Departamento de Física Experimental

Distribuição de Probabilidade

Distribuição Gaussiana

(*Cosmos*)

01-02 abril de 2014

Paulo R. Pascholati

Cosmos – São Paulo

Prólogo

Esta apresentação é parte das distribuições de probabilidade de interesse em Física: Distribuição Binomial, Distribuição de Poisson e Distribuição de Gauss (Gauβ). Enfatizando o que já foi adiantado na apresentação de distribuição binomial:

- a distribuição binomial é geralmente aplicada a experimentos onde há um número pequeno de possíveis eventos;
- a distribuição de Poisson é apropriada para descrever experimentos cujos eventos são contagens onde os dados representam um certo número de eventos observados por unidade de intervalo; e
- A distribuição de Gauss ou normal descreve as observações aleatórias de uma vasta gama de experimentos.

Distribuição Gaussiana

Expressão

A distribuição gaussiana é uma aproximação para a distribuição binomial no caso limite em que o número de diferentes possibilidades de resultados N se torna infinitamente grande enquanto que a probabilidade de sucesso de cada uma deles é finitamente grande de modo que $Np \gg 1$[1].

A distribuição gaussiana é também um caso limite da distribuição de Poisson em que μ se torna grande.

A expressão da distribuição gaussiana é:

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

onde os parâmetros μ e σ são, respectivamente, a média e o desvio padrão da distribuição.

[1] Abordagens diferentes de dedução da distribuição gaussiana podem ser encontradas em: http://scipp.ucsc.edu/∼haber/ph116C/NormalApprox.pdf e O.A.M. Helene e V.R. Vanin, "Tratamento Estatísticos de Dados em Física Experimental", Editôra Edgar Blücher, 1981, páginas 34 a 38.

Distribuição Gaussiana

Propriedades

- É uma probabilidade de distribuição contínua na variável x.
- É simétrica em relação à média. Mediana, moda e média são iguais.
- Não é adimensional.
- A probabilidade $P(x)$ de se encontrar x em um intervalo Δx é $g(x)\Delta x$. Adicionalmente
- É unimodal, a derivada primeira é positiva para $x < \mu$ e negativa para $x > \mu$.
- Tem dois pontos de inflexão(onde a segunda derivada é zero e muda de sinal) em $x - \mu$ e $x + \mu$.
- É infinitamente diferenciável.

Distribuição Gaussiana
Propriedades

O máximo do valor da gaussiana acontece quando $x = \mu$

$$g(\mu) = \frac{1}{\sigma\sqrt{2\pi}} = \frac{0,3989}{\sigma} \tag{2}$$

outros valores

$$g(\mu \pm \sigma) = e^{-1/2} g(\mu) = 0,606 g(\mu) \tag{3}$$

$$g(\mu \pm 2\sigma) = e^{-2} g(\mu) = 0,153 g(\mu) \tag{4}$$

$$g(\mu \pm 3\sigma) = e^{-9/2} g(\mu) = 0,011 g(\mu) \tag{5}$$

Distribuição Gaussiana
Normalização

A normalização da distribuição gaussiana é obtida por:

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx = 1 \tag{6}$$

Para calcular a integral faz-se a mudança de variável $\frac{x-\mu}{\sqrt{2}\sigma} = x'$ segue que $x = \sqrt{2}\sigma \; x'$, assim

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x'^2} \, dx' \tag{7}$$

A integral da equação 7 sem o fator multiplicativo, chamada equação de Gauss, pode ser resolvida elevando-a ao quadrado

$$\left(\int_{-\infty}^{+\infty} e^{-x'^2} \, dx' \right)^2 = \int_{-\infty}^{+\infty} e^{-x'^2} \, dx' \cdot \int_{-\infty}^{+\infty} e^{-y'^2} \, dy' \tag{8}$$

Distribuição Gaussiana

Normalização

ou

$$\left(\int_{-\infty}^{+\infty} e^{-x'^2}\, dx'\right)^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x'^2+y'^2)}\, dx'dy' \qquad (8)$$

