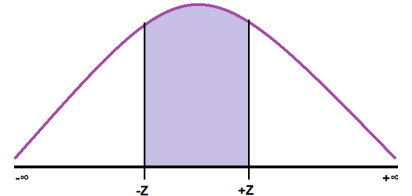


Figura 5: $W_Z(z) = P_Z(-z \leq Z \leq z)$

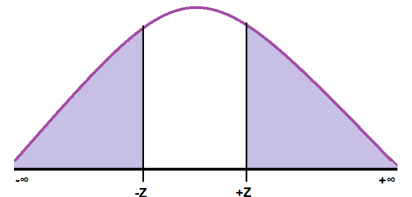


Figura 6: $2R_Z(z) = P_Z(|Z| \geq z)$

Queste calcolano la probabilità corrispondente all'area ombreggiata della densità $f(z)$ come rappresentato nelle figure accanto.

$S_Z(z), R_Z(z), F_Z(-z), W_Z(z), 2R_Z(z)$ definiscono nella pratica le probabilità di interesse per qualsiasi problema che riguardi la distribuzione normale.

In buona sostanza, le relazioni ricavate precedentemente sono tutte calcolabili facilmente a partire da $F_Z(z)$.

Nella forma più generale la CDF $F_Z(z)$ della normale ridotta viene spesso denotata $F_Z(z) = \Phi(z)$ e in generale:

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right). \quad (12)$$

Infatti, invertendo la trasformazione standard, si ha che $X = \sigma Z + \mu$, dunque:

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(\sigma Z + \mu \leq x) = \\ &= P(Z \leq \frac{x - \mu}{\sigma}) = \Phi\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

Abbiamo un ultimo rilevante problema: la $\Phi(z)$ come scritta in equazione (5) non è calcolabile analiticamente!

Soluzioni possibili:

- se stiamo lavorando in simulazione software, usiamo qualche funzione basata su approssimazione numerica dell'integrale
- se stiamo lavorando "carta e penna", utilizziamo le tabelle della forma normale standard.

Un esempio semplificato di tabella della normale standard è mostrato in Figura 7

Se usiamo come "entry point" nella tabella i valori $z = 1, 2, 3$ (x nella tabella in Figura 7) e leggiamo i corrispondenti valori di $W_Z(z) = P_Z(-z \leq Z \leq +z)$, tenendo presente che per la normale standard $\sigma = 1$ e dunque $z = z\sigma$, ne risulta che :

1. i valori di z compresi tra $-1\sigma \leq z \leq 1\sigma$ corrispondono a $\approx 68,3\%$ della distribuzione;
2. i valori di z compresi tra $-2\sigma \leq z \leq 2\sigma$ corrispondono a $\approx 95,6\%$ della distribuzione;
3. i valori di z compresi tra $-3\sigma \leq z \leq 3\sigma$ corrispondono a $\approx 99,7\%$ della distribuzione.

Questa proprietà della distribuzione normale viene detta **Legge 3σ** , che riprenderemo in seguito.

Si noti che sarebbe bastato avere la colonna relativa a $F_Z(z)$, per ottenere $W_Z(z) = 2F_Z(z) - 1$. Anzi, tipicamente le tabelle utilizzate di solito riportano solo il valore compreso tra $Z = 0$ e $Z = z$ essendo per simmetria della distribuzione $P_Z(-\infty \leq Z \leq 0) = 0.5$. Nelle tabelle complete sulle righe (in prima colonna) si accede al valore di $F_Z(z)$ al primo decimale ($F_Z(z = 0.1), F_Z(z = 0.2) \dots$ ecc.); sulla stessa riga, nelle colonne successive si considerano con maggiore precisione i valori di $z = 0.01, 0.02, 0.03, \dots$ da comporre con il primo decimale per poter calcolare valori più precisi, quali $F_Z(z = 1.63)$, ecc.

Un esempio è riportato in Figura 8, dove, ad esempio $F_Z(z = 1.63)$ si ricava

1. accedendo per riga in $z = 1.6$
2. posizionandosi sulla colonna indicizzata da $z = 0.03$
3. leggendo il valore $P_Z(0 \leq Z \leq z = 1.63) = 0.4484$
4. calcolando $F_Z(z = 1.63) = P_Z(-\infty \leq Z \leq 0) + P_Z(0 \leq Z \leq z = 1.63) = 0.5 + 0.4484 = 0.9484$

4 Legge 3σ

La Legge 3σ in sintesi ci dice che la densità di probabilità gaussiana è concentrata soprattutto in un intervallo di uno o due σ attorno al valore $X = \mu$ e diventa pressoché nulla per $|X - \mu| > 3\sigma$.

