Normal distribution

What does it mean that x goes from $-\infty$ to $-\infty$?

Gauss was indeed 'embarassed' about that!

"The function just found cannot, it is true, express rigorously the probabilities of the errors: for since the possible errors are in all cases confined within certain limits, the probability of errors exceeding those limits ought always be zero, while our formula always gives some value. However, this defect, which every analytical function must, from its nature, labor under, is of no importance in practice, because the value of function decreases so rapidly, when $h\,\Delta$ [$\propto (x_i - \mu)/\sigma$', in modern notation] *has acquired a considerable magnitude, that it can safely be considered as vanishing. Besides, the nature of the subject never admits of assigning with absolute rigor the limits of error."*

→ Remember Laplace's **good sense**!

Gauss' derivation of the Gaussian

Gauss' problem (expressed in modern terms):

What is the more general form of the likelihood such that the maximum of the posterior of μ is equal to the arithmetic average of the observed values (and the function has some 'good' mathematical properties)?

- 'likelihood' $\leftrightarrow$ 'error function'

$$f(x_i \,|\, \mu) = \varphi(x_i - \mu)$$

- x_i independent and affected by errors of the same kind:

$$f(\boldsymbol{x} \,|\, \mu) = \varphi(x_1 - \mu) \cdot \varphi(x_2 - \mu) \cdot \;\cdots\; \cdot \varphi(x_n - \mu)$$

- Then Gauss makes a reasoning that we would call *Bayesian* (without reference, so obvious he considered the reasoning)

$$f(\mu \,|\, \boldsymbol{x}) \propto f(\boldsymbol{x} \,|\, \mu) \cdot f_0(\mu)$$

Gauss' derivation of the Gaussian (cont.d)

Note: the concept of prior (*"ante eventum cognitum"*) was very clear and natural to Gauss, opposed to the concept of posterior (*"post eventum cognitum"*).

Then he makes two assumptions (besides that φ is continuous and infinite times derivable):

1. All values of μ are considered *a priori* (*"ante illa observationes"*) equally likely (*"... aeque probabilia fuisse"*).
2. The maximum *a posteriori* (*"post illas observationes"*) is given by $\mu = \overline{x}$, arithmetic average of the n observed values.

1. $\rightarrow f(\mu\,|\,\boldsymbol{x}) \propto f(\boldsymbol{x}\,|\,\mu) = \prod_i \varphi(x_i - \mu)$

2. $\rightarrow \frac{1}{z_i}\frac{\varphi'(z_i)}{\varphi(z_i)} = k \;\rightarrow\; \varphi(z_i) \propto e^{\frac{k}{2}\,z_i^2} = e^{-h^2\,z_i^2}$ (*he was Gauss. . .*)

with $z_i = x_i - \overline{x} = x_i - \mu$ and redefining $k/2 = -h^2$ in order to make evident that the function had a maximum at $z_i = 0$.

Gauss' derivation of the Gaussian (cont.d)

The searched error function was of the kind, using the original notation, with $\Delta = z$ [note '$\varphi\,\Delta$' for $\varphi(\Delta)$],

$$\varphi\,\Delta \propto e^{-hh\,\Delta\Delta}$$

But the normalization was still missing $\longrightarrow$ *Laplace ex machina*:

"since, by the elegant theorem first discovered by ill. LAPLACE, integrale

$$\int e^{-hh\,\Delta\Delta}\mathrm{d}\Delta\,,$$

a $\Delta = -\infty$ usque ad $\Delta = +\infty$, fiat $= \frac{\sqrt{\pi}}{h}$, (denotando per π semicircumferentiam circuli cuius radius 1),

functio nostra fiet

$$\varphi\,\Delta = \frac{h}{\sqrt{\pi}}e^{-hh\,\Delta\Delta}\text{ "} \longrightarrow \varphi(\Delta) = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{\Delta^2}{2\,\sigma^2}}$$

$(h^2 = \frac{1}{2\sigma^2})$