Fazendo então a transformação de coordenadas cartesianas para coordenadas polares $r = x'^2 + y'^2$ com o elemento de volume $rd\theta dr = dx'dy'$ o segundo termo da equação 8

$$\int_0^{+\infty}\int_0^{2\pi} e^{-r^2}\, rd\theta dr = 2\pi \int_0^{+\infty} e^{-r^2}\, rdr \qquad (9)$$

ou

$$\pi \int_0^{+\infty} e^{-r^2}\, 2rdr = -\pi e^{-z}\Big|_0^{+\infty} = -\pi(e^{-\infty} - e^{-0}) = \pi \qquad (10)$$

Distribuição Gaussiana
Normalização

Assim a equação 7 torna

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x'^2}\, dx' = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1 \tag{7}$$

ou

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\, dx = 1 \tag{6}$$

Distribuição Gaussiana
Média

$$\text{média} = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ dx \tag{11}$$

Colocando em lugar de x $(x - \mu + \mu)$ e rearrando os termos

$$\text{média} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x - \mu + \mu) e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ dx \tag{12}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{+\infty} (x-\mu) e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{-\infty}^{+\infty} \mu e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \right) \tag{13}$$

Distribuição Gaussiana
Média

A primeira integral da equação 13 é nula porque é integral de uma função ímpar em um intervalo simétrico em relação ao valor $x - \mu = 0$.
A segunda integral, usando a normalização da distribuição pela normalização, equação 6, resulta em:

$$\mu \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx = \mu \cdot 1 = \mu \qquad (14)$$

Assim

$$\text{média} = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx = \mu \qquad (11')$$

Distribuição Gaussiana
Média

Outra forma de obter o resultado da primeira integral da equação 13. Fazendo a mudança de variável
$\frac{1}{\sqrt{2}}\left(\frac{x-\mu}{\sigma}\right) = x' \quad \rightarrow \quad \frac{1}{\sigma\sqrt{2}}dx = dx' \quad \rightarrow \quad dx = \sigma\sqrt{2}dx'$

$$I_1 = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-\mu)e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx \tag{15}$$

$$I_1 = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma\sqrt{2}x'e^{-x'^2}\ \sigma\sqrt{2}dx' = \frac{\sigma\sqrt{2}}{\sqrt{\pi}}\int_{-\infty}^{+\infty}x'e^{-x'^2}\ dx' \tag{16}$$

Agora fazendo a seguinte transformação de variável na equação 16
$x'^2 = z \quad \rightarrow \quad 2x'dx' = dz$

$$\int x'e^{-x'^2}\ dx' = \frac{1}{2}\int e^{-z}\ dz = -\frac{1}{2}e^{-z} = -\frac{1}{2}e^{-x'^2} \tag{17}$$

Distribuição Gaussiana
Média

Finalmente, a primeira integral da equação 13 resulta :

$$I_1 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-\mu) e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \tag{15}$$

$$= -\frac{\sigma\sqrt{2}}{\sqrt{\pi}} \frac{1}{2} e^{-x'^2}\Big|_{-\infty}^{+\infty} = -\frac{\sigma}{\sqrt{2\pi}} e^{-x'^2}\Big|_{-\infty}^{+\infty} = 0 \tag{18}$$

Distribuição Gaussiana

Variança

$$\text{variança} = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \tag{19}$$

Fazendo uma mudança de variável como na equação 6 obtem-se

$$\text{variança} = \int_{-\infty}^{+\infty} 2\sigma^2 x'^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-x'^2} \, \sqrt{2}\sigma dx' \tag{20}$$

$$= \int_{-\infty}^{+\infty} 2\sigma^2 x'^2 \frac{1}{\sqrt{\pi}} e^{-x'^2} \, dx' = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x'^2 e^{-x'^2} \, dx' \tag{21}$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x' \left(x' e^{-x'^2} \, dx' \right) \tag{22}$$