Ricaviamo esplicitamente la Legge 3σ calcolando la probabilità che una VA $X \sim \mathcal{N}(\mu, \sigma)$ abbia uno scarto dalla media:

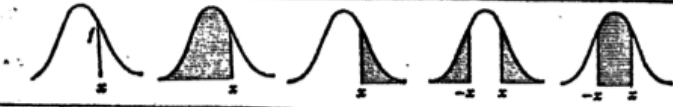
$$|X - \mu| \leq k\sigma$$

ovvero

$$P(|X - \mu| \leq k\sigma) = P(-k\sigma \leq X - \mu \leq k\sigma) = P(\mu - k\sigma \leq X \leq \mu + k\sigma)$$

come rappresentato in Figura 10

TABLE 4-4. UNIT NORMAL DISTRIBUTION



z	$f(z)$	$F(z)$	$R(z)$	$2R(z)$	$W(z)$
.0	.3989	.5000	.5000	1.0000	0
.1	.3970	.5398	.4602	.9203	.0797
.2	.3910	.5793	.4207	.8415	.1585
.3	.3814	.6179	.3821	.7642	.2358
.4	.3683	.6554	.3446	.6892	.3108
.5	.3521	.6915	.3085	.6171	.3829
.6	.3332	.7257	.2743	.5485	.4515
.7	.3123	.7580	.2420	.4839	.5161
.8	.2897	.7881	.2119	.4237	.5763
.9	.2661	.8159	.1841	.3681	.6319
1.0	.2420	.8413	.1587	.3173	.6827
1.1	.2179	.8643	.1357	.2713	.7287
1.2	.1942	.8849	.1151	.2301	.7699
1.3	.1714	.9032	.0968	.1936	.8064
1.4	.1497	.9192	.0808	.1615	.8385
1.5	.1295	.9332	.0668	.1336	.8664
1.6	.1109	.9452	.0548	.1096	.8904
1.7	.0940	.9554	.0446	.0891	.9109
1.8	.0790	.9641	.0359	.0719	.9281
1.9	.0656	.9713	.0287	.0574	.9426
2.0	.0540	.9772	.0228	.0455	.9545
2.1	.0440	.9821	.0179	.0357	.9643
2.2	.0355	.9861	.0139	.0278	.9722
2.3	.0283	.9893	.0107	.0214	.9786
2.4	.0224	.9918	.0082	.0164	.9836
2.5	.0175	.9938	.0062	.0124	.9876
2.6	.0136	.9953	.0047	.0093	.9907
2.7	.0104	.9965	.0035	.0069	.9931
2.8	.0079	.9974	.0026	.0051	.9949
2.9	.0060	.9981	.0019	.0037	.9963
3.0	.0044	.9987	.0013	.0027	.9973
Fractiles					
1.2816	.1755	.9000	.1000	.2000	.8000
1.6449	.1031	.9500	.0500	.1000	.9000
1.9600	.0584	.9750	.0250	.0500	.9500
2.0537	.0484	.9800	.0200	.0400	.9600
2.3263	.0267	.9900	.0100	.0200	.9800
2.5758	.0145	.9950	.0050	.0100	.9900

