Distribuzioni

Distribuzione binomiale

Parecchi problemi riguardano prove ripetute ed indipendenti di un processo per cui l'esito di ogni singola prova è dicotomico:

- Si o no
- Testa o croce
- Colpito o mancato
- Ecc
- In generale : successo o fallimento

Esempio : - lancio della moneta n volte
- numero di bambini nati da un gruppo di n madri in attesa
- numero di colpi centrati dopo aver tirato n palle a un bersaglio fisso

Si vuol conoscere la probabilità di k successi (o fallimenti) in n prove , indipendentemente dall'ordine con cui capitano.

- Assumendo che per ogni evento E la probabilità di un successo sia p e di un insuccesso sia 1-p = q e che non cambino tra un tentativo e l'altro, tale probabilità è descritta dalla *distribuzione binomiale*

$$\boxed{B(n,k) = \binom{n}{k} p^k q^{n-k}.}$$

- Infatti, supponiamo che l'evento E sia il risultato del lancio di una moneta. I modi in cui può accadere sono : testa o croce

-Sia P(T) = p e P (C)= q

-La probabilità che in n lanci i primi k diano T e gli altri n-k diano C è

$$p^k q^{n-k}$$

-Questa è solo una possibile sequenza per la realizzazione di k volte T e n-k volte C.

-Tutti possibili modi sono le combinazioni di n elementi prese a k a k.
Il calcolo combinatorio ci fornisce la formula seguente:

$$B(n,k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{(n-k)!\ k!} p^k q^{n-k}$$

-*Nota*: questo è un termine dello sviluppo di un binomio elevato alla potenza n:

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

dove $\binom{n}{k}$ = combinazioni di n elementi a k a k = $\frac{n \cdot (n-1) \ldots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$
(con $k!$ si intende $k \cdot (k-1) \ldots 2 \cdot 1$, per definizione si ha $0! = 1! = 1$).

Proprietà della distribuzione binomiale

•Normalizzazione.

$\sum_{k=0}^{n} B(n,k) = 1$. Infatti:

$$\sum B(n,k) = \sum \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1^n = 1.$$

•Calcolo media

$$\begin{aligned}
\mu &= E[k] = \sum_{k=0}^{n} kB(n,k) = \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=1}^{n} k\frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 2\cdot 1} p^k q^{n-k} \\
&= np\sum_{k=1}^{n} \frac{(n-1)(n-2)\dots(n-k+1)}{(k-1)(k-2)\dots 2\cdot 1} p^{k-1} q^{n-k}.
\end{aligned}$$

Ponendo

$h = k-1$ e $m = n-1$ si ottiene:

$$\mu = np\sum_{h=0}^{m} \frac{m(m-1)\dots(m-h+1)}{h(h-1)\dots 2\cdot 1} p^h q^{m-h} = np.$$

Ricorda: $\binom{n}{k} = \frac{n\cdot(n-1)\dots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$

Proprietà della distribuzione binomiale

•Si dimostra (vedi appendice) che la varianza

$$\sigma^2 = E[(k-\mu)^2] = E[k^2] - \mu^2 = npq$$

Esempio: Sia E il lancio di un dado.
Sia n=4 il numero di volte che si lancia il dado.

-La probabilità che un numero , per esempio il 5, esca 0,1,2,3,4, volte è data dalla distribuzione binomiale con

n= 4 e k = 0,1,2,3,4 rispettivamente

Calcolo esplicito di B(4,k) per ogni valore di k

$$\begin{aligned}
B(4,0) &= \binom{4}{0}\left(\frac{5}{6}\right)^4 = \frac{5^4}{6^4} = \frac{625}{6^4},\\
B(4,1) &= \binom{4}{1}\frac{1}{6}\left(\frac{5}{6}\right)^3 = 4\frac{5^3}{6^4} = \frac{500}{6^4},\\
B(4,2) &= \binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 = \frac{4\cdot 3}{2}\frac{5^2}{6^4} = \frac{150}{6^4},\\
B(4,3) &= \binom{4}{3}\left(\frac{1}{6}\right)^3\frac{5}{6} = \frac{4\cdot 3\cdot 2}{3\cdot 2}5\frac{1}{6^4} = \frac{20}{6^4},\\
B(4,4) &= \binom{4}{4}\left(\frac{1}{6}\right)^4 = \frac{1}{6^4},\\
\mu &= np = 4\frac{1}{6} \approx 0.667,\\
\sigma &= \sqrt{4\frac{1}{6}\frac{5}{6}} \approx 0.7.
\end{aligned}$$

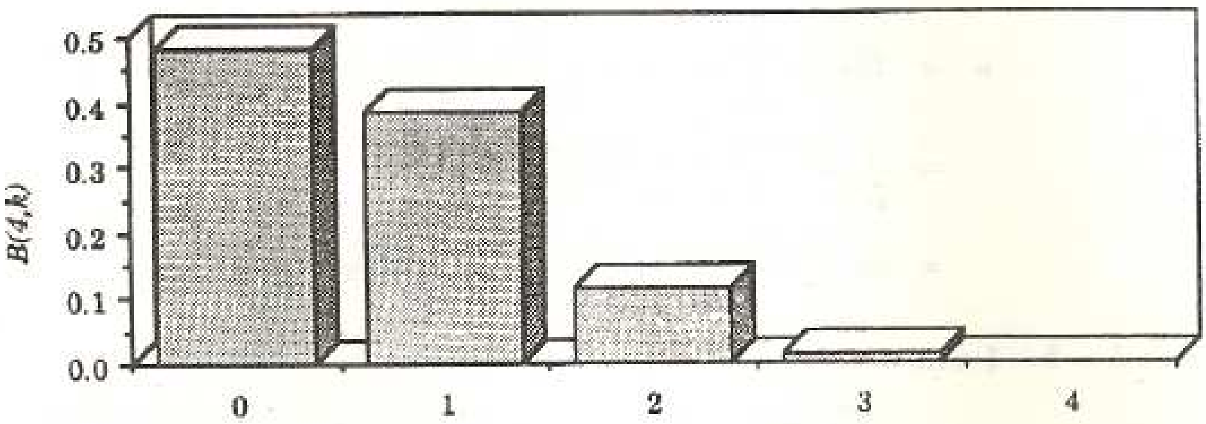


Figura 3.1: grafico di $B(4,k)$.

