distribuzione normale

Si tratta della più importante distribuzione di *variabili continue*, in quanto:

1. si può assumere come comportamento di molti fenomeni casuali, tra cui gli errori accidentali;

2. è la forma limite di molte altre distribuzioni di probabilità;

3. trasformando opportunamente delle v.c. non normali, si possono ottenere nuove variabili distribuite normalmente;

4. sotto determinate condizioni, delle somme di v.c. possono essere approssimate da una distribuzione normale (teorema del limite centrale).

distribuzione normale

Una v.c. con media μ e varianza σ^2 (parametri della distribuzione) ha una distribuzione normale se la sua densità è data da

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

dove:

σ deviazione standard

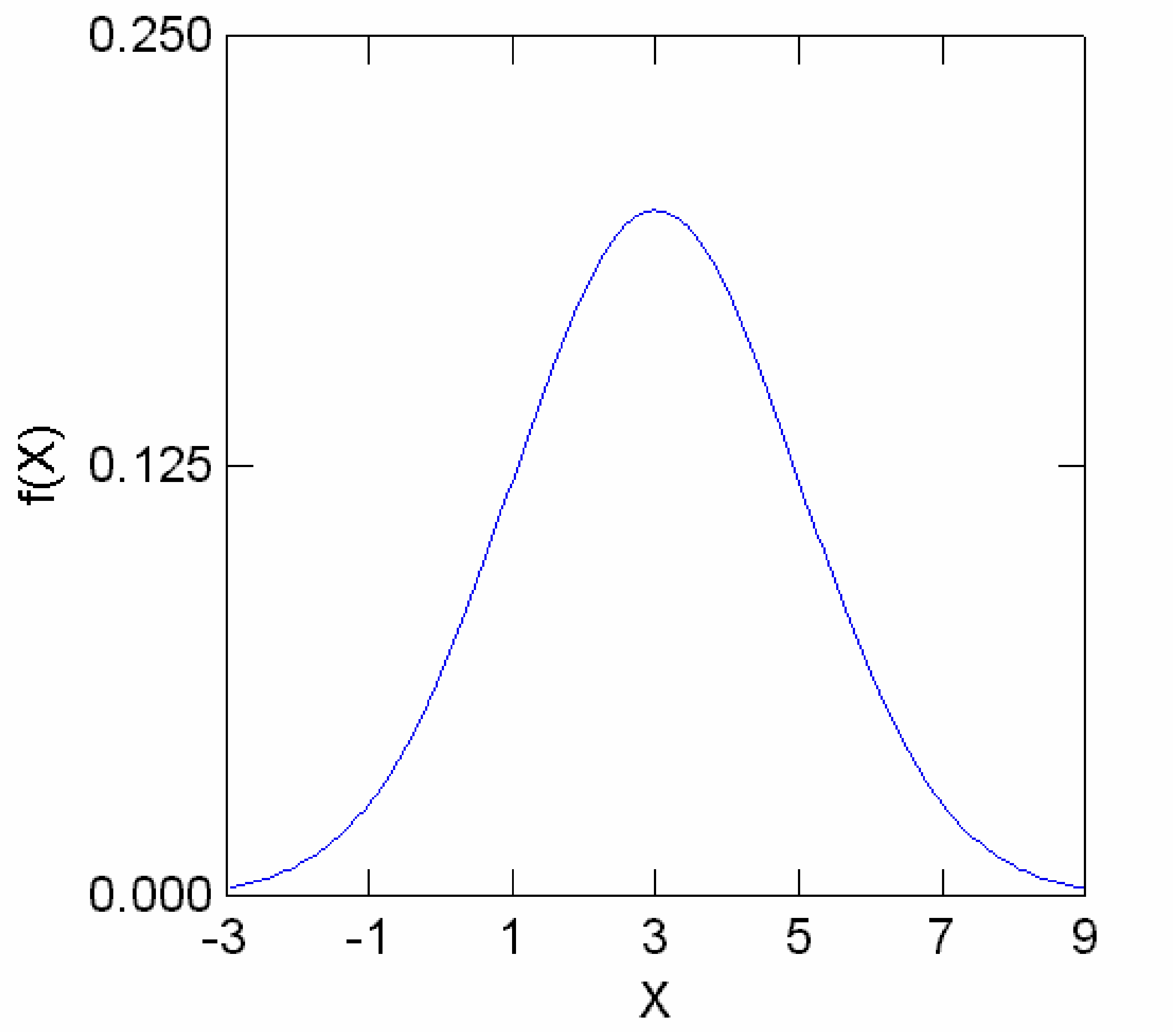
π costante = 3,146...

e base di logaritmi naturali = 2,183...

$(x-\mu)^2$ scarto dalla media elevato al quadrato

distribuzione normale

Rappresentazione grafica della distribuzione normale con media $\mu = 3$ e varianza $\sigma^2 = 4$



proprietà

1. la **distribuzione normale** è una distribuzione continua, con valori compresi tra -∞ e +∞;

2. la curva da essa descritta è simmetrica rispetto alla media (punto di ordinata massima): $f(\mu)$

3. per valori di X che vanno a -∞ oppure a +∞ la curva tende a zero senza mai toccare l'asse delle ascisse (la probabilità di ottenere valori "molto" lontani dalla media è "molto bassa");

4. è crescente per valori di X che vanno da -∞ a μ, decrescente per i valori da μ a +∞

5. presenta due punti di flesso in corrispondenza a $\mu-\sigma$ e $\mu+\sigma$, punti in cui la curva da convessa diventa concava.

ruolo dei parametri μ e σ (1/2)

- μ è la media della normale (la punta della campana) è un indicatore di posizione. Il suo variare “sposta” la campana sulla retta dei valori