Figura 7: Tabella semplificata della normale ridotta

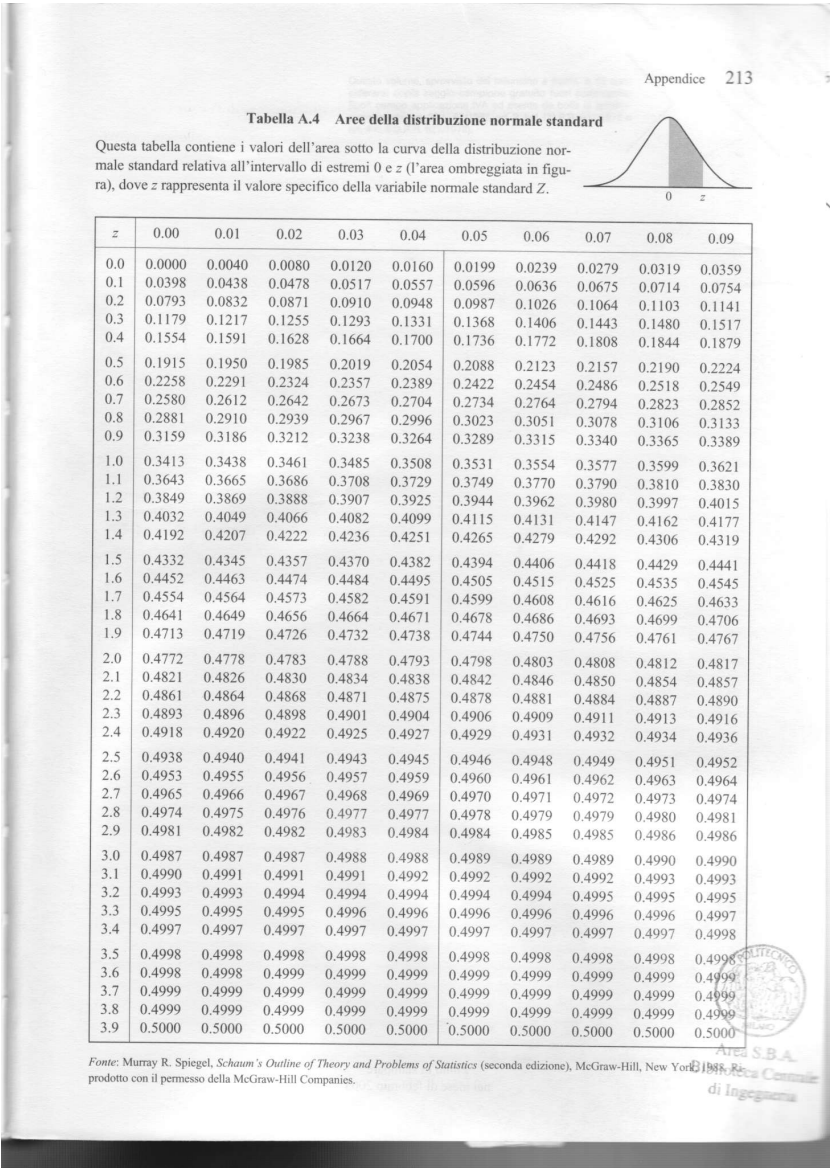


Figura 8: Tabella completa della normale ridotta

Standardizziamo X :

$$z = \frac{x - \mu}{\sigma}$$

da cui

$$dz = \frac{1}{\sigma} dx \rightarrow dx = \sigma dz$$

I limiti di integrazione sono pertanto:

$$x_2 = \mu + k\sigma \rightarrow z_2 = \frac{\mu + k\sigma - \mu}{\sigma} = k$$

$$x_1 = \mu - k\sigma \rightarrow z_1 = \frac{\mu - k\sigma - \mu}{\sigma} = -k$$

e dunque:

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) = \int_{\mu - k\sigma}^{\mu + k\sigma} N(x|\mu, \sigma) dx = \int_{-k}^{+k} f(z) dz = W(k)$$

I tre casi di interesse per la Legge 3 σ sono quelli relativi a $k = 1, 2, 3$. Per completezza ricaviamo il valore complementare $P(|X - \mu| \geq k\sigma)$ e lo confrontiamo con il limite generale ottenuto dal Teorema di Tchebychev, $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$:

$$-k = 1 \rightarrow W(1) = \int_{-1}^{+1} f(z) dz = 0.6827 \rightarrow P(|X - \mu| \geq 1\sigma) = 0.3173 \leq \frac{1}{1^2} = 1$$

$$-k = 2 \rightarrow W(2) = \int_{-2}^{+2} f(z) dz = 0.9545 \rightarrow P(|X - \mu| \geq 2\sigma) = 0.0455 \leq \frac{1}{2^2} = 0.25$$

$$-k = 3 \rightarrow W(3) = \int_{-3}^{+3} f(z) dz = 0.9973 \rightarrow P(|X - \mu| \geq 3\sigma) = 0.0027 \leq \frac{1}{3^2} = 0.11$$

La Legge 3 - σ é sintetizzata graficamente nella Figura 11



Figura 9: La regione ombreggiata indica la probabilità dello scarto $|X - \mu| \leq k\sigma$