Esercizio (per casa): costruire sperimentalmente la distribuzione B(4,k) operando così:

- Fare N prove, ogni prova sia un lancio di 4 dadi.
- Contare quanti 5 escono in una prova.
- Contare quante prove contengono k 'cinque'.
- Dividere per il numero di prove, in modo da ottenere la frequenza relativa, che è una stima della probabilità.
- Rappresentare le frequenze relative e confrontarle con quelle teoriche.

Esercizio : Una madre da il permesso alla figlia di andare a comprare una maglietta .
Il commesso prepara una file di 15 bellissime magliette e la figlia le prova 1 ad 1 e con una probabilità 0.7 le scarta perché non le piacciono, altrimenti le compra.

Qual è la probabilità che torni a casa con 6 magliette?

R : La scelta è binaria: compra – non compra . Si usa binomiale

$$B(15,6) = \begin{bmatrix} 15 \\ 6 \end{bmatrix} (0.3)^6 (0.7)^{15-6} = \frac{15!}{6!\,9\,!} (0.3)^6 (0.7)^9 =$$

$$=0.147$$

Definizioni del numero e

$$e \stackrel{\text{def}}{=} \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \approx 2.71828182845904523536028747135266249775572470\ldots$$

Si possono anche dimostrare le seguenti uguaglianze

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

Da cui per x = 1 si ottiene una formula da cui è facile valutare *e*

$$e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} + \cdots$$

Distribuzioni

Distribuzione di Poisson

La distribuzione di Poisson si presenta come forma limite della distribuzione binomiale quando il numero medio dei successi è molto più piccolo del numero degli eventi possibili. Vediamo.

Esempio : variabile casuale : numero di studenti che si immatricolano ogni anno a chimica.

-Supponiamo di avere sul monitor del computer un elenco di tutti gli immatricolati a 'scienze MFN' e di scorrerlo alla ricerca degli immatricolati a chimica.

-Ogni studente dell'elenco (N=1000, buona stima degli immatricolati) rappresenta un tentativo : si ha un successo quando lo studente è di chimica.

-La probabilità a priori di un successo per ogni singola prova può essere valutata dal rapporto:
Nro medio studenti che ogni anno si iscrivono a chimica / nro totale degli studenti immatricolati a scienze
Si può stimare che tale frazione valga $p = 35/1000$.

-Si parte dalla binomiale con l'ipotesi $p \ll 1$, $n \gg 1$, $n \gg k$

$$B(n,k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{(n-k)!\ k!}\ p^k\ q^{n-k}$$

-Con l'ipotesi p<<1, n>>1, n>>k il rapporto di fattoriali diventa:

$$\frac{n!}{k!(n-k)!} = \frac{1}{k!}\ \frac{n!}{(n-k)!} = \frac{1}{k!}\ n(n-1)(n-2)..(n-k+1) \cong \frac{n^k}{k!}$$

Ricordando che $\mu = np$ e quindi che $\mu^k = (np)^k$ si ha

$$P(k,n) = \frac{\mu^k}{k!}\ (1-p)^{n-k}$$

L'ultimo termine può essere riscritto, con l'uso del teorema binomiale come

$$(1-p)^{n-k} = (1-p)^n (1-p)^{-k} \cong (1-p)^n (1+pk)$$

Poiché $n = \mu/p$

$$(1-p)^{n-k} = [(1-p)^{1/p}]^{\mu}\ (1+kp)$$

Passando al limite per p ->0

$$(1-p)^{n-k} = e^{-\mu}$$

e infine

$$\boxed{p(k) = \frac{\mu^K}{k!} e^{-\mu}}$$

Proprietà della distribuzione poissoniana

$$\boxed{P(a,k) = \frac{a^k}{k!}e^{-a}}$$

•Normalizzazione.

$$\sum_{k=0}^{\infty} P(a,k) = \sum_{k=0}^{\infty} \frac{a^k}{k!}e^{-a} = e^{-a}\sum_{k=0}^{\infty}\frac{a^k}{k!} = e^{-a}\cdot e^{a} = 1$$

dove si è usato la definizione di *e*. $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}.$

•Calcolo media

$$\mu = E[k] = \sum_{k=0}^{\infty} k\,P(a,k) =$$

$$= \sum_{k=0}^{\infty} k\frac{a^k}{k!}e^{-a} = a\sum_{k=1}^{\infty}\frac{a^{k-1}}{(k-1)!}e^{-a}$$

$$= a\sum_{h=0}^{\infty}\frac{a^h}{h!}e^{-a} = a.$$

dove si è posto $h = k-1$, cioè $k = h+1$.

Proprietà della distribuzione poissoniana

$$\boxed{P(a,k) = \frac{a^k}{k!}e^{-a}}$$

•calcolo della varianza

$$\begin{aligned}
\sigma^2 &= E[(k-\mu)^2] = E[k^2] - \mu^2, \\
E[k^2] &= \sum_{k=0}^{\infty} k^2 P(\mu,k) = \mu \sum_{k=1}^{\infty} k \frac{\mu^{k-1}}{(k-1)!} e^{-\mu} \\
&= \mu \sum_{h=0}^{\infty} (h+1) \frac{\mu^h}{h!} e^{-\mu} \\
&= \mu \left[\Sigma h \cdot P(\mu,h) + \Sigma P(\mu,h)\right] \\
&= \mu(\mu+1) = \mu^2 + \mu.
\end{aligned}$$

Quindi $\sigma^2 = \mu^2 + \mu - \mu^2 = \mu.$

$$\mu = a, \qquad \sigma^2 = \mu = a, \qquad \sigma = \sqrt{\mu}.$$

Esercizio: vengono trasmessi 10^5 bits binari (0, 1) e la probabilità di errore nella trasmissione di ogni bit è 10^{-3} , indipendentemente dagli altri .
Calcolare la probabilità che *k* bits siano errati

R : Si ha una poissoniana con media $\mu = n*p = 10^{5} * 10^{-3}$

$$P(k) = e^{-\mu} * \mu^{k} / (k!) = = e^{-100} * 100^{k} / (k!)$$

Distribuzioni

Distribuzione di Gauss

- *Premessa matematica*: studio della funzione $f(x) = e^{-x^2}$

❶ $f(x)$ è pari cioè $f(x) = f(-x)$.

