

# Distribuições

## **Contínuas**

- Normal
- Gama
- Valores Extremos
- Exponencial

## **Discretas**

- Binomial
- Poisson
- Geométrica

# Funções

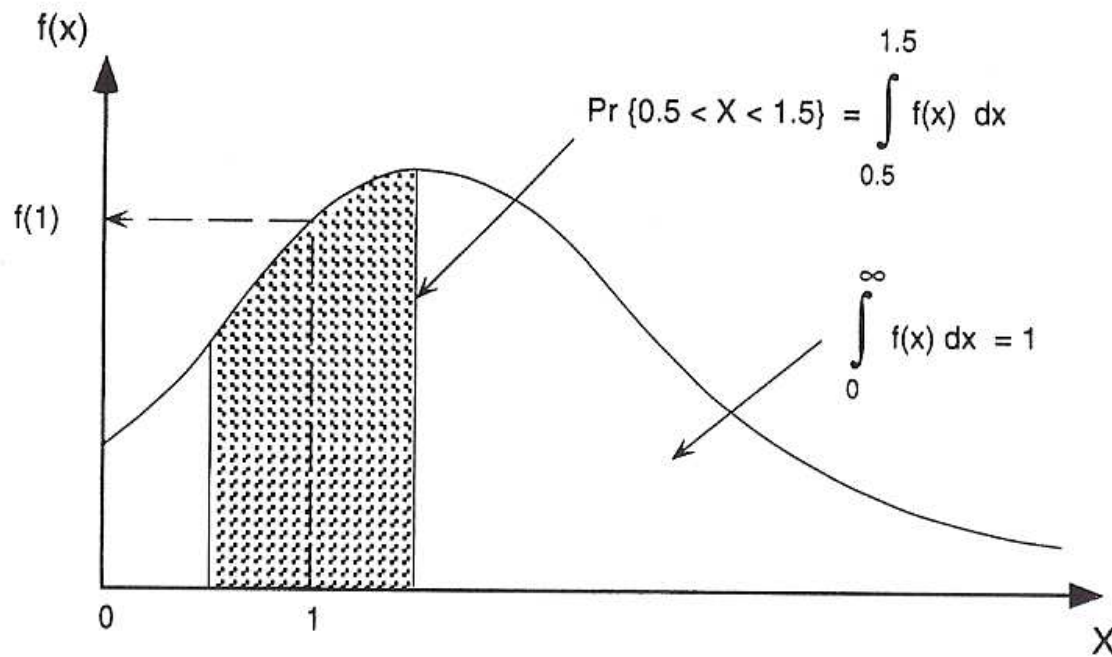
- A diferença entre as distribuições contínuas e discretas é que as distribuições discretas envolvem somas sobre uma função de probabilidade descontínua, enquanto a probabilidade para variáveis aleatórias contínuas envolve integração sobre funções contínuas denominadas **FUNÇÃO DENSIDADE DE PROBABILIDADE**, ou (PDFs).
- *Convencionalmente, a PDF para uma variável aleatória  $X$  é denominada  $f(x)$ .*

# PDF ou FDP

- Uma função densidade de probabilidade, PDF, tem que satisfazer a seguinte condição.

$$\int_x f(x) = 1$$

- Além disso,  $f(x)$  não pode ser negativo para todos os valores de  $x$ .



A  $f(x)$  tem significado quando pensamos em calcular probabilidades para valores de uma variável aleatória em uma vizinhança não infinitesimal em torno de um ponto, por exemplo  $X=1$ .

- Uma idéia relacionada com a PDF é aquela de uma função de distribuição cumulativa (CDF). A CDF é uma função da variável aleatória  $X$ , dada pela integral da PDF até um valor particular de  $x$ . Convencionalmente, CDFs são denominadas  $F(x)$ :