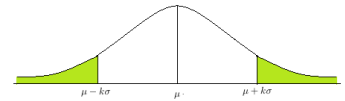


Figura 10: La regione ombreggiata indica la $P(|X - \mu| \geq k\sigma)$

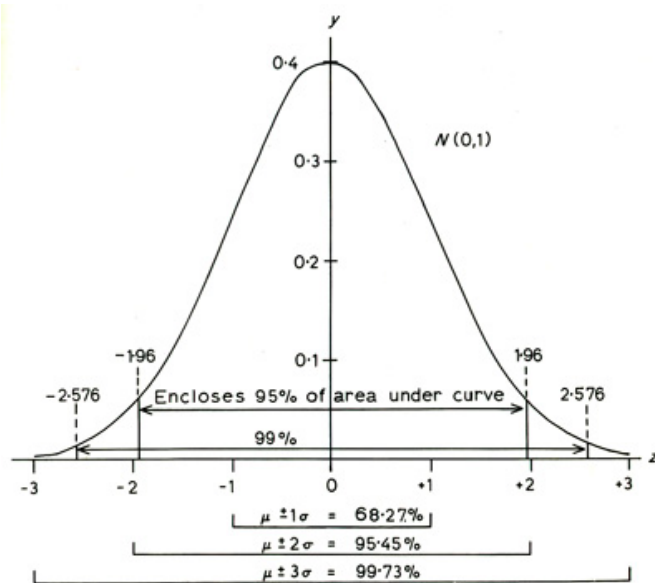


Figura 11: Distribuzione dei dati secondo la legge 3 - σ

5 Esempi di utilizzo della Normale e delle tabelle della CDF

Esempio 5.1 Il contenuto reale di birra che un apposito macchinario mette in fusti da 5 litri può essere considerato come una variabile aleatoria avente una distribuzione normale con una deviazione standard pari a 0.05 litri.

Se solo il 2% dei fusti contiene meno di 5 litri, quale dovrebbe essere il contenuto medio dei fusti?

Per determinare μ tale che $\Phi(z) = F_X\left(\frac{5-\mu}{\sigma}\right) = 0.02$, si cerca nella tabella della distribuzione normale standard il valore più vicino a 0.02. Si trova la migliore approssimazione per 0.0202 che corrisponde a $z = -2.05$.

Qualora la tabella a disposizione consideri solo i valori per $z \geq 0$, si faccia uso della proprietà della cumulativa standardizzata $\Phi(-z) = 1 - \Phi(z)$; in tal caso la migliore approssimazione si ha per $1 - \Phi(2.05) = 1 - 0.9798 = 0.0202$.

Pertanto:

$$z = \frac{5 - \mu}{\sigma} = \frac{5 - \mu}{0.05} = -2.05.$$

Risolviendo l'equazione in μ , si trova che $\mu = 5.1$ litri.

Esempio 5.2 Vediamo in pratica come utilizzando il teorema di De Moivre - Laplace, la distribuzione normale possa essere utilizzata per approssimare il calcolo di una distribuzione binomiale. In questo caso bisogna sempre tener presente che la Binomiale è una distribuzione discreta per cui ha senso calcolare il valore di probabilità $P(K = k)$ in un punto k , mentre la Normale è una densità continua dove $P(X = x)$ ha misura nulla e possiamo solo calcolare un valore di probabilità in un intervallo $P(X \in dx) \approx f(x)dx$

Supponiamo di avere in magazzino una scorta di 100 chip di memoria di cui statisticamente il 20% risulta difettoso.

Vogliamo calcolare la probabilità che esattamente 15 pezzi siano difettosi

Il problema è modellabile mediante una distribuzione binomiale dove:

$$100 \text{ pezzi} \rightarrow n = 100$$

$$20\% \text{ difettosi} \rightarrow p = 0.20$$

Poiché n è sufficientemente grande approssimiamo con una Gaussiana $\mathcal{N}(x \mid \mu = np, \sigma^2 = npq)$ dove

$$\mu = np = 20$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0.20 \cdot 0.8} = 4$$

Il problema ci richiede di calcolare $P_{\text{Bin}}(X = 15)$. Approssimando con la Normale dobbiamo tener presente che per densità continue non ha senso determinare $P_{\mathcal{N}}(X = 15)$ giacché ha misura nulla: applichiamo dunque la **correzione di continuità**:

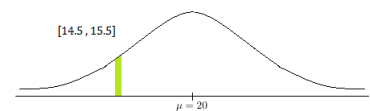


Figura 12: Correzione di continuità per approssimare $P_{\text{Bin}}(X = 15)$

$$P_{Bin}(X = 15) \approx P_N(14.5 \leq X \leq 15.5) = \Phi\left(\frac{15.5 - \mu}{\sigma}\right) - \Phi\left(\frac{14.5 - \mu}{\sigma}\right)$$

Procediamo alla standardizzazione

$$z_1 = \frac{14.5 - 20}{4} = -1.38$$

$$z_2 = \frac{15.5 - 20}{4} = -1.13$$

Usando z_1, z_2 , e accedendo alle tabella della CDF

$$\Phi\left(\frac{15.5 - \mu}{\sigma}\right) - \Phi\left(\frac{14.5 - \mu}{\sigma}\right) = F_Z(-1.13) - F_Z(-1.38) = 1 - F_Z(1.13) - 1 + F_Z(1.38) = 0.0454$$

Esempio 5.3 L'altezza di un uomo segue una distribuzione normale di media 178 cm con una deviazione dalla media di 8 cm. L'altezza di una donna ha invece una media di 165 cm con una deviazione pari a 7 cm.