❷ $f(x)$ è definita per $x \in [-\infty \ldots +\infty]$.

❸ $\forall x,\ f(x) \leq f(0) = 1$, quindi $x = 0$ è un punto di massimo.

❹ $\lim_{x \to \pm\infty} f(x) = 0$.

❺ $f'(x) = -2x \cdot e^{-x^2} = 0$ solo per $x = 0$,

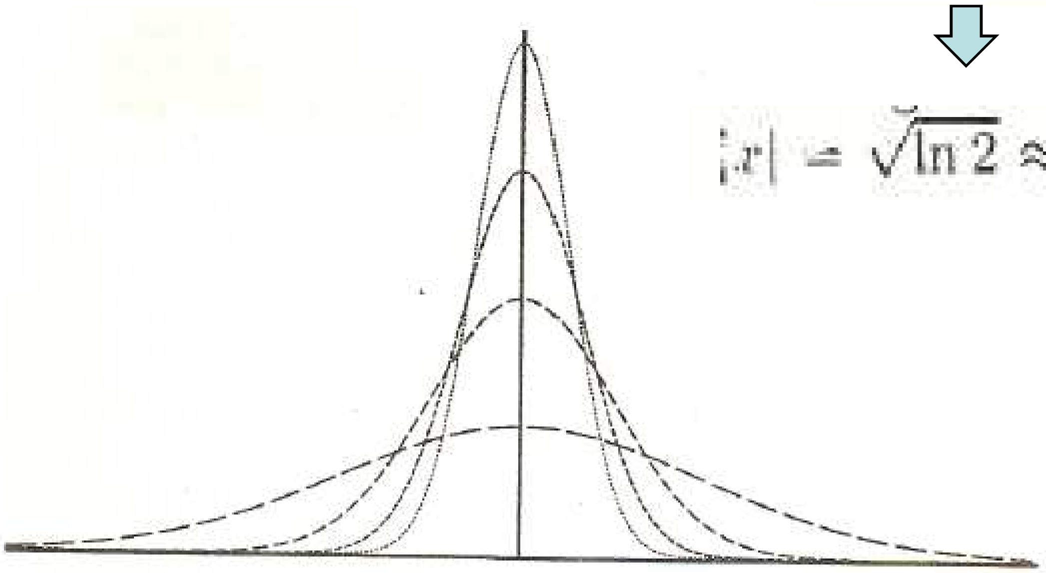
$f' > 0$ per $x < 0$ e $f' < 0$ per $x > 0$,

❻ $f(x)$ ha dunque la tipica forma di una curva a campana:

semilarghezza a metà altezza.
il valore per cui $f(x) = 1/2$: $e^{-x^2} = 1/2$

$$x^2 = \ln 2$$

$$|x| = \sqrt{\ln 2} \approx 0.83:$$



❼ $f(x)$ decresce in modo estremamente rapido al crescere di x,

per es. $f(3) = e^{-9} \approx 1.234 \times 10^{-4}$.

❽ $\int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi}$.

Infatti, chiamando I questo integrale, si ha

$$I^2 = \int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx \cdot \int\limits_{-\infty}^{+\infty} e^{-y^2}\,dy = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} e^{-(x^2+y^2)}\,dx\,dy$$

che è un integrale doppio su tutto il piano xy.

Passando in coordinare polari

l'elemento di superficie diventa $r\,dr\,d\theta$:

$$I^2 = \int\limits_{0}^{2\pi} d\theta \cdot \int\limits_{0}^{\infty} r e^{-r^2}\,dr = 2\pi \cdot \left[-\frac{e^{-r^2}}{2}\right]_0^{\infty} = \pi.$$

Distribuzione di Gauss

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

con σ e μ costanti di cui si chiarirà il significato.

E' *continua*, e rappresenta una *densità* di *probabilità*

E' legata alla funzione e^{-x^2} col cambiamento di variabile

$$t = \frac{1}{\sqrt{2}} \frac{x-\mu}{\sigma}$$

si vede che $G(t) = \mathrm{Cte} \cdot e^{-t^2}$

La semilarghezza a metà altezza è:

$$t = \frac{1}{\sqrt{2}} \frac{x-\mu}{\sigma} = \sqrt{\ln 2}, \qquad x - \mu = \sigma\sqrt{2\ln 2} \approx 1.2\sigma.$$

Ne segue che μ rappresenta il centro della campana

e σ dà una misura della larghezza.

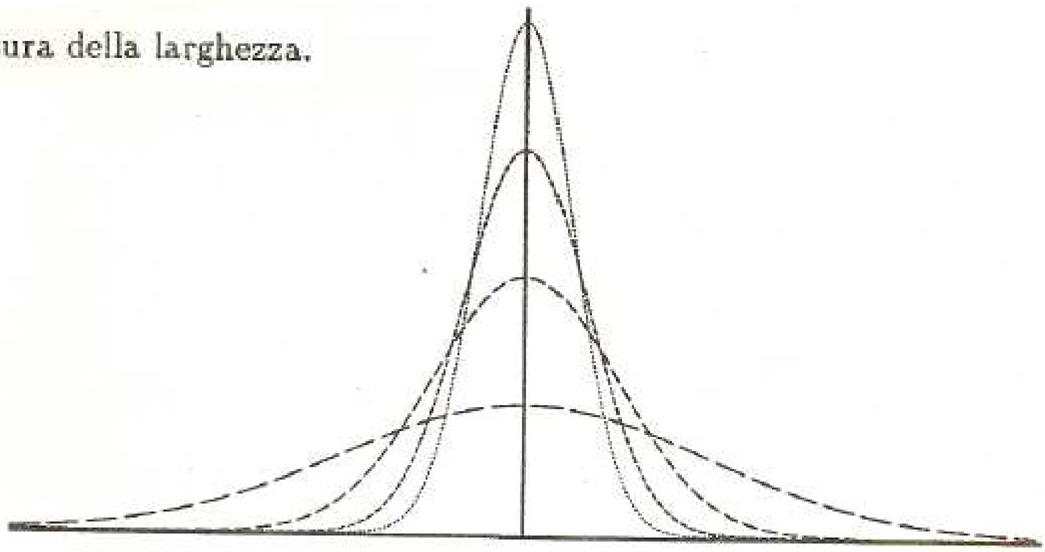


Figura 3.2: Grafici della funzione $G(x)$ per $\sigma = 1, \frac{4}{3}, 2, 4$ e $\mu = 0$.

