# La funzione di distribuzione Gaussiana normale

Nicola Morganti

25 aprile 2004

## Indice



## 1 Proprietà fondamentali

L'utilizzo della funzione di distribuzione Gaussiana normale, è quasi d'obbligo nel descrivere l'andamento degli errori di molte misure. Infatti anche se gli errori individuali non seguono questa distribuzione, le medie di gruppi di queste misure hanno una distribuzione gaussiana per un numero elevato di gruppi. Abbiamo principalmente due modi per affrontare il problema:

1. Un risultato che può essere ottenuto matematicamente da considerazioni elementari;

2. Utilizzare una formula empirica che ben approssima gli errori random.

Consideriamo per ora questa distribuzione come un fatto assodato e analizziamo le conseguenze. Assumiamo cioè come nota una distribuzione di probabilità avente il seguente andamento:

$$f(x) = Ae^{-h^2(x-m)^2}$$

dove $A$ è il valore massimo della funzione, $m$ il valore di $x$ per cui la funzione è massima e $h$ un parametro associato alla larghezza della funzione. Ma che significato ha questa $f(x)$ ? Come già anticipato, dire che rappresenta la probabilità di ottenere il valore $x$ non è corretto: tale probabilità infatti è nulla (funzione continua ed illimitata). È più opportuno discutere della probabilità di ottenere un valore compreso tra $x$ e $x + \Delta x$. E quindi la corretta interpretazione è che per un certo valore $x$ ed intervallo $dx$, $f(x)dx$ è la probabilità di osservare un valore nell'intervallo $(x, x + dx)$. Ricordiamo che la probabilità può essere espressa come:

$$P(a, b) = \int_a^b f(x)dx$$

inoltre la probabilità totale vale sempre 1 quindi:

$$P(-\infty, +\infty) = \int_{\infty}^{\infty} f(x)dx = 1$$

La condizione di normalità impone che:

$$\int_{\infty}^{\infty} A e^{-h^2(x-m)^2} dx = 1$$

la quale è soddisfatta per un solo valore di $A$. Eseguiamo una trasformazione di variabili e poniamo:

$$h(x-m) = z \tag{1}$$

otteniamo quindi:

$$A \int_{\infty}^{\infty} e^{-z^2} dz = h$$

poichè l'integrale assume un valore noto, pari a $\sqrt{\pi}$, otteniamo:

$$A = \frac{h}{\sqrt{\pi}} \quad \text{e quindi} \quad f(x) = \frac{h}{\sqrt{\pi}} e^{-h^2(x-m)^2}$$

Determiniamo quindi alcune caratteristiche come il valor medio:

$$\overline{x} = \int_{\infty}^{\infty} x f(x) dx = \frac{h}{\sqrt{\pi}} \int_{\infty}^{\infty} x e^{-h^2(x-m)^2} dx$$

Utilizzando ancora la trasformazione (1) otteniamo:

$$\overline{x} = \frac{1}{\sqrt{\pi}} \int_{\infty}^{\infty} \left( \frac{z}{h} + m \right) e^{-z^2} dz$$

Il primo termine tende a zero, rimane quindi:

$$\overline{x} = \frac{m}{\sqrt{\pi}} \int_{\infty}^{\infty} e^{-z^2} dz = \frac{m}{\sqrt{\pi}} \sqrt{\pi} = m$$

che è un risultato scontato osservando la curva. Determiniamo da ultimo il valore della varianza:

$$\sigma^2 = \int_{\infty}^{\infty} (x-m)^2 f(x) dx = \int_{\infty}^{\infty} \frac{h}{\sqrt{\pi}} (x-m)^2 e^{-h^2(x-m)^2}$$

Eseguendo la medesima sostituzione di variabile (1), si ottiene:

$$\sigma^2 = \frac{1}{h^2 \sqrt{\pi}} \int_{\infty}^{\infty} z^2 e^{-z^2} dz = \frac{1}{2h^2} \quad \text{e quindi} \quad \sigma = \frac{1}{\sqrt{2}h}$$

Possiamo quindi concludere che la deviazione standard è quindi inversamente proporzionale ad $h$:

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

## 2 Standard Normal Density Function

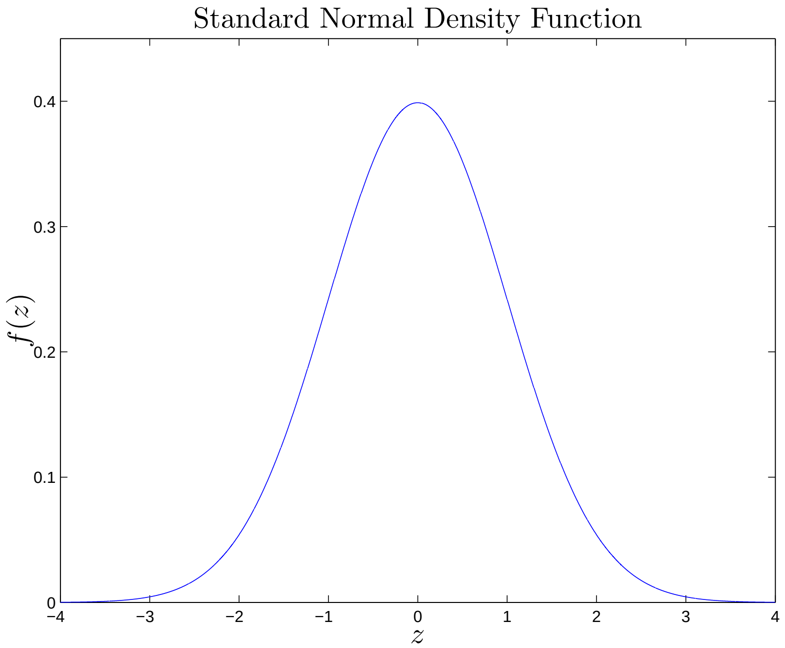
Mediante la sostituzione:

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

otteniamo la *Standard Normal Density Function*, ovvero:

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

che rappresenta la densità di probabilità per una variabile random $z$ con media nulla e deviazione standard unitaria. Questa viene integrata numericamente ed i valori sono disponibili tabulati come nella tabella 1. Possiamo rappresentarla graficamente:




Vediamo appunto come il valore massimo sia in corrispondenza di $z = 0$. In questo modo abbiamo che:

$$P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2) = P\left(\frac{x_1 - \mu}{\sigma} \le \frac{x - \mu}{\sigma} \le \frac{x_2 - \mu}{\sigma}\right)$$

Utilizzando i valori integrati numericamente riportati in tabella 1 è possibile valutare i livelli di confidenza (intervalli di confidenza), ovvero la probabilità che i valori di $z$ cadano negli intervalli seguenti:

| | |
|---|---|
| $\sigma$ | 68.26% |
| $2\sigma$ | 95.44% |
| $3\sigma$ | 99.74% |
| $3.5\sigma$ | 99.96% |

Questi intervalli si utilizzano per classificare i vari strumenti di misura. Di solito nelle specifiche di uno strumento, viene indicato anche l'intervallo di confidenza a cui appartiene. Se questo non è specificato si sottintende l'appartenenza all'intervallo $3\sigma$.

Tabella 1: Area sottesa alla funzione di distribuzione Gaussiana normale da $z = 0$ a $z$