1) Qual é, nella stessa popolazione, la proporzione di uomini la cui altezza é superiore a 185 cm?

Sia U la variabile aleatoria che denota l'altezza degli uomini.

Vogliamo determinare $P(U > 185)$, dove la distribuzione é Normale con $\mu = 178, \sigma = 8$. Standardizzando le variabili e usando le tabelle della distribuzione normale

$$P(U > 185) = P\left(\frac{U - 178}{8} > \frac{185 - 178}{8}\right) = P(Z < 0.875) \approx 0.19$$

2) Qual é la proporzione di donne che sono piú alte della metà degli uomini?

Sia D la variabile aleatoria che denota l'altezza delle donne.

Sappiamo che media e mediana nella distribuzione normale coincidono dunque

$$S_U(u) = P(U > u) = 0.5,$$

quando $u = \mu = 178$

Pertanto, si tratta di determinare $P(D > 178)$:

$$P(D > 178) = P\left(Z > \frac{178 - 165}{7}\right) = P(Z > 1.857) \approx 0.032$$

Solo il 3.2% delle donne é piú alto di metà degli uomini.

6 Appendice: calcolo dell'integrale di Gauss

Per poter "manipolare" matematicamente la Gaussiana, é necessario saper risolvere l'integrale di Gauss:

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx.$$

Per calcolarlo, si sviluppa I^2 :

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} e^{-x^2} dx \cdot \int_{-\infty}^{+\infty} e^{-y^2} dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2} \cdot e^{-y^2} dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy \end{aligned}$$

Si passa poi in coordinate polari (r, θ) dove $x = r \cos \theta$ $y = r \sin \theta$ corrisponde a θ .

Nella trasformazione, l'elemento infinitesimo di area dA deve conservarsi, ovvero:

$$dA = dx dy = |J| dr d\theta$$

dove $|J|$ è il determinante della matrice J detta matrice jacobiana.

$$J = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}.$$

Gli elementi della matrice jacobiana sono le derivate prime parziali di $x = x(r, \theta)$, $y = y(r, \theta)$ rispetto alle nuove coordinate (r, θ) . Ricordando la definizione di derivata parziale di una generica funzione $f = f(x, y)$

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}.$$

$$\frac{\partial f}{\partial y} = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h}.$$

applicandola al nostro caso otteniamo

$$\frac{\partial x}{\partial r} = \cos \theta.$$

$$\frac{\partial x}{\partial \theta} = -r \sin \theta.$$

$$\frac{\partial y}{\partial r} = \sin \theta.$$

$$\frac{\partial y}{\partial \theta} = r \cos \theta.$$

La matrice Jacobiana diventa:

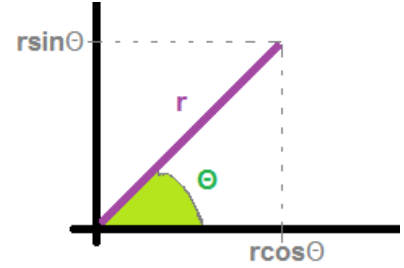


Figura 13: Coordinate polari

$$J = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}.$$

Calcoliamo il determinante:

$$\begin{aligned} |J| &= r \cos \theta \sin \theta - (-r \sin \theta) \sin \theta \\ &= r \cos^2 \theta + r \sin^2 \theta \\ &= r (\underbrace{\cos^2 \theta + \sin^2 \theta}_1) \\ &= r \end{aligned}$$

Quindi:

$$dx dy = r dr d\theta$$

Ora possiamo completare il calcolo dell'integrale di Gauss. Consideriamo il dominio di r e θ : $0 \leq r \leq +\infty$ e $0 \leq \theta \leq 2\pi$:

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy \\ &= \int_0^{+\infty} \int_0^{2\pi} e^{-r^2} r dr d\theta \\ &= \underbrace{\int_0^{2\pi} d\theta}_{2\pi} \cdot \int_0^{+\infty} r \cdot e^{-r^2} dr \\ &= \pi \cdot \int_0^{+\infty} 2r \cdot e^{-r^2} dr \\ &= \pi \cdot [e^{-r^2}]_{+\infty}^0 \\ &= \pi \cdot (1 - 0) \\ &= \pi \end{aligned}$$

Infine:

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

Analogamente (con sostituzione di variabili) si calcola il seguente integrale:

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi} \quad (13)$$