E' *normalizzata*, come tutte le distribuzioni

$$\int_{-\infty}^{+\infty} G(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx = 1$$

Si dimostra usando il cambiamento di variabile $t = \frac{1}{\sqrt{2}}\frac{x-\mu}{\sigma}$
e il calcolo precedente di I^2

Media $$E[x] = \int_{-\infty}^{+\infty} x \cdot G(x)\,dx = \mu.$$

Si dimostra per esempio calcolando direttamente l'integrale
Nota: la costante μ è proprio la *media*

Varianza $$E[(x-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 G(x)\,dx = \sigma^2.$$

Col cambiamento di variabile $t = \frac{1}{\sqrt{2}}\frac{x-\mu}{\sigma}$

da cui $dx = \sigma\sqrt{2}\,dt$, si ha

$$\begin{aligned} E[(x-\mu)^2] &= \int_{-\infty}^{+\infty} (x-\mu)^2 G(x)\,dx = \int_{-\infty}^{+\infty} 2t^2\sigma^2 \cdot \frac{e^{-t^2}}{\sigma\sqrt{2\pi}} \cdot 2\sigma\sqrt{2}\,dt \\ &= \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t \cdot 2te^{-t^2}\,dt \\ &= \frac{\sigma^2}{\sqrt{\pi}} \left(\left| -te^{-t^2} \right|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-t^2}\,dt \right) \\ &= \sigma^2. \end{aligned}$$

Nota: la costante σ è proprio la *deviazione standard*

La probabilità che la variabile *x* sia compresa in un generico intervallo [0, a] è

$$P(0 \le x \le a) = \int_0^a G(x)\,dx.$$

Questo integrale non ha espressione analitica

Tavole dell'integrale per una variabile gaussiana z in forma standard cioè con media μ=0 e deviazione standard σ =1

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

e la funzione di distribuzione diventa $G(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

Esempio sull'uso delle tavole: supponiamo di avere una variabile gaussiana con media μ = 20 e deviazione standard σ = 4

(a) Qual'è la probabilità che $22 \le x \le 24$?

$$z = \frac{x - 20}{4}, \qquad z_< = \frac{22-20}{4} = \frac{1}{2}, \quad z_> = \frac{24-20}{4} = 1.$$

Per cui $P(22 \le x \le 24) = P(1/2 \le z \le 1) = \int_{z_<}^{z_>} G(z)\,dz =$

$$= \int_0^{z_>} G(z)\,dz - \int_0^{z_<} G(z)\,dz \approx$$

$$0.3413 - 0.1915 = 0.1498.$$

(*b*) Qual'è la probabilità che $15 \leq x \leq 25$?

$$z = \frac{x - 20}{4}, \qquad \begin{array}{l} z_< = \frac{15-20}{4} = -1.25 \\ z_> = \frac{25-20}{4} = +1.25. \end{array}$$

Per cui $P(15 \leq x \leq 25) = P(-1.25 \leq z \leq 1.25) = \int_{z_<}^{z_>} G(z)\,dz =$

$$= 2 \int_0^{1.25} G(z)\,dz \approx 2 \cdot 0.39435 = \; 0.7870.$$

(*c*) Qual'è la probabilità che $x > 30$?

$$z = \frac{x - 20}{4}, \qquad \begin{array}{l} z_< = \frac{30-20}{4} = 2.5 \\ z_> = \frac{\infty-20}{4} = \infty. \end{array}$$

Per cui $P(x > 30) = P(x > 2.5) = \int_{z_<}^{z_>} G(z)\,dz =$

$$= \int_0^{z_>} G(z)\,dz - \int_0^{z_<} G(z)\,dz \approx 0.5 - 0.4946 = 0.0054.$$

Nota 1: la distribuzione di Gauss si ottiene formalmente come caso limite della binomiale quando *n*->∞ e *p* resta costante .

In pratica quando *np* e *n (1-p)* sono ≥5 la gaussiana è già una buona approssimazione della binomiale.

Poiché la binomiale ha media *np* e deviazione standard $\sqrt{np(1-p)}$
La corrispondente gaussiana è:

$$G(x) = \frac{1}{\sqrt{2\pi \cdot np(1-p)}} \exp\left[-\frac{1}{2}\frac{(x-np)^2}{np(1-p)}\right]$$

Nota 2: si può dimostrare che la distribuzione di Gauss si ottiene anche come limite della poissoniana per grandi valori della media μ

In questo caso la deviazione standard vale $(\mu)^{1/2}$
Per cui la corrispondente gaussiana è:

$$\frac{1}{\sqrt{2\pi \cdot np(1-p)}} e^{-\frac{1}{2}\frac{(x-np)^2}{np(1-np)}}$$