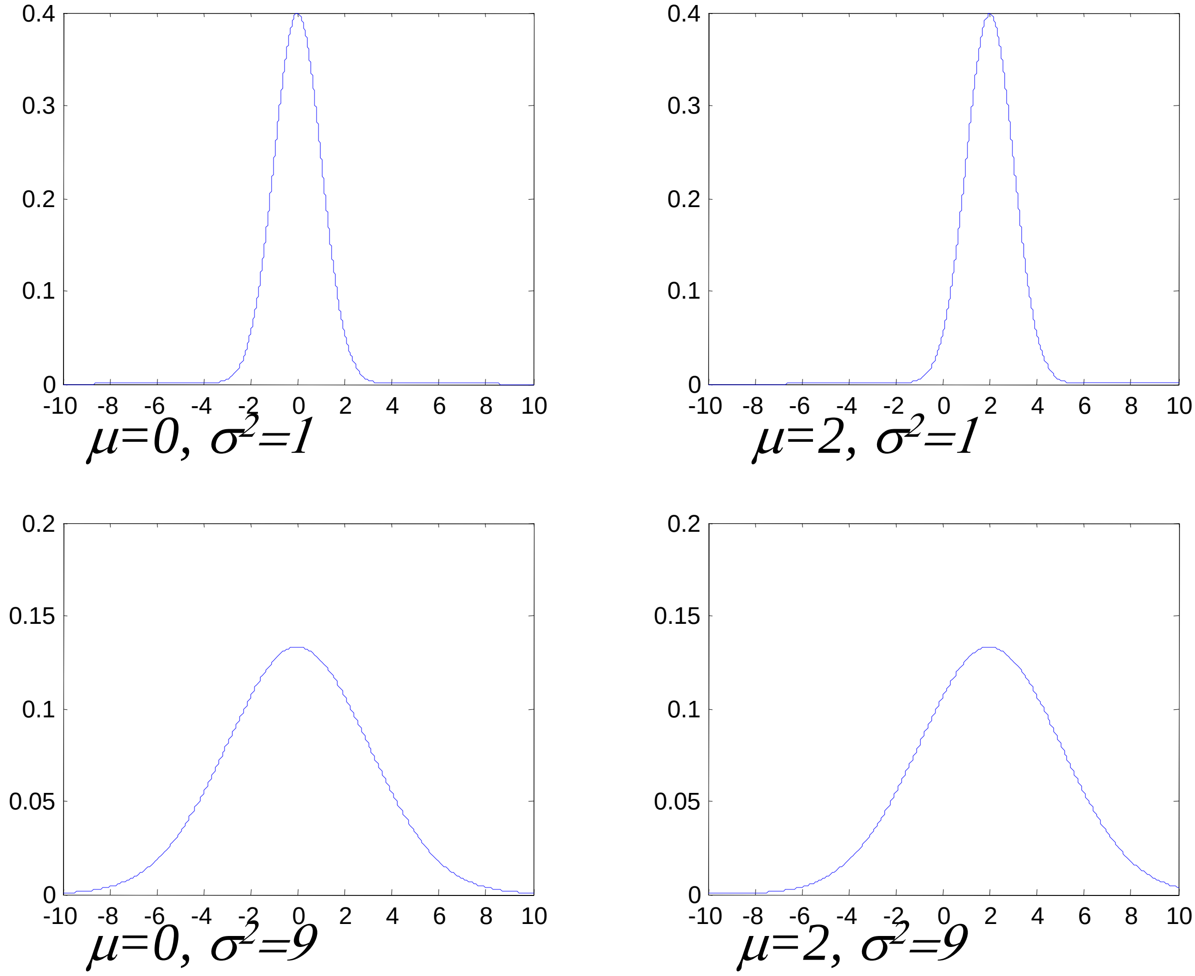
- σ^2 è la varianza è un indicatore di dispersione, è legata all’”apertura” della campana (valori più alti indicano distribuzioni più disperse).

…

ruolo dei parametri μ e σ (2/2)

Variazioni di μ →

Variazioni di σ →



distribuzione normale standardizzata

Si tratta di una distribuzione semplificata, e che consente di tabulare i valori di probabilità in apposite tavole;

si ottiene mediante una trasformazione dei dati con la formula

$$z = \frac{(X - \mu)}{\sigma}$$

di conseguenza la funzione di densità diventa:

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

con $\mu = 0$ e $\sigma = 1$.

proprietà

1. la **distribuzione normale standardizzata** è una distribuzione continua, con valori compresi tra -∞ e +∞, con **media zero** e **deviazione standard 1**;

2. la curva da essa descritta è perfettamente simmetrica rispetto al punto di ordinata massima:

3. per valori di X che vanno a -∞ oppure a +∞ la curva tende a zero senza mai toccare l'asse delle ascisse;

4. è crescente per valori di X che vanno da -∞ a *0*, decrescente per i valori da *0* a +∞

5. presenta due punti di flesso in corrispondenza a *-1* e *+1*, punti in cui la curva da convessa diventa concava.

tavole della normale

I valori delle aree della distribuzione normale standardizzata sono tabulati in apposite tavole;

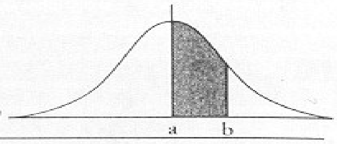
tali tavole vengono utilizzate per due scopi:

a) per calcolare l’area compresa tra due determinati valori della variabile studiata;

b) per determinare la proporzione di punteggi compresi tra due valori di una variabile casuale (distribuita normalmente)

la tavola della normale

TAV. B. *Aree della distribuzione normale standard tra a = 0 e b > 0*

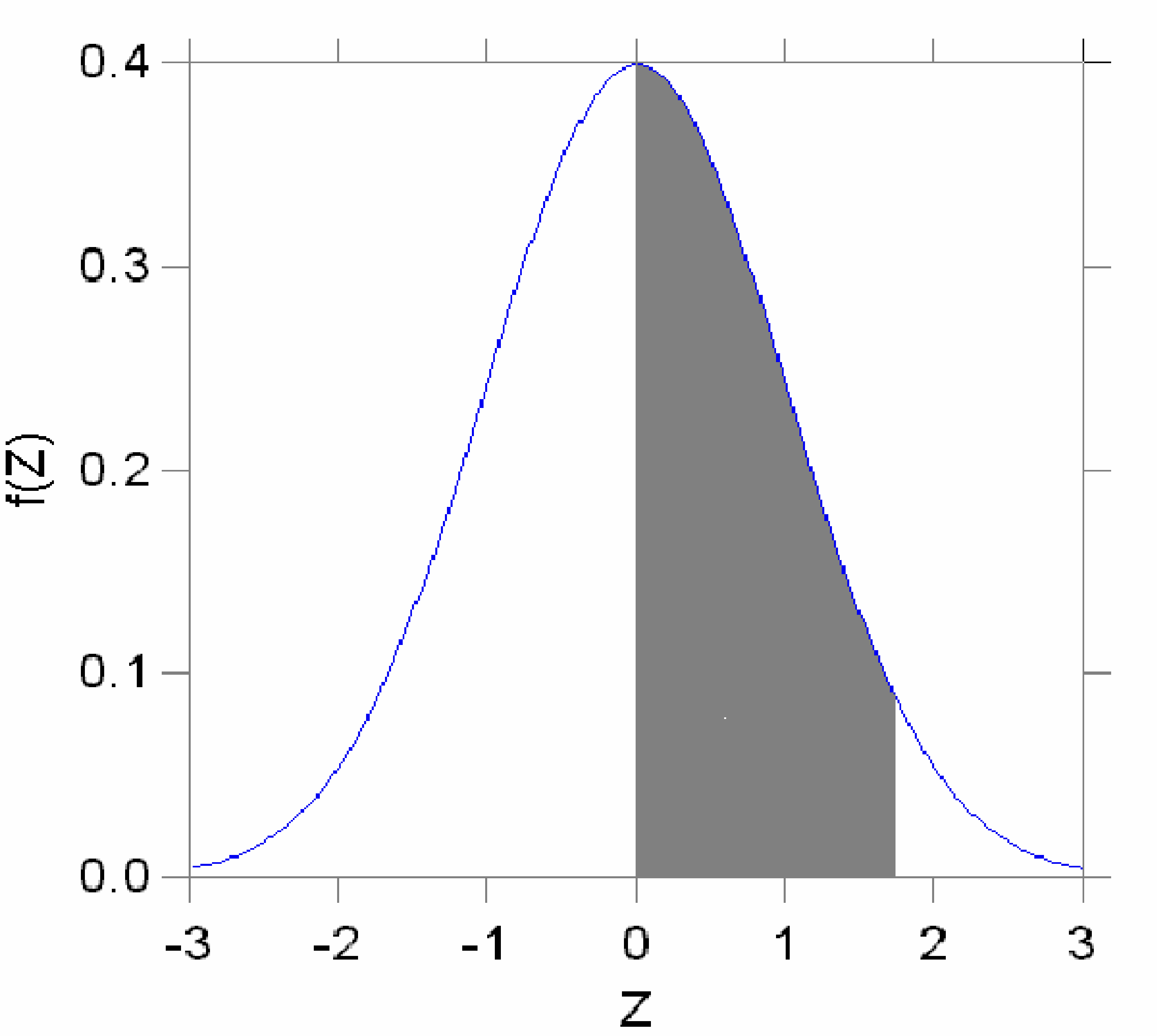


z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	8.032	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4351	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Fonte: Cristante, Lis, Sambin [1980].