Distribuição Gaussiana
Variança

Resolvendo a segunda integral da equação 22 por partes[2], $u = x'$ e $dv = x'e^{-x'^2}$ $(v = -\frac{1}{2}e^{-x'^2})$, se obtem:

$$\frac{2\sigma^2}{\sqrt{\pi}}\left[\left(-\frac{1}{2}x'e^{-x'^2}\right)\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty}\left(-\frac{1}{2}e^{-x'^2}\,dx'\right)\right] \tag{23}$$

A primeira integral da equação 23 é nula pela regra de L'Hôpital e a segunda integral é $\frac{\sqrt{\pi}}{2}$ (integral de Gauss, equação 7).
Assim tem-se que:

$$\text{variança} = \frac{2\sigma^2}{\sqrt{\pi}}\frac{\sqrt{\pi}}{2} = \sigma^2 \tag{24}$$

e o desvio padrão

$$\text{desvio padrão} = \sigma \tag{25}$$

[2] $\int udv = uv - \int vdu$

Distribuição Gaussiana

Largura à Meia Altura – FWHM Full Width at Half Maximum

$$g(x_{FWHM}) = \frac{1}{2} g(\mu) \tag{26}$$

$$\frac{g_\mu(x_{FWHM})}{g_\mu(x)} = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_{FWHM}-\mu}{\sigma}\right)^2}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu-\mu}{\sigma}\right)^2}} \tag{27}$$

$$\frac{1}{2} = \frac{e^{-\frac{1}{2}\left(\frac{x_{FWHM}-\mu}{\sigma}\right)^2}}{e^{-\frac{1}{2}\left(\frac{0}{\sigma}\right)^2}} \tag{28}$$

Distribuição Gaussiana

Largura à Meia Altura – FWHM Full Width at Half Maximum

$$\frac{1}{2} = e^{-\frac{1}{2}\left(\frac{x_{FWHM}-\mu}{\sigma}\right)^2} \tag{29}$$

Fazendo o logaritmo natural em ambos os lados da igualdade

$$-ln(2) = -\frac{1}{2}\left(\frac{x_{FWHM}-\mu}{\sigma}\right)^2 \tag{30}$$

$$2ln(2) = \left(\frac{x_{FWHM}-\mu}{\sigma}\right)^2 \tag{31}$$

$$x_{FWHM} = \mu \pm \sigma\sqrt{2ln(2)} \tag{32}$$

$$\Gamma = (x_{FWHM}+\mu) - (x_{FWHM}-\mu) = 2\sigma\sqrt{2ln(2)} = 2,354\sigma \tag{33}$$

Distribuição Gaussiana

Largura à Meia Altura – FWHM Full Width at Half Maximum

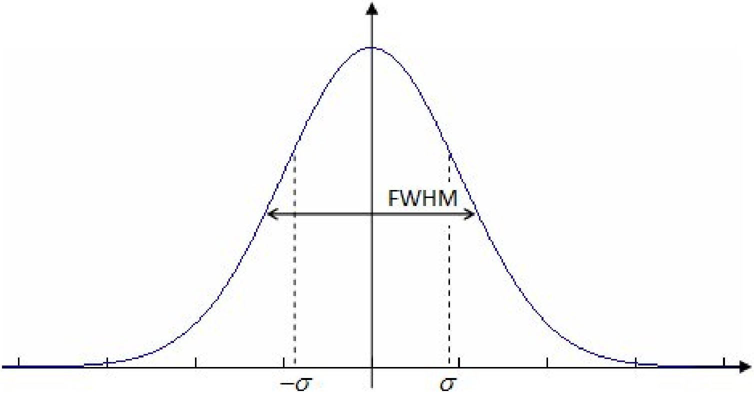


Figura : Distribuição gaussiana com largura-a-meia-altura e sigma.