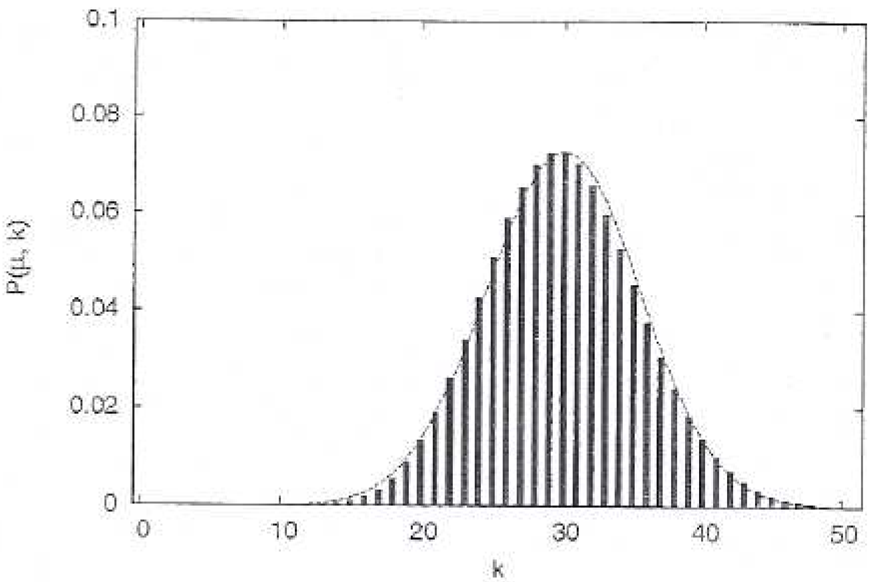


Figura 3.8: *Confronto tra una Poissoniana con media $\mu = 30$ ed una Gaussiana con media $\mu = 30$ e deviazion e standard $\sigma = \sqrt{30}$. Le due distribuzioni sono virtualmente indistinguibili.*

Esempio (nota1): Si calcoli la probabilità che , lanciando 10 monete, il numero t di teste sia compreso tra 3 e 6.

R: Usando la ***statistica binomiale*** si ha, con n = 10 e p=1/2

$$B(t) = \binom{10}{t}(1/2)^t(1 - 1/2)^{10-t}$$

$$B(3 \le t \le 6) = B(3) + B(4) + B(5) + B(6),$$

$$B(3) = \binom{10}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^7 = \frac{15}{128}$$

$$B(4) = \binom{10}{4}\left(\frac{1}{2}\right)^{10} = \frac{105}{512}$$

$$B(5) = \binom{10}{5}\left(\frac{1}{2}\right)^{10} = \frac{63}{256}$$

$$B(6) = \binom{10}{6}\left(\frac{1}{2}\right)^{10} = \frac{105}{512}$$

$$B(3 \le t \le 6) \approx 0.7734$$

Per usare ***l'approssimazione gaussiana*** , t deve essere considerata *continua* e non discreta. In questo schema $3 \le t \le 6$ diviene $2.5 \le t \le 6.5$

$$\mu = np = 10 \cdot \frac{1}{2} = 5 \qquad \sigma = \sqrt{np(1-p)} = \sqrt{5 \cdot \frac{1}{2}} \approx 1.58$$

Passando alla variabile standard gli estremi di integrazione divengono:

$$z_1 = \frac{2.5 - 5}{1.58} \approx -1.58$$

$$z_2 = \frac{6.5 - 5}{1.58} \approx 0.95$$

ed infine:

$$P(2.5 \le k \le 6.5) = 0.4429 + 0.3289 = 0.7718$$

Che è molto vicino al valore 0.7734 trovato prima.

Esempio (nota2): Supponiamo che in un dato campione di materiale radiattivo vi siano in media 30 decadimenti al secondo.
Qual è la probabilità che in un certo intervallo di 1 secondo il numero di decadimenti sia compreso tra 30 e 35.

R: Usando la ***statistica di Poisson*** la probabilità di avere k decadimenti è $P(30,k) = \frac{30^k}{k!}e^{-30}$

$$\begin{aligned}
P(30,30) &= \frac{30^{30}}{30!}e^{-30} \approx 0.0726 \\
P(30,31) &= \frac{30^{31}}{31!}e^{-30} \approx 0.0703 \\
P(30,32) &= \frac{30^{32}}{32!}e^{-30} \approx 0.0659 \\
P(30,33) &= \frac{30^{33}}{33!}e^{-30} \approx 0.0599 \\
P(30,34) &= \frac{30^{34}}{34!}e^{-30} \approx 0.0529 \\
P(30,35) &= \frac{30^{35}}{35!}e^{-30} \approx 0.0453
\end{aligned}$$

$$P(30 \le k \le 35) = \sum_{k=30}^{35} P(30,k) \approx 0.3745$$

Per usare ***l'approssimazione gaussiana*** , analogamente all'esempio precedente 29.5$\le k \le$ 35.5

$$\mu = 30 \qquad \sigma = \sqrt{\mu} \approx 5.48$$

Passando alla variabile standard gli estremi di integrazione divengono:

$$\begin{aligned}
z_1 &= \frac{29.5-30}{5.48} \approx -0.09 \\
z_2 &= \frac{35.5-30}{5.48} \approx 1.00
\end{aligned}$$

ed infine:

$$P(29.5 \le k \le 35.5) = 0.0359 + 0.3413 = 0.3772$$

Distribuzione del χ^2

Sia x la somma dei quadrati di n variabili gaussiane standard ($\mu = 0$, $\sigma = 1$) indipendenti

$$x = \sum_{i=1}^{n} z_i^2 \qquad z_i \text{ normale } (0, 1).$$

Questa quantità si chiama χ^2 ad n gradi di libertà, è definito nell'intervallo $[0 \dots +\infty]$ e la sua funzione di distribuzione è:

$$\boxed{\wp(n, x) = C_n x^{\frac{n-2}{2}} e^{-\frac{x}{2}}}$$

dove C_n è una costante di normalizzazione che vale:

$$C_n = \frac{1}{2^{n/2}\Gamma(n/2)}.$$

Per grandi valori di n ($n \geq 30$), $\sqrt{2\chi^2} - \sqrt{2n-1}$ è quasi distribuito normalmente con media zero e deviazione standard uno.