esempio 3

Supponiamo di voler calcolare l'area compresa tra le ordinate $z = 0$ e $z = 1,96$.



la tavola della normale

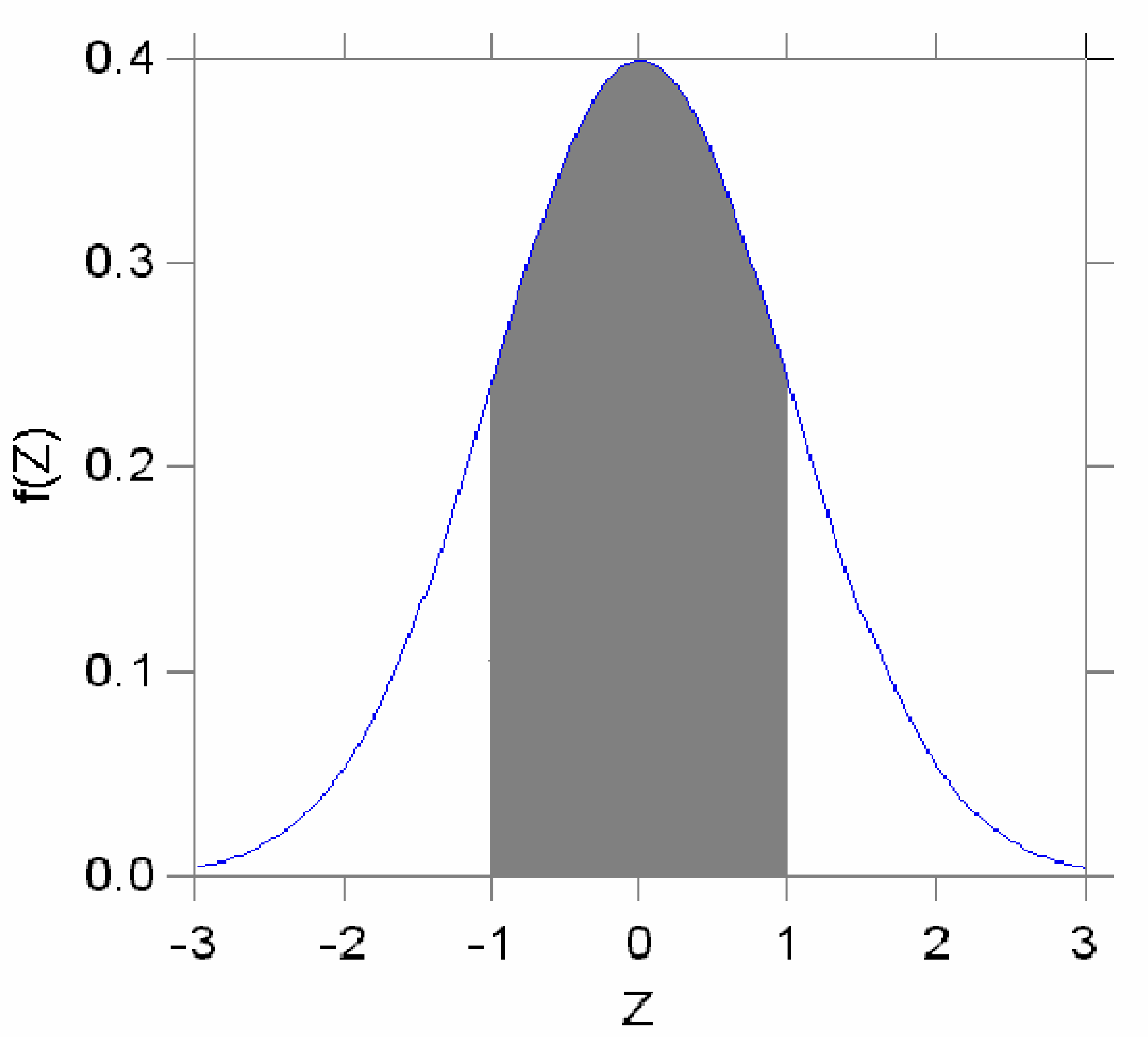
TAV. B. *Aree della distribuzione normale standard tra $a = 0$ e $b > 0$*

z	0	1	2	3	4	5	6	7
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675
0.2	.0793	8.032	.0871	.0910	.0948	.0987	.1026	.1064
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147
1.4	.4192	.4207	.4222	.4236	.4351	.4265	.4279	.4292
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808

L'area compresa tra ***z* = 0** e ***z* = 1,96** è *0,475*

esempio 4

Supponiamo di voler calcolare l'area compresa tra le ordinate $z = -1$ e $z = +1$



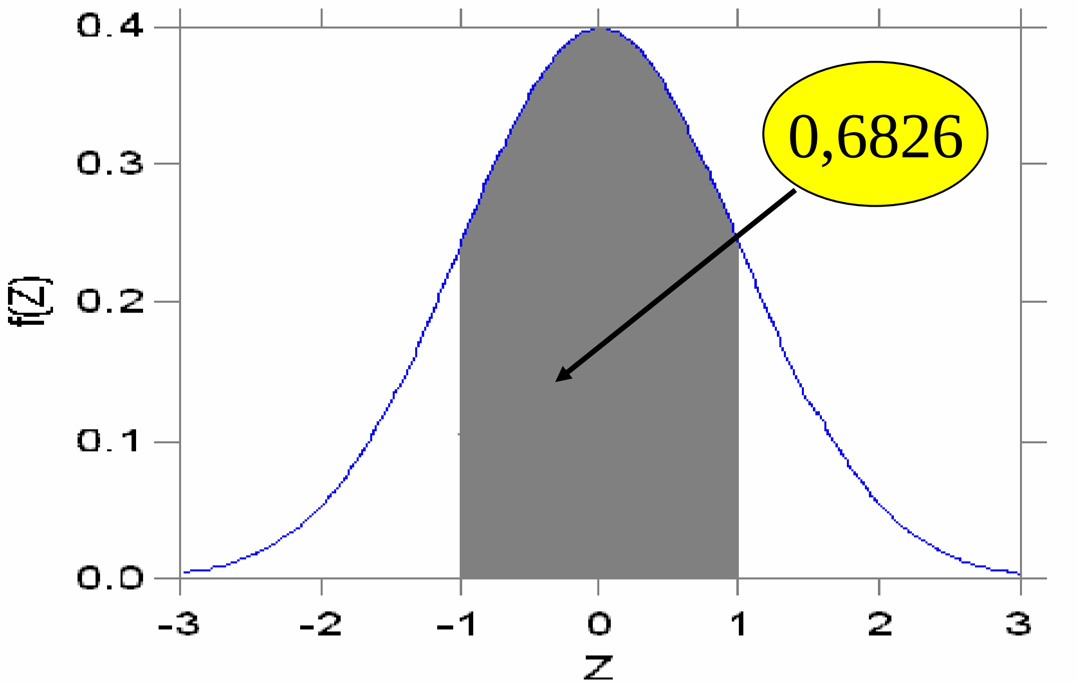
esempio 4$_{(2)}$

Dalla lettura della tavola rileviamo che l'area compresa tra $z = 0$ e $z = +1$ è pari a 0,3413;

dal momento che la curva è simmetrica e centrata sullo zero, l'area compresa tra $z = -1$ e $z = 0$ sarà identica e pari anch'essa a 0,3413;

per ottenere l'area cercata sarà sufficiente sommare le due aree:

0,3413 + 0,3413 = 0,6826



esempio 5

Data una serie di 500 punteggi distribuiti normalmente con media 100 e deviazione standard 15, si stimi quanti possano essere i punteggi compresi tra 88 e 130.