Distribuição Gaussiana

Níveis de Confiança e Intervalos de Confiança

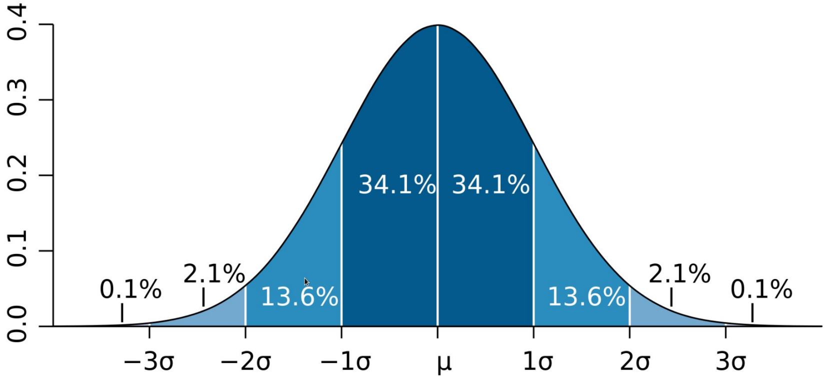


Figura : Distribuição gaussiana com intervalos de confiança em função de sigma.

Distribuição Gaussiana

Distribuição Gaussiana Padrão

A distribuição gaussiana padrão, que é de muito uso prático, é obtida fazendo a mudança de variável de $z = (x - \mu)/\sigma$ e $\sigma = 1$

$$g_{\mu}(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{z^2}{2}\right)} \tag{34}$$

Distribuição Gaussiana
Distribuição Gaussiana Padrão

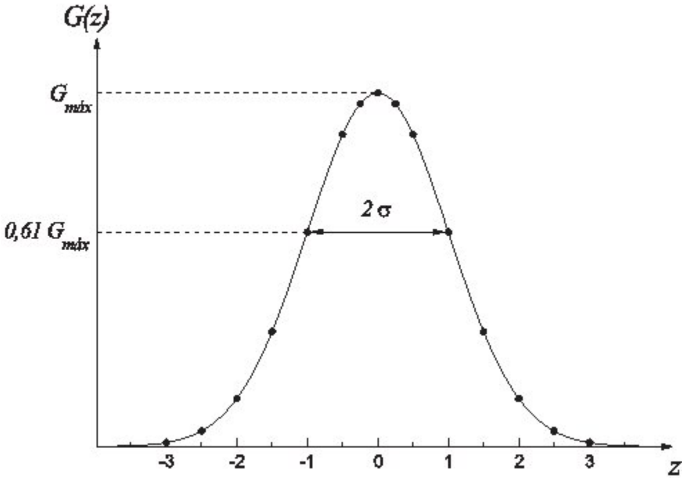


Figura : Gaussiana padrão, $\mu = 0$ e $\sigma = 1$.

Distribuição Gaussiana

Distribuição Gaussiana Padrão

Tabela : Valores de z e $g(z)$ para a distribuição normal padrão.

z	$g(z)$
0	0,3989
0,25	0,3867
0,5	0,3521
1,0	0,2420
1,5	0,1295
2,0	0,0540
2,5	0,0175
3,0	0,0044

Distribuição Gaussiana
Distribuição Gaussiana Padrão

Para colocar a curva gaussiana sobre um histograma basta calcular x para cada valor de z da tabela, onde

$$z = \frac{(x - \bar{x})}{s}, \tag{35}$$

com $x = \bar{x} \pm sz$, e valor $\Delta P(x)$ neste ponto x é dado pela equação

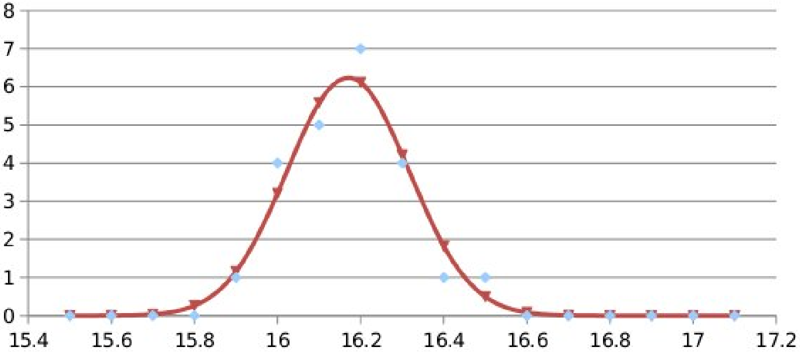
$$\Delta P(x) = \frac{\Delta x}{s} g(z) \tag{36}$$