Il valor medio della distribuzione del χ^2 è uguale al numero di gradi di libertà, la varianza è uguale a due volte il numero di gradi di libertà:

$$\mu = n, \qquad \sigma^2 = 2n$$

Anche per la distribuzione del χ^2 si trovano delle tavole che danno sia i valori della densità $\wp(n, x)$ sia i valori della probabilità

$$\mathcal{P}(n, x) = \int_0^x \wp(n, \xi)\, d\xi.$$

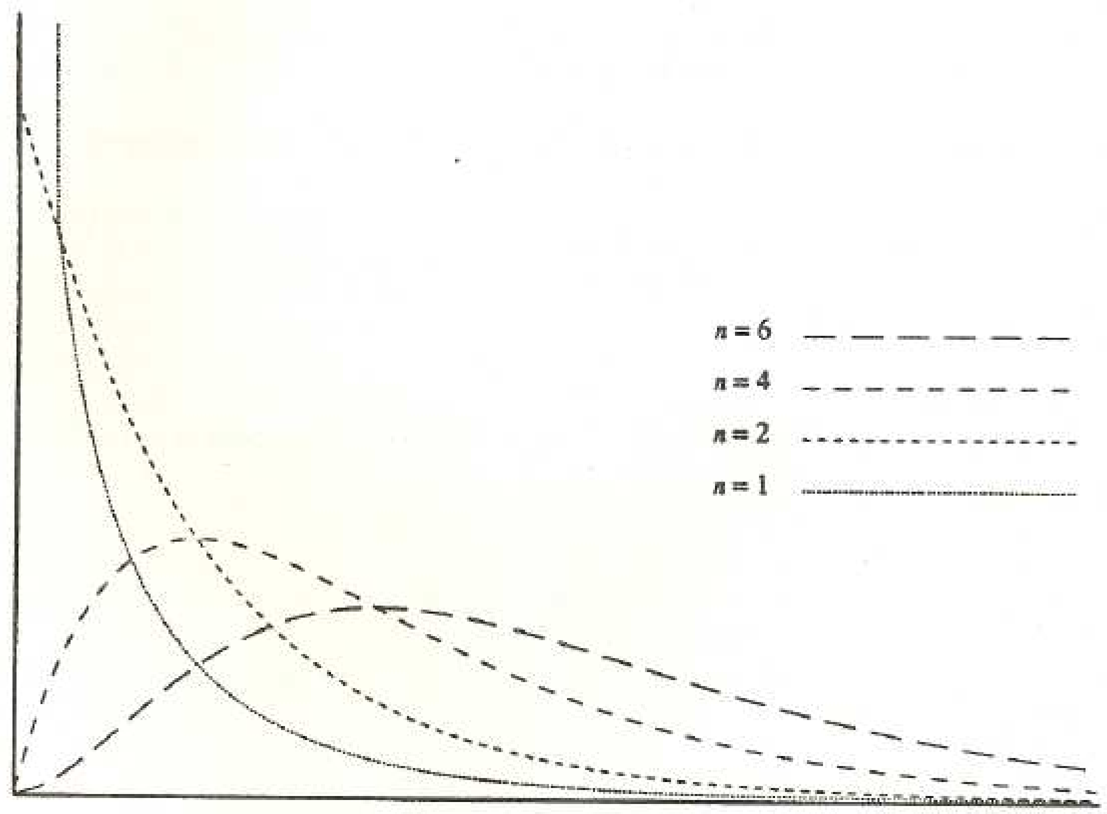


Figura 3.3: distribuzione del χ^2 per diversi gradi di libertà.

Tabella riassuntiva

distribuzione (*normalizzata ad 1*)	media	varianza
$B(n,k) = \binom{n}{k} p^k q^{n-k}$	np	npq
$P(\mu,k) = \frac{\mu^k}{k!} e^{-\mu}$	μ	μ
$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2
$p(n,x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{\frac{n-2}{2}} e^{-\frac{x}{2}}$	n	$2n$

Per $p \ll 1$ e $n \gg 1$ ma $np = \mu$ la binomiale si può approssimare con una Poissoniana:

$$\binom{n}{k} p^k q^{n-k} \approx \frac{(np)^k}{k!} e^{-np}$$

Per $np \gg 1$ la binomiale si può approssimare con una Gaussiana:

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi\, np(1-p)}} \exp\left\{-\frac{1}{2}\frac{(x-np)^2}{np(1-p)}\right\}$$

Appendice 1: calcolo della varianza per la distribuzione binomiale

$$\sigma^2 = E[(k-\mu)^2] = E[k^2] - \mu^2$$

$$E[k^2] = \sum_{k=0}^{n} k^2 \frac{n(n-1)\dots(n-k+1)}{k!} p^k q^{n-k}$$

$$= np \sum_{k=1}^{n} k \frac{(n-1)\dots(n-k+1)}{(k-1)!} p^{k-1} q^{n-k}$$

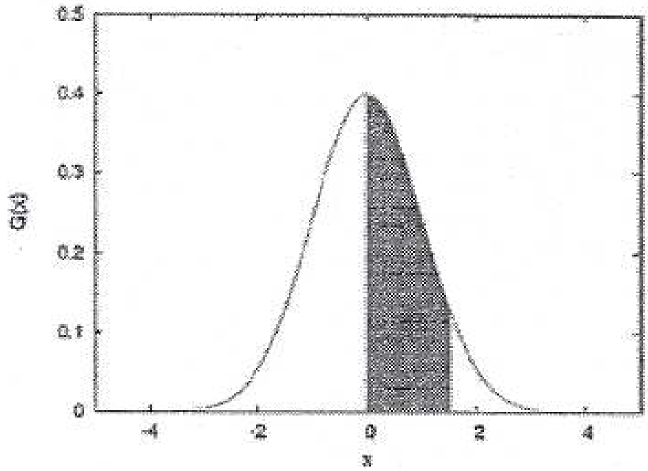
$$= np \sum_{h=0}^{m} (h+1) B(m,h) = np[\Sigma h \cdot B(m,h) + \Sigma B(m,h)] = np(mp+1).$$

Quindi $\sigma^2 = np(np - p + 1) - (np)^2 = np - np^2 = np(1-p) = npq.$

Ricorda: $\begin{pmatrix} n \\ k \end{pmatrix} = \dfrac{n \cdot (n-1)\dots(n-k+1)}{k!} = \dfrac{n!}{k!(n-k)!}$

Appendice 2

Integrale normale degli errori - I



Tavole della funzione $f(x) = erf(x)$.

Nota: le righe identificano le prime due cifre della x, mentre le colonne identificano la seconda cifra decimale della x stessa.

Esempio: la probablilità che una variabile gaussiana in forma standard sia comresa tra 0 e 1.52 è 0.4357 (sedicesima riga, terza colonna).