• calcoliamo i *punti zeta* corrispondenti a 88 e 130, che saranno:

$$z(88) = \frac{88 - 100}{15} = -0,8$$

$$z(130) = \frac{130 - 100}{15} = +2$$

esempio 5$_{(2)}$

• dalla lettura delle tavole ricaviamo che l'area compresa tra $z = -0{,}8$ e $z = 0$ è pari a 0,2881;

• l'area compresa tra $z = 0$ e $z = +2$ è 0,4772;

• l'area complessiva tra $z = -0{,}8$ e $z = +2$ sarà 0,2881 + 0,4772 = 0,7653;

• tale valore 0,7653 può essere letto sia come proporzione dei casi compresi tra i valori 88 e 130, sia come la probabilità che il punteggio di un soggetto cada all'interno di tale intervallo;

• in conclusione, per ottenere il numero di punteggi che ci si attende nell'intervallo compreso tra 88 e 130 si calcola:

$$0{,}7653 \times 500 = 382{,}65 \cong 383$$

distribuzione chi-quadrato

Si tratta di una distribuzione continua con densità data dalla relazione:

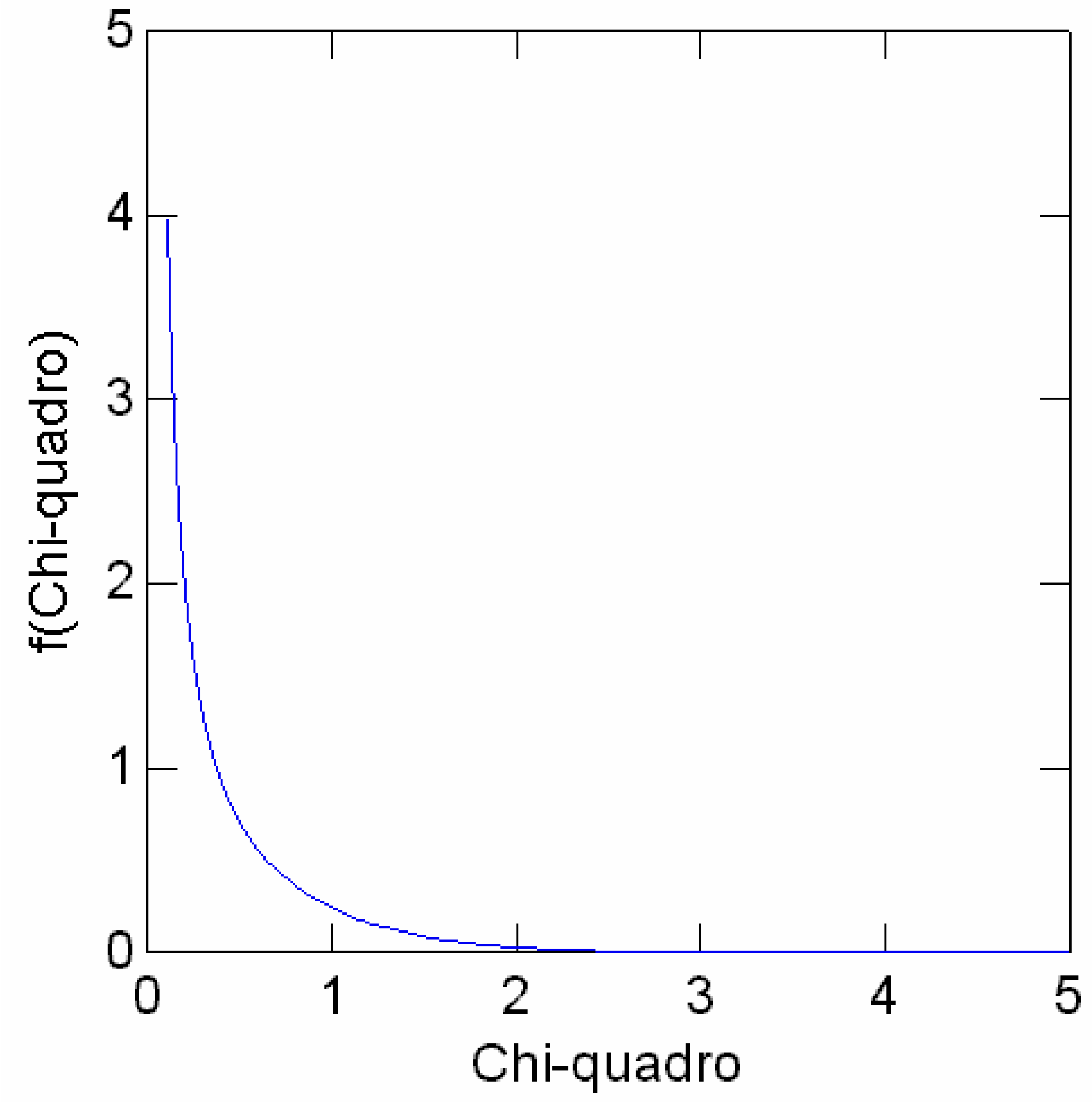
$$f\left(\chi^{2}\right)=C_{(\nu)}\left(\chi^{2}\right)^{\left(\frac{\nu}{2}\right)-1} e^{-\left(\frac{\chi^{2}}{2}\right)}$$

dove:

ν	gradi di libertà della distribuzione
$C_{(\nu)}$	costante tale da assicurare che l'area delimitata sotto la curva sia pari a 1