O valor assim obtido é a estimativa gaussiana para a probabilidade, que pode ser comparado à freqüência relativa obtida experimentalmente, no respectivo ponto x.
Caso não seja usada a frequência relativa a Equação 36 se torna

$$\Delta P(x) = \frac{\Delta x}{s} g(z) N \tag{37}$$

Distribuição Gaussiana
Distribuição Gaussiana Padrão

Figura : Intervalo de tempo de dez períodos de um pêndulo simples com $\overline{10\,T} = 16,22$ e $s = 0,15$.



18,30	18,29	18,25	18,11	18,10	18,07	18,06	18,47	18,09	18,17
18,34	18,10	18,28	18,54	17,95	18,38	18,18	17,82	18,24	17,92
18,09	18,26	18,13	18,29	18,24	18,32	18,13	18,29	18,47	18,14
18,30	18,04	17,89	17,54	17,72	18,11	17,88	18,25	18,34	17,80
18,35	18,31	18,29	18,18	18,19	17,96	18,26	18,23	18,20	18,57
18,30	18,11	18,01	18,50	18,45	18,39	18,29	18,20	18,03	18,08
17,89	18,03	18,46	18,15	18,32	18,20	18,14	18,65	18,23	18,37
18,47	17,73	18,15	18,08	18,53	18,29	18,21	17,97	18,34	18,37
17,99	18,32	18,41	18,32	18,22	18,30	18,20	18,26	18,06	18,08
18,65	18,60	18,60	18,54	18,10	18,19	18,15	18,32	18,27	17,88
18,07	18,10	18,15	18,38	17,82	18,38	18,03	17,79	18,28	17,88
18,17	18,54	18,09	18,24	18,80	18,27	18,06	18,32	18,51	18,4
18,46	18,18	18,46	18,33	18,28	18,23	18,02	18,54	18,56	18,19
17,92	18,04	18,29	18,33	18,45	17,99	18,37	18,16	17,81	18,45
18,32	18,24	18,23	18,18	18,34	18,23	18,21	18,16	18,46	18,12
18,38	18,13	18,25	18,31	18,13	18,42	18,29	17,90	18,08	17,58
18,34	18,37	18,06	18,22	18,15	18,13	18,07	18,29	17,93	17,98
18,06	18,03	18,45	18,37	17,66	18,31	17,74	18,45	18,09	18,40
18,32	18,25	18,12	18,36	18,00	18,36	18,23	17,94	18,11	18,27
18,00	18,28	18,46	17,91	18,13	18,95	18,36	17,92	18,29	18,63

Distribuição Gaussiana
Distribuição Gaussiana Padrão

$$\overline{10T} = 18,2108$$
$$s_{10T} = 0,2172$$
$$s_{\overline{10T}} = 0,0154$$

$$\overline{10T} = 18,211$$
$$s_{10T} = 0,2172$$
$$s_{\overline{10T}} = 0,015$$

Distribuição Gaussiana
Distribuição Gaussiana Padrão

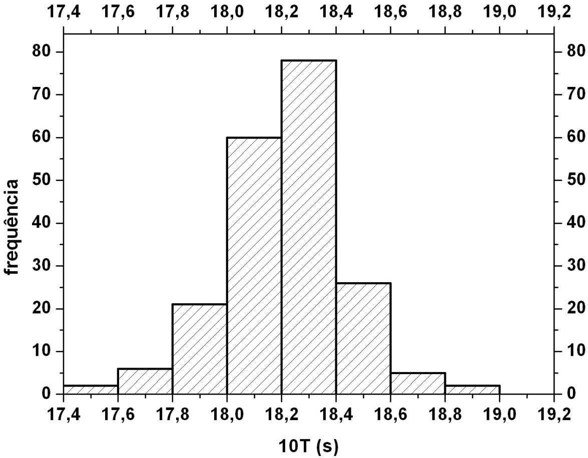


Figura : 100 medições de 10 oscilações de um pêndulo. Histograma com canal de 0,20s