$\int_0^x$ Valori %

x	0	1	2	3	4	5	6	7	8	9
0.0	0.000	0.399	0.798	1.197	1.595	1.994	2.392	2.790	3.188	3.586
0.1	3.983	4.380	4.776	5.172	5.567	5.962	6.356	6.749	7.142	7.535
0.2	7.926	8.317	8.706	9.095	9.483	9.871	10.257	10.642	11.026	11.409
0.3	11.791	12.172	12.552	12.930	13.307	13.683	14.058	14.431	14.803	15.173
0.4	15.542	15.910	16.276	16.640	17.003	17.364	17.724	18.082	18.439	18.793
0.5	19.146	19.497	19.847	20.194	20.540	20.884	21.226	21.566	21.904	22.240
0.6	22.575	22.907	23.237	23.565	23.891	24.215	24.537	24.857	25.175	25.490
0.7	25.804	26.115	26.424	26.730	27.035	27.337	27.637	27.935	28.230	28.524
0.8	28.814	29.103	29.389	29.673	29.955	30.234	30.511	30.785	31.057	31.327
0.9	31.594	31.859	32.121	32.381	32.639	32.894	33.147	33.398	33.646	33.891
1.0	34.134	34.375	34.614	34.849	35.083	35.314	35.543	35.769	35.993	36.214
1.1	36.433	36.650	36.864	37.076	37.286	37.493	37.698	37.900	38.100	38.298
1.2	38.493	38.686	38.877	39.065	39.251	39.435	39.617	39.796	39.973	40.147
1.3	40.320	40.490	40.658	40.824	40.988	41.149	41.309	41.466	41.621	41.774
1.4	41.924	42.073	42.220	42.364	42.507	42.647	42.785	42.922	43.056	43.189
1.5	43.319	43.448	43.574	43.699	43.822	43.943	44.062	44.179	44.295	44.408
1.6	44.520	44.630	44.738	44.845	44.950	45.053	45.154	45.254	45.352	45.449
1.7	45.543	45.637	45.728	45.818	45.907	45.994	46.080	46.164	46.246	46.327
1.8	46.407	46.485	46.562	46.638	46.712	46.784	46.856	46.926	46.995	47.062
1.9	47.128	47.193	47.257	47.320	47.381	47.441	47.500	47.558	47.615	47.670
2.0	47.725	47.778	47.831	47.882	47.932	47.982	48.030	48.077	48.124	48.169
2.1	48.214	48.257	48.300	48.341	48.382	48.422	48.461	48.500	48.537	48.574
2.2	48.610	48.645	48.679	48.713	48.745	48.778	48.809	48.840	48.870	48.899
2.3	48.928	48.956	48.983	49.010	49.036	49.061	49.086	49.111	49.134	49.158
2.4	49.180	49.202	49.224	49.245	49.266	49.286	49.305	49.324	49.343	49.361
2.5	49.379	49.396	49.413	49.430	49.446	49.461	49.477	49.492	49.506	49.520
2.6	49.534	49.547	49.560	49.573	49.585	49.598	49.609	49.621	49.632	49.643
2.7	49.653	49.664	49.674	49.683	49.693	49.702	49.711	49.720	49.728	49.736
2.8	49.744	49.752	49.760	49.767	49.774	49.781	49.788	49.795	49.801	49.807
2.9	49.813	49.819	49.825	49.831	49.836	49.841	49.846	49.851	49.856	49.861
3.0	49.865	49.869	49.874	49.878	49.882	49.886	49.889	49.893	49.896	49.900
3.1	49.903	49.906	49.910	49.913	49.916	49.918	49.921	49.924	49.926	49.929
3.2	49.931	49.934	49.936	49.938	49.940	49.942	49.944	49.946	49.948	49.950
3.3	49.952	49.953	49.955	49.957	49.958	49.960	49.961	49.962	49.964	49.965
3.4	49.966	49.968	49.969	49.970	49.971	49.972	49.973	49.974	49.975	49.976
3.5	49.977	49.978	49.978	49.979	49.980	49.981	49.981	49.982	49.983	49.983
3.6	49.984	49.985	49.985	49.986	49.986	49.987	49.987	49.988	49.988	49.989
3.7	49.989	49.990	49.990	49.990	49.991	49.991	49.992	49.992	49.992	49.992
3.8	49.993	49.993	49.993	49.994	49.994	49.994	49.994	49.995	49.995	49.995
3.9	49.995	49.995	49.996	49.996	49.996	49.996	49.996	49.996	49.997	49.997

Appendice 3

Uso delle tabelle dell'integrale della distribuzione normale standardizzata

Esistono tabelle dell'integrale della gaussiana standardizzata, espresso in genere come

$$T(z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-\frac{z'^2}{2}} \mathrm{d}z' .$$

Da queste tabelle è possibile calcolare qualsiasi integrale facendo uso delle proprietà di simmetria della funzione di Gauss.

Un esempio è la tabella che vi è stata fornita a suo tempo.

Essa va letta nel seguente modo:

il valore di z, fino alla prima cifra decimale, è riportato nella prima colonna; la seconda cifra decimale è indicata nella prima riga delle altre colonne e, in corrispondenza di essa, è riportato il valore dell'integrale;