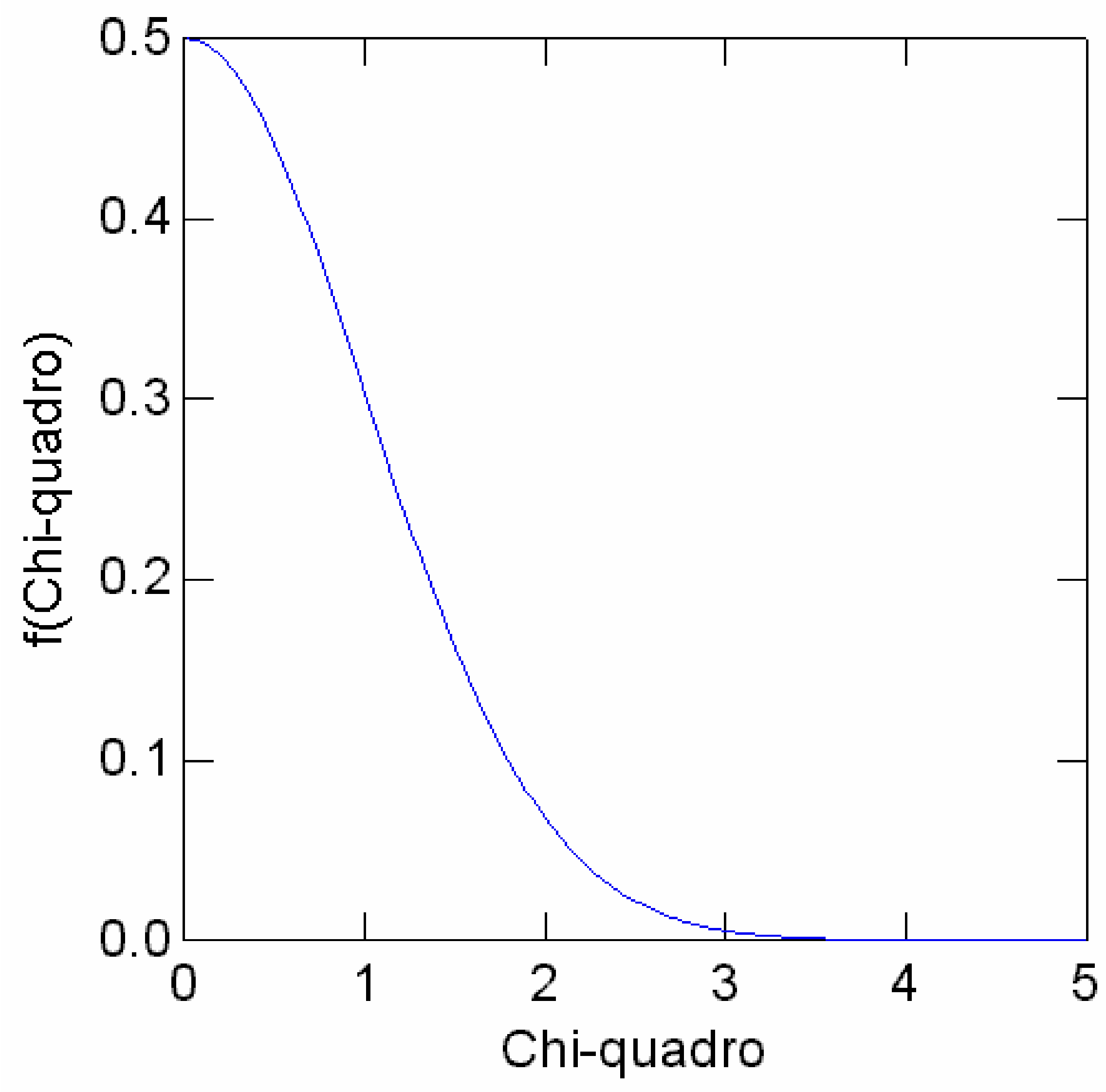
distribuzione chi-quadrato

Quando $\nu = 1$ il grafico della curva è



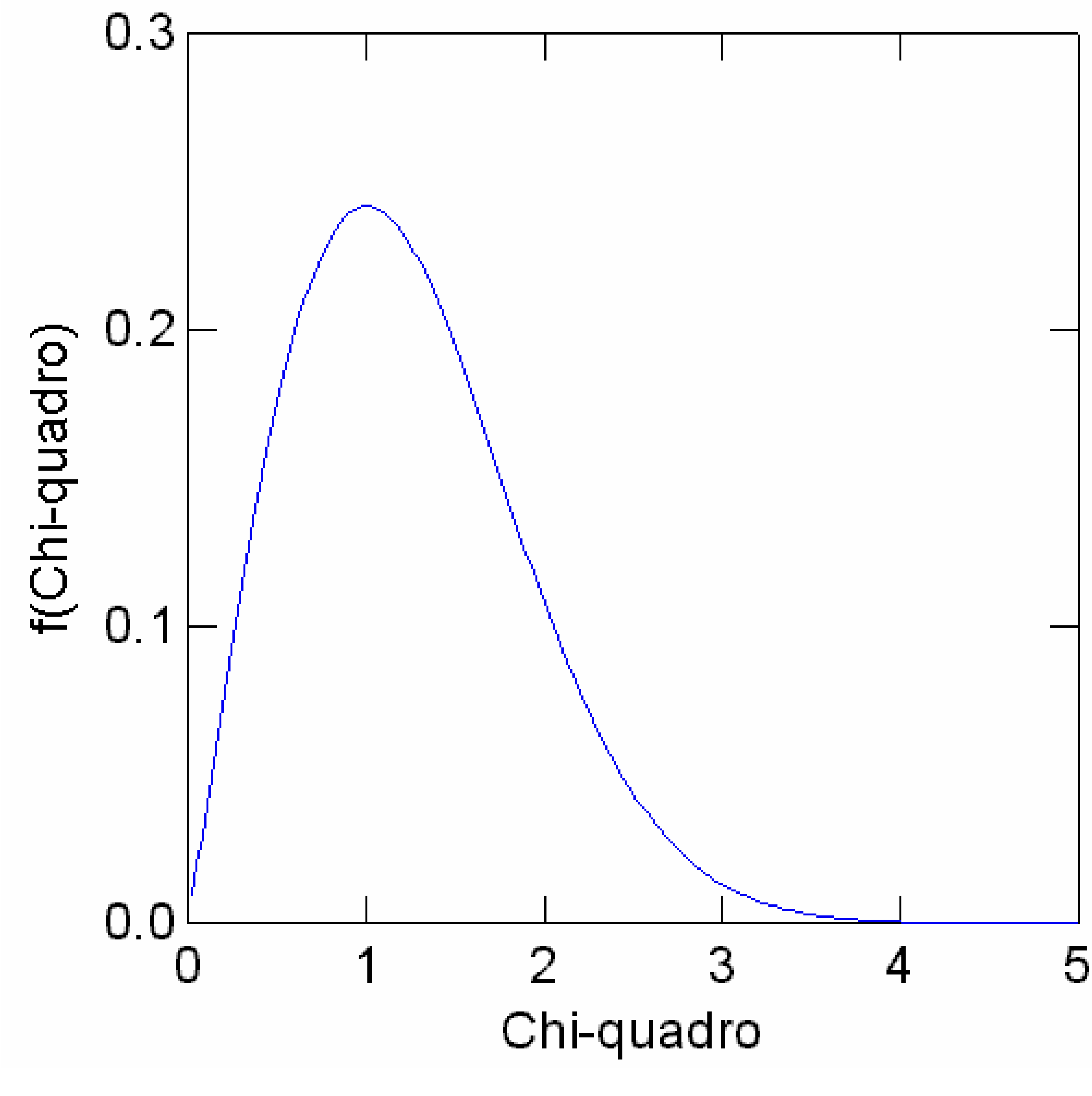
distribuzione chi-quadrato

Quando $\nu = 2$ il grafico della curva è



distribuzione chi-quadrato

Quando $\nu = 3$ il grafico della curva è



proprietà

1. la **distribuzione χ^2** è una distribuzione continua, con valori compresi tra 0 e $+\infty$;

2. al crescere dei gradi di libertà la curva tende ad assumere la forma della normale;