Distribuição Gaussiana
Distribuição Gaussiana Padrão

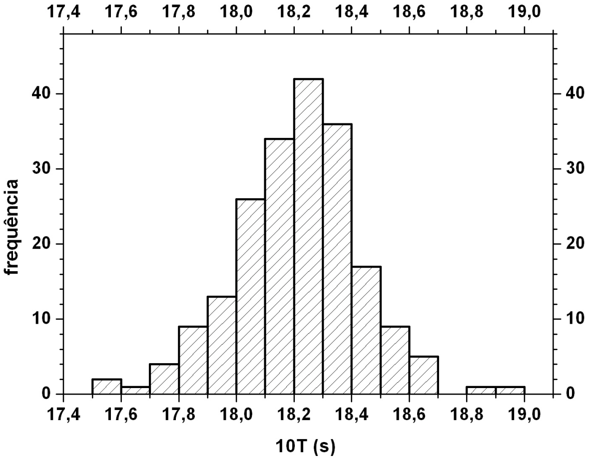


Figura : 100 medições de 10 oscilações de um pêndulo. Histograma com canal de 0,10s.

Distribuição Gaussiana
Distribuição Gaussiana Padrão

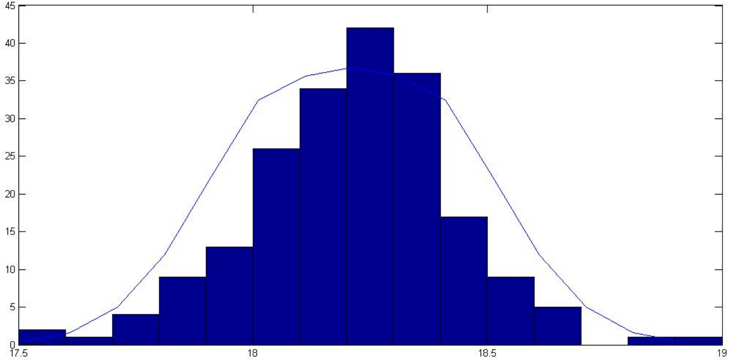


Figura : Cem medições de 10 oscilações de um pêndulo. Histograma com canal de 0,10s e função gaussiana sobreposta. Má sobreposiçao, área do histograma muito diferente da área sob a guassiana.

Distribuição Gaussiana
Distribuição Gaussiana Padrão

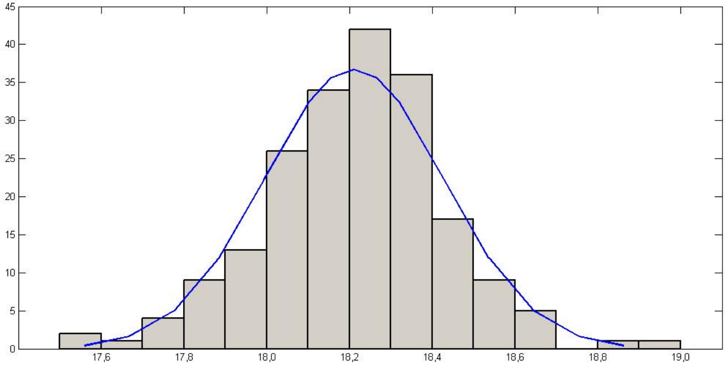


Figura : Cem medições de 10 oscilações de um pêndulo. Histograma com canal de 0,10s e função gaussiana sobreposta. Boa sobreposiçao, área do histograma muito próxima da área sob a gaussiana.