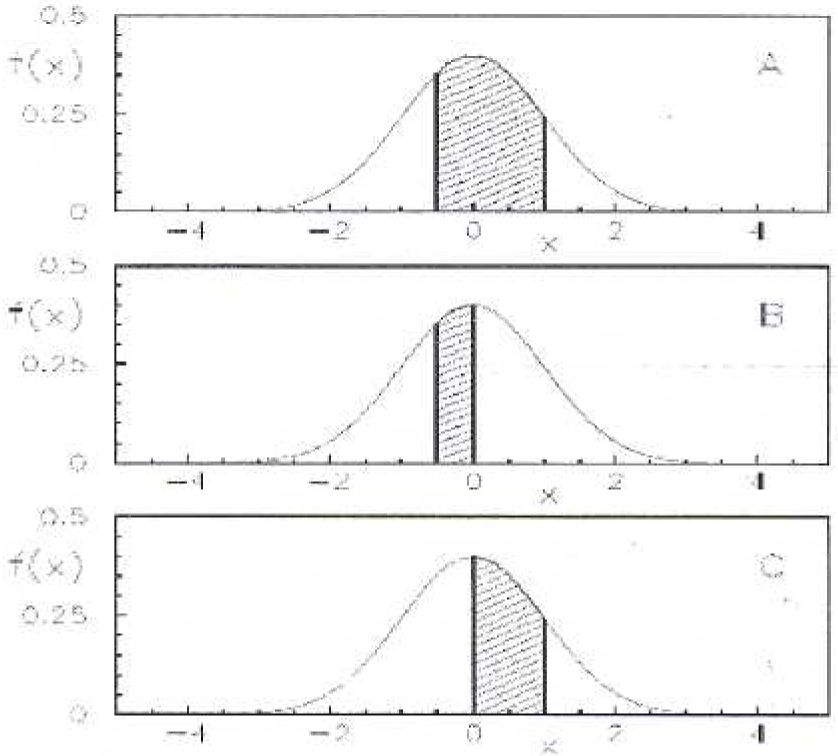


Figura 8.8: Esempio di calcolo dell'integrale della funzione normale standardizzata. L'integrale della figura A è pari alla somma di quelli di B e C, leggibili dalle tabelle.

La simmetria della distribuzione normale permette di valutare dalle tavole l'integrale su un intervallo qualsiasi. Facciamo alcuni esempi di integrali calcolati fra

z_1 e z_2, con $z_1 < z_2$:

- z_1 e z_2 positivi:

$$P(z_1 < Z < z_2) = T(z_2) - T(z_1);$$

- z_1 e z_2 negativi:

$$P(z_1 \leq Z \leq z_2) = T(-z_2) - T(-z_1);$$

- z_1 negativo e z_2 positivo:

$$P(z_1 \leq Z \leq z_2) = T(z_2) + T(-z_1)$$

Alcuni esempi sono mostrati in figura 8.8. Riportiamo la soluzione numerica ottenuta dalla tabella:

$$P(-0.5 \leq Z \leq 1.0) = 0.19146 + 0.34134 = 0.53280\,.$$

la figura sottostante illustra alcuni valori notevoli dell'integrale della gaussiana: notare che essi non dipendono dai valori di media e deviazione standard

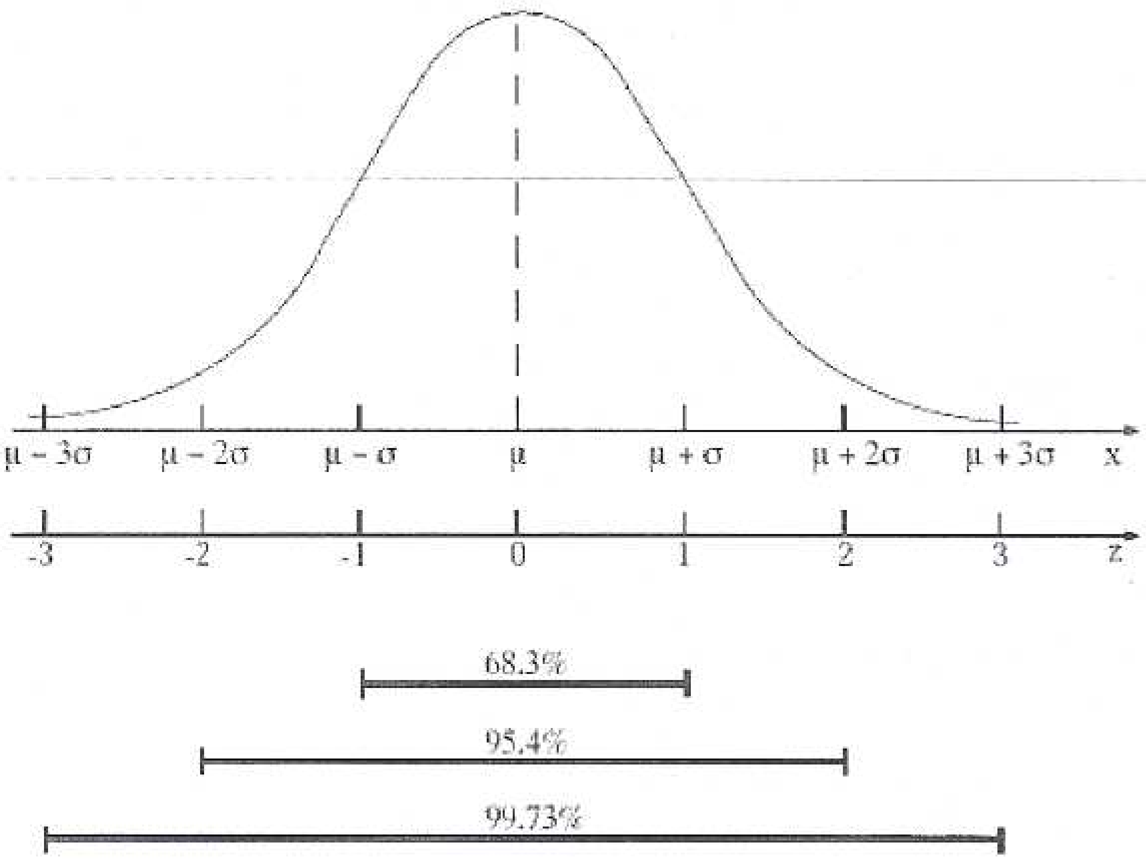
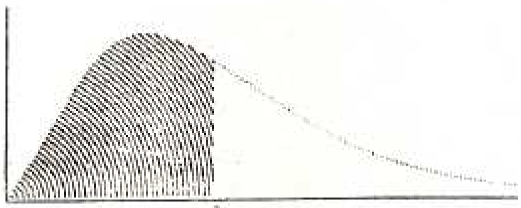


Figura 8.9: Distribuzione normale standardizzata e intervalli di probabilità.

Appendice 4

Valori per la distribuzione del chi-quadro

Valori dei χ^2_p ad n gradi di libertà per i quali l'area ombreggiata è p.



n	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.5}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.155	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	20.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	39.3	24.5	29.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.7	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.3	59.2
100	140.2	135.2	129.6	124.3	118.5	109.1	99.5	90.1	82.4	77.9	74.2	70.1	67.3

L. Andreozzi
<andreozz@df.unipi.it>