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1 | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2 | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3 | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4 | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5 | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |
| 0.6 | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7 | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8 | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9 | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.3340 | 0.3365 | 0.3389 |
| 1.0 | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1 | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.3790 | 0.3810 | 0.3830 |
| 1.2 | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.3980 | 0.3997 | 0.4015 |
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |
| 1.8 | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.4693 | 0.4699 | 0.4706 |
| 1.9 | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | 0.4750 | 0.4756 | 0.4761 | 0.4767 |
| 2.0 | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.4808 | 0.4812 | 0.4817 |
| 2.1 | 0.4821 | 0.4826 | 0.4830 | 0.4834 | 0.4838 | 0.4842 | 0.4846 | 0.4850 | 0.4854 | 0.4857 |
| 2.2 | 0.4861 | 0.4864 | 0.4868 | 0.4871 | 0.4875 | 0.4878 | 0.4881 | 0.4884 | 0.4887 | 0.4890 |
| 2.3 | 0.4893 | 0.4896 | 0.4898 | 0.4901 | 0.4904 | 0.4906 | 0.4909 | 0.4911 | 0.4913 | 0.4916 |
| 2.4 | 0.4918 | 0.4920 | 0.4922 | 0.4925 | 0.4927 | 0.4929 | 0.4931 | 0.4932 | 0.4934 | 0.4936 |
| 2.5 | 0.4938 | 0.4940 | 0.4941 | 0.4943 | 0.4945 | 0.4946 | 0.4948 | 0.4949 | 0.4951 | 0.4952 |
| 2.6 | 0.4953 | 0.4955 | 0.4956 | 0.4957 | 0.4959 | 0.4960 | 0.4961 | 0.4962 | 0.4963 | 0.4964 |
| 2.7 | 0.4965 | 0.4966 | 0.4967 | 0.4968 | 0.4969 | 0.4970 | 0.4971 | 0.4972 | 0.4973 | 0.4974 |
| 2.8 | 0.4974 | 0.4975 | 0.4976 | 0.4977 | 0.4977 | 0.4978 | 0.4979 | 0.4979 | 0.4980 | 0.4981 |
| 2.9 | 0.4981 | 0.4982 | 0.4982 | 0.4983 | 0.4984 | 0.4984 | 0.4985 | 0.4985 | 0.4986 | 0.4986 |
| 3.0 | 0.4987 | 0.4987 | 0.4987 | 0.4988 | 0.4988 | 0.4989 | 0.4989 | 0.4989 | 0.4990 | 0.4990 |
| 3.1 | 0.4990 | 0.4991 | 0.4991 | 0.4991 | 0.4992 | 0.4992 | 0.4992 | 0.4992 | 0.4993 | 0.4993 |
| 3.2 | 0.4993 | 0.4993 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4995 | 0.4995 | 0.4995 |
| 3.3 | 0.4995 | 0.4995 | 0.4995 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4997 |
| 3.4 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4998 |
| 3.5 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 |
| 3.6 | 0.4998 | 0.4998 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 |
| 3.7 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 |
| 3.8 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 |
| 3.9 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 |
| 4.0 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 | 0.5000 |

## 3 Esempio applicativo

Il risultato di un test presenta una distribuzione gaussiana, con media pari a 10 e deviazione standard pari ad 1. Trovare la probabilità che una singola misura sia rispettivamente:

1. compresa tra 9 e 12;
2. compresa tra 8 e 9.55;
3. $\leq 9$;
4. $> 12$

**Caso 1** Calcoliamo gli estremi di integrazione $z_1$ e $z_2$:

$$z_1 = \frac{9-10}{1} = -1 \quad z_2 = \frac{12-10}{1} = 2$$

Sfruttiamo il fatto che la funzione è simmetrica e spezziamo l'integrale in due intervalli:

$$[-1:0] \text{ e } [0:2]$$

abbiamo quindi, utilizzando i valore della tabella 1:

$$P(-1,0) = P(0,1) = 0.3413$$

$$P(0,2) = 0.4772$$

la probabilità che la misura sia compresa tra 9 e 12 sarà quindi:

$$P(-1,0) + P(0,2) = 0.3413 + 0.4772 = 0.8185 = 81.85\%$$

**Caso 2** Procedendo come nel caso precedente otteniamo

$$z_1 = \frac{8-10}{1} = -2 \; z_2 = \frac{9.55-10}{1} = -0.45$$

Non abbiamo bisogno di dividere l'intervallo in due parti in quanto sia $z_1$ che $z_2$ hanno lo stesso segno. Utilizzando sempre la stessa tabella calcoliamo la probabilità richiesta come:

$$P(-2,-0.45) = P(0,2) - P(0,0.45) = 0.4772 - 0.1736 = 0.3036 = 30.36\%$$

**Caso 3** Questa volta gli estremi sono $z_1 = -\infty$ e $z_2 = -1$, mentre la probabilità sarà:

$$P(-\infty,-1) = P(0,\infty) - P(1,0) = 0.5 - 0.3413 = 0.1587 = 15.87\%$$

**Caso 4** Allo stesso modo del caso precedente abbiamo:

$$z_1 = \frac{12-10}{1} = 2 \quad z_2 = \infty$$

e quindi la probabilità sarà:

$$P(2,\infty) = P(0,\infty) - P(2,0) = 0.5 - 0.4772 = 0.0228 = 2.28\%$$