3. dato che la v.c. χ^2 è generata da una somma dei quadrati di n valori indipen-denti di una v.c. normale standardizzata, quando tali valori non sono indipendenti è necessario stabilire le condizioni che li vincolano; sottraendo ad n tali vincoli, si ottiene il numero di **gradi di libertà**, cioè il numero di valori, tra loro indipendenti, della v.c. normale standardizzata.

distribuzione t di Student

Si tratta di una distribuzione continua con densità data dalla relazione:

$$f(t) = C\left(1 + \frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

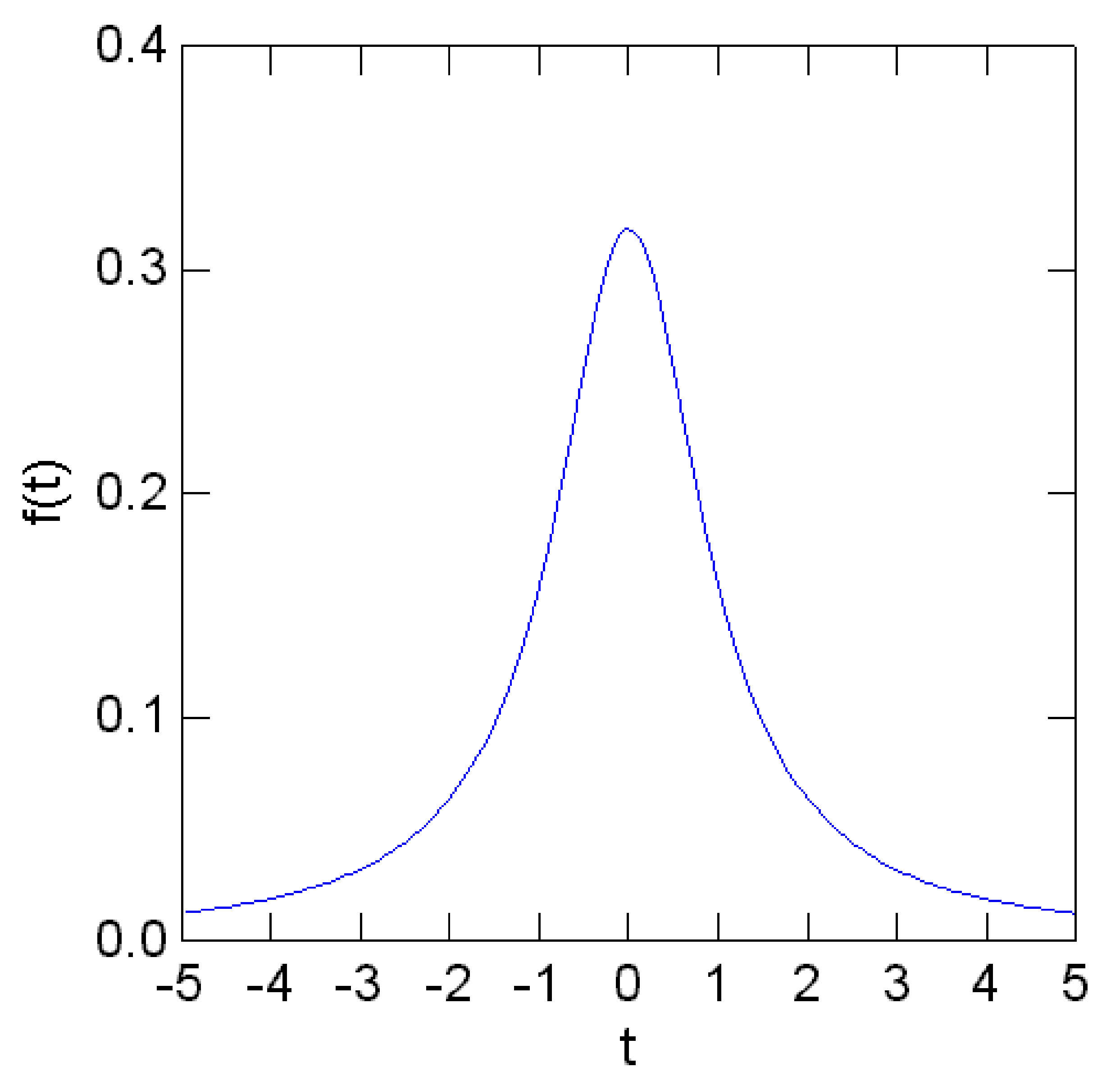
dove:

ν gradi di libertà della distribuzione

C costante tale da assicurare che l'area delimitata sotto la curva sia pari a 1

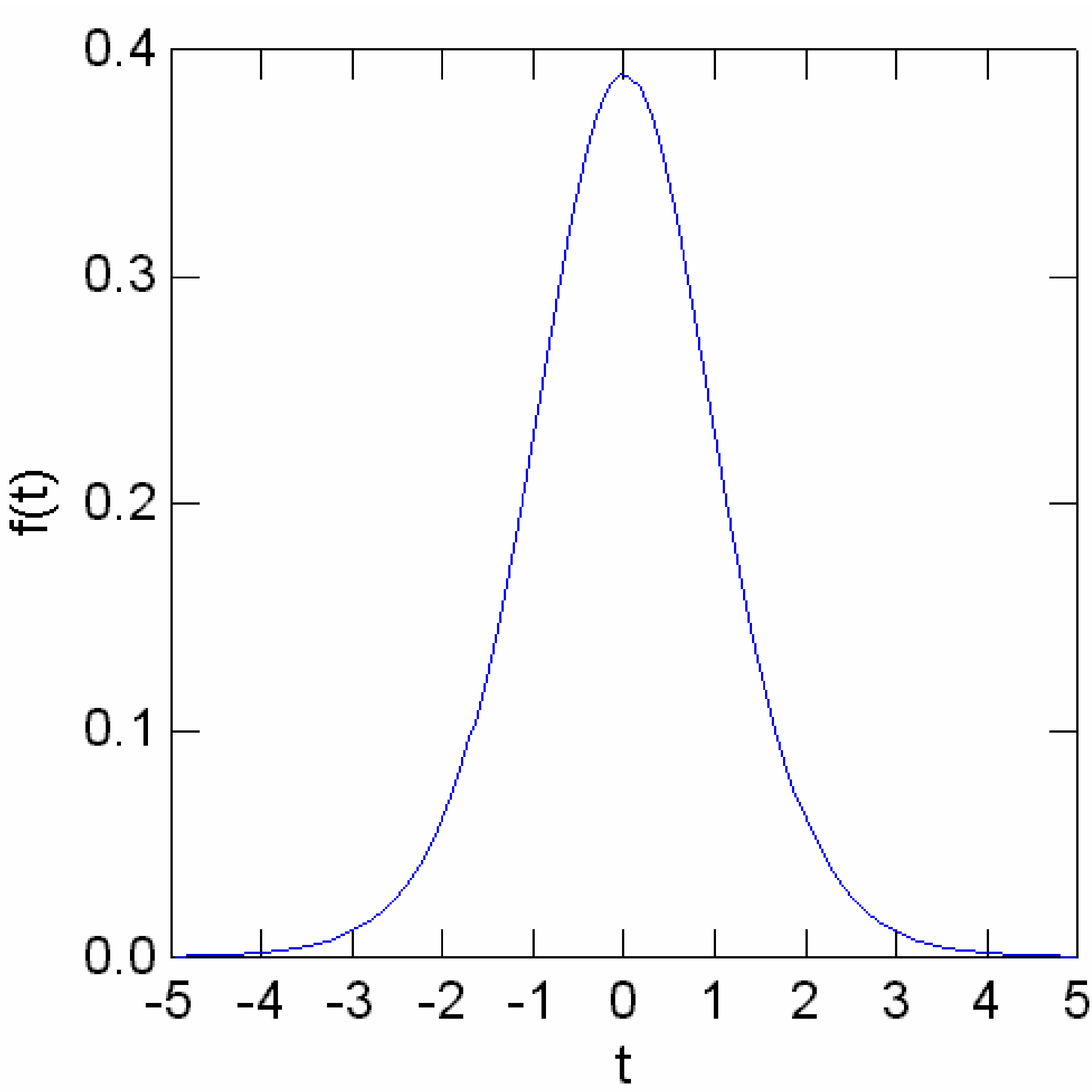
distribuzione t di Student

Quando $\nu = 1$ il grafico della curva è



distribuzione t di Student

Quando $\nu = 10$ il grafico della curva è



proprietà

1. la **distribuzione t** è una distribuzione continua, con valori compresi tra $-\infty$ e $+\infty$;

2. si rivela particolarmente utile nello studio di fenomeni casuali relativi a campioni piccoli ($n < 30$);

3. il valore dei gradi di libertà è dato da $\nu = n - 1$

4. con $\nu \to \infty$ la distribuzione tende alla distribuzione *normale*

distribuzione F di Snedecor

Si tratta di una distribuzione continua con densità data dalla relazione:

$$f\left(F\right)=C_{\left(\nu_1,\nu_2\right)}\frac{F^{\left(\frac{\nu_1}{2}\right)-1}}{\left(1+\frac{\nu_1}{\nu_2}F\right)^{\frac{\nu_1+\nu_2}{2}}}$$

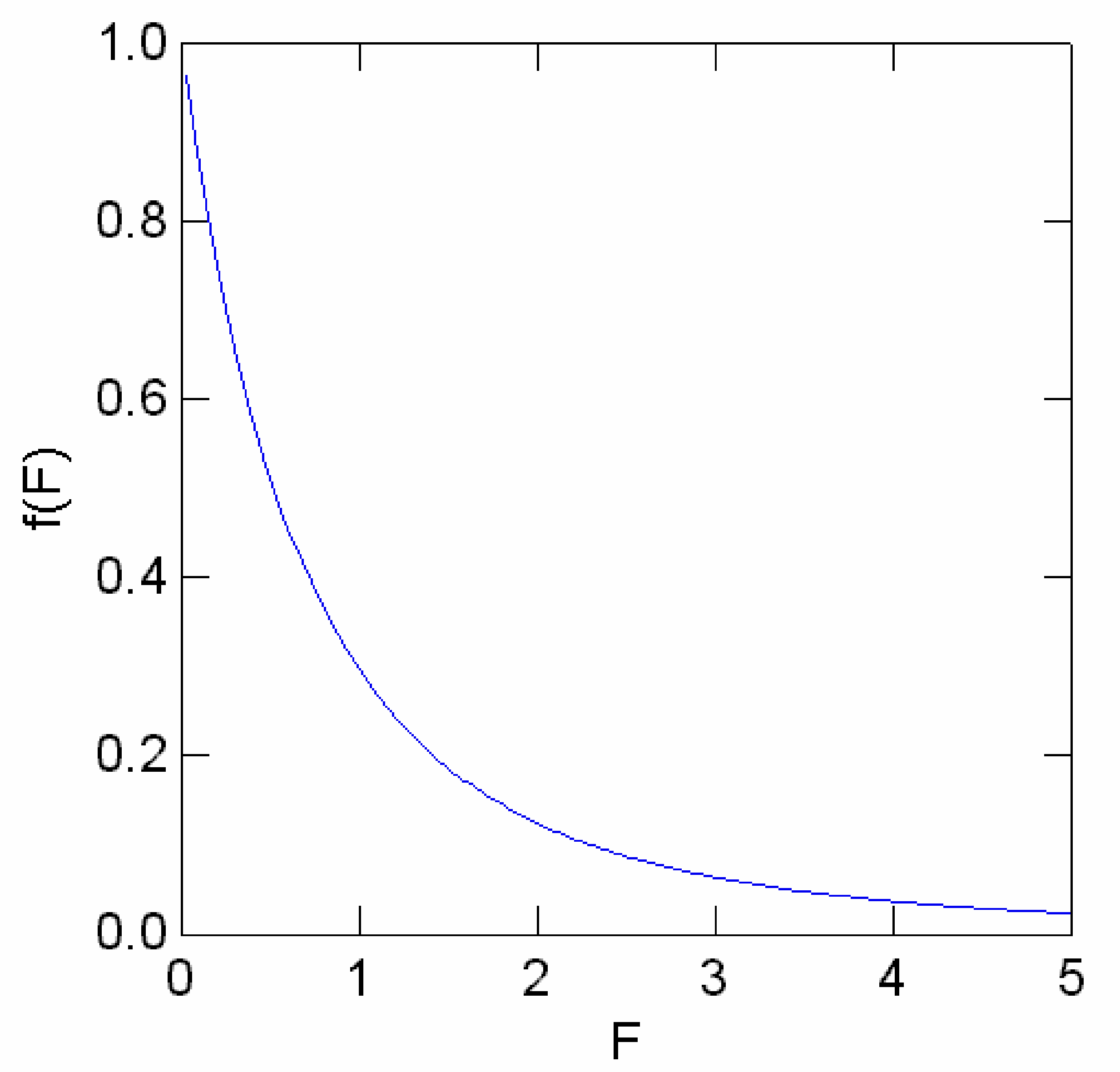
dove:

ν_1,ν_2 gradi di libertà della distribuzione

$C_{\left(\nu_1,\nu_2\right)}$ costante tale da assicurare che l'area delimitata sotto la curva sia pari a 1