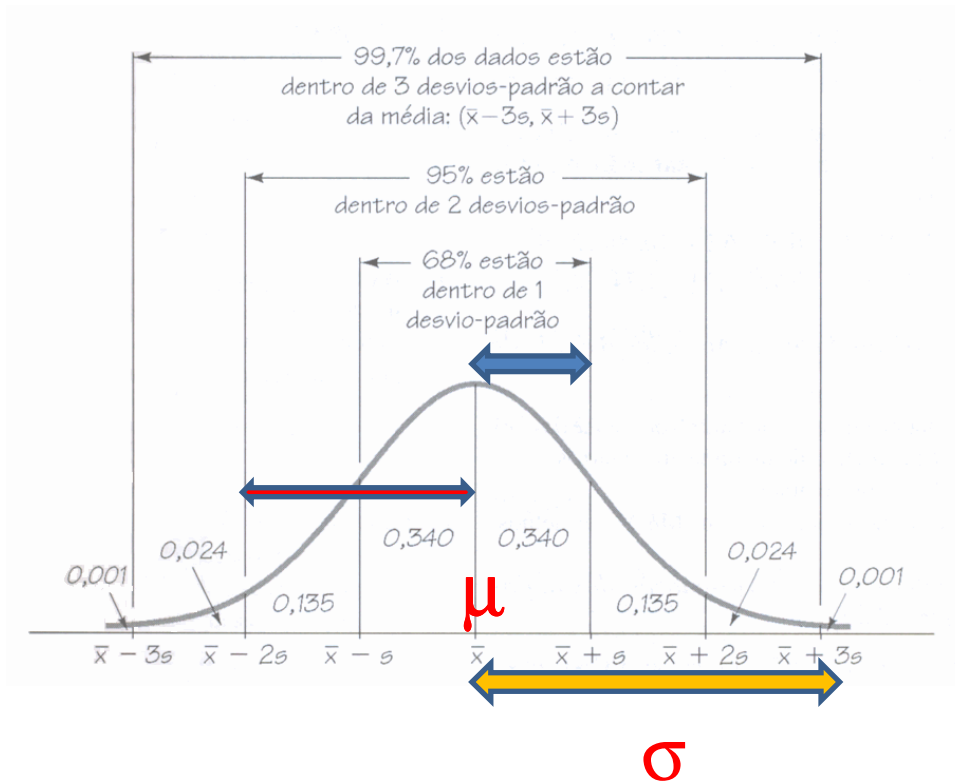
$$F(x) = \Pr\{X \leq x\} = \int_{X \leq x} f(x) dx$$

# Distribuição Normal

Uma variável aleatória contínua é representada por uma distribuição normal se a sua distribuição for:

- Simétrica
- A forma gráfica é similar a um sino

# Distribuição Normal



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

para  $-\infty < X < +\infty$

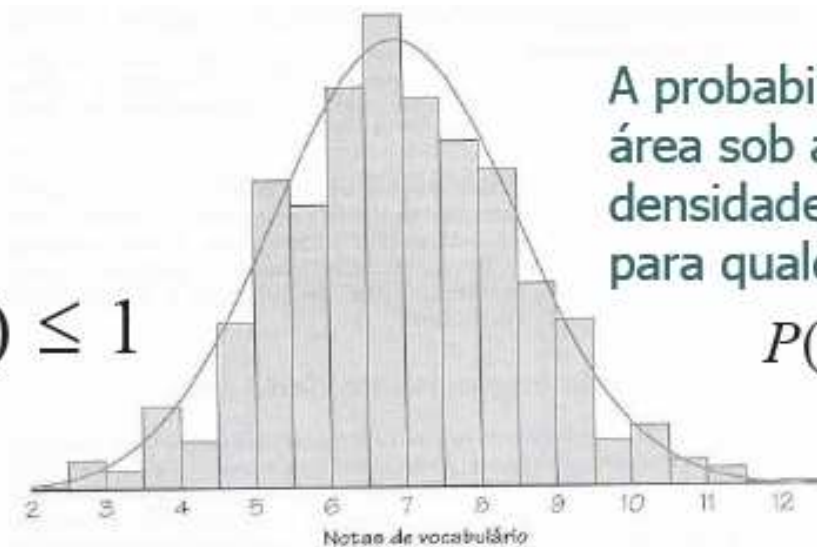
$\mu$  média da população

$\sigma$  desvio-padrão da população

A função densidade pode ser compreendida como uma extensão natural do histograma.

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

$$0 \leq P(x) \leq 1$$