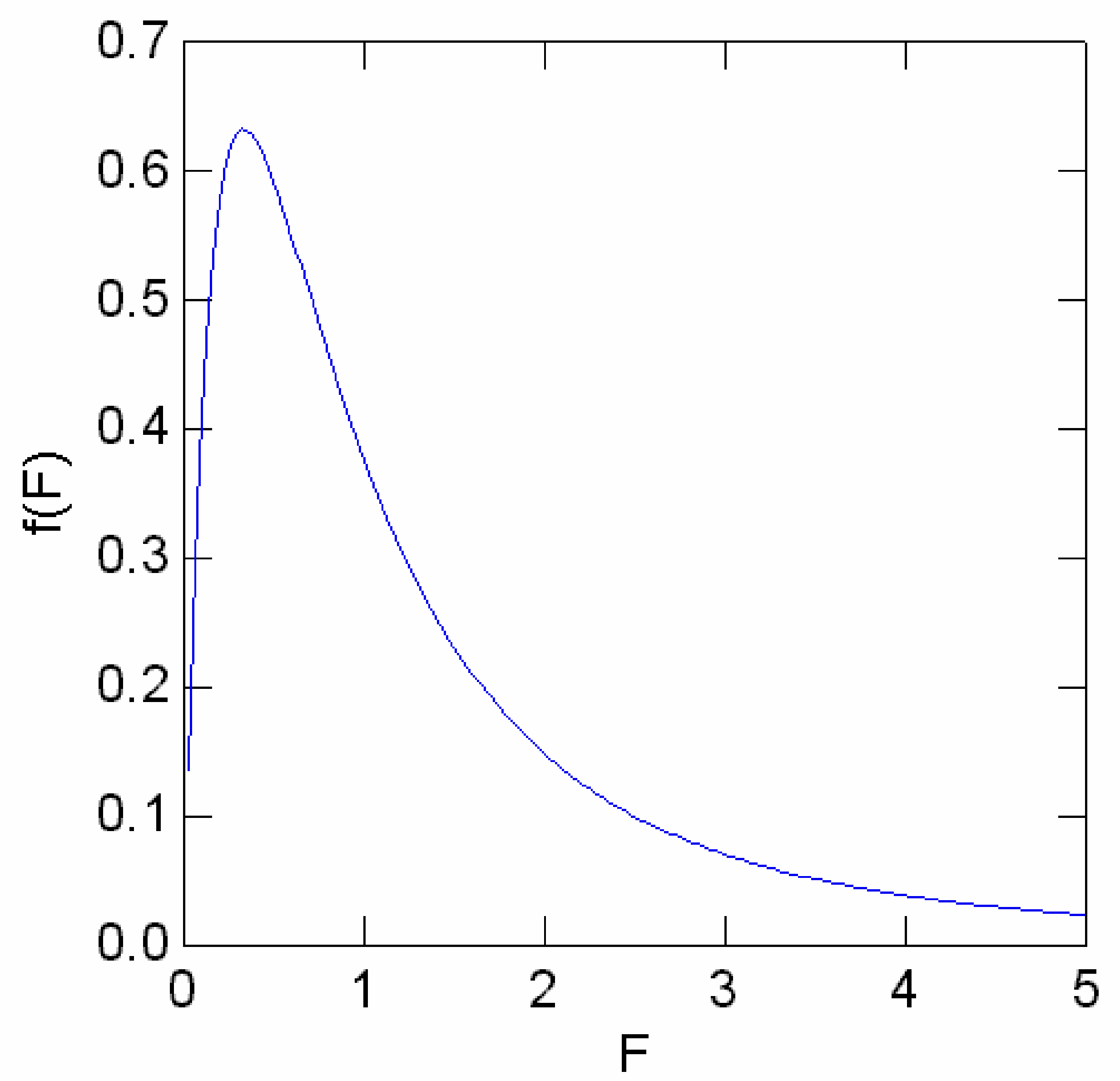
<u>*distribuzione F di Snedecor*</u>

Quando $\nu_1 = 2$ e $\nu_2 = 4$ il grafico della curva è



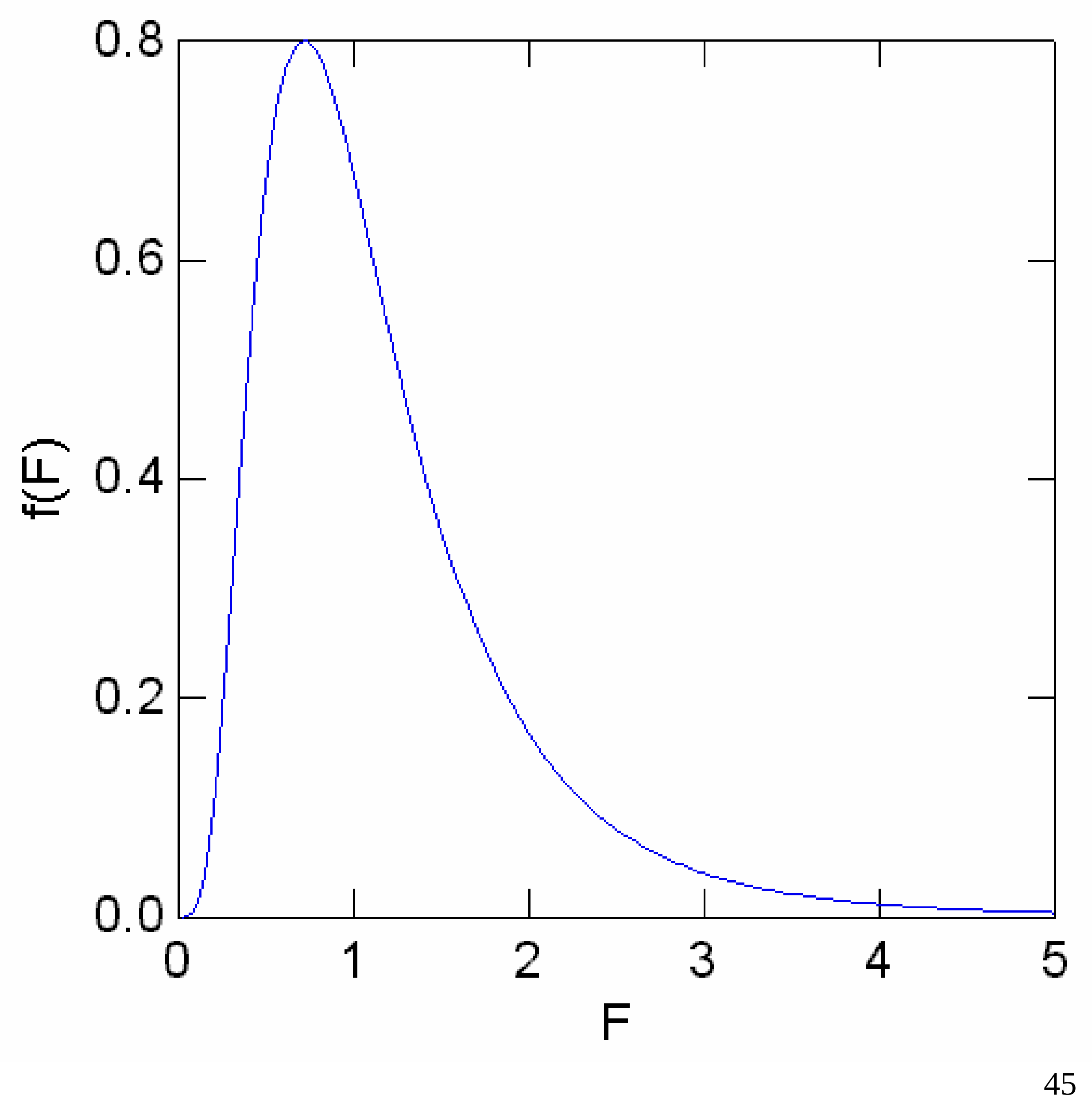
distribuzione F di Snedecor

Quando $\nu_1 = 4$ e $\nu_2 = 4$ il grafico della curva è



distribuzione F di Snedecor

Quando $\nu_1 = 12$ e $\nu_2 = 12$ il grafico della curva è



proprietà

1. la **distribuzione F** è una distribuzione continua, con valori compresi tra 0 e $+\infty$;

2. si dimostra che il rapporto tra due varianze campionarie si distribuisce con questa forma;

3. con $\nu_1 = 1$ e $\nu_2 \to \infty$ la distribuzione tende alla normale standardizzata

4. con $\nu_1 = 1$ e $\nu_2 = \nu$ la distribuzione è uguale alla distribuzione *t di Student*

5. con $\nu_1 = \nu$ e $\nu_2 \to \infty$ la distribuzione tende a quella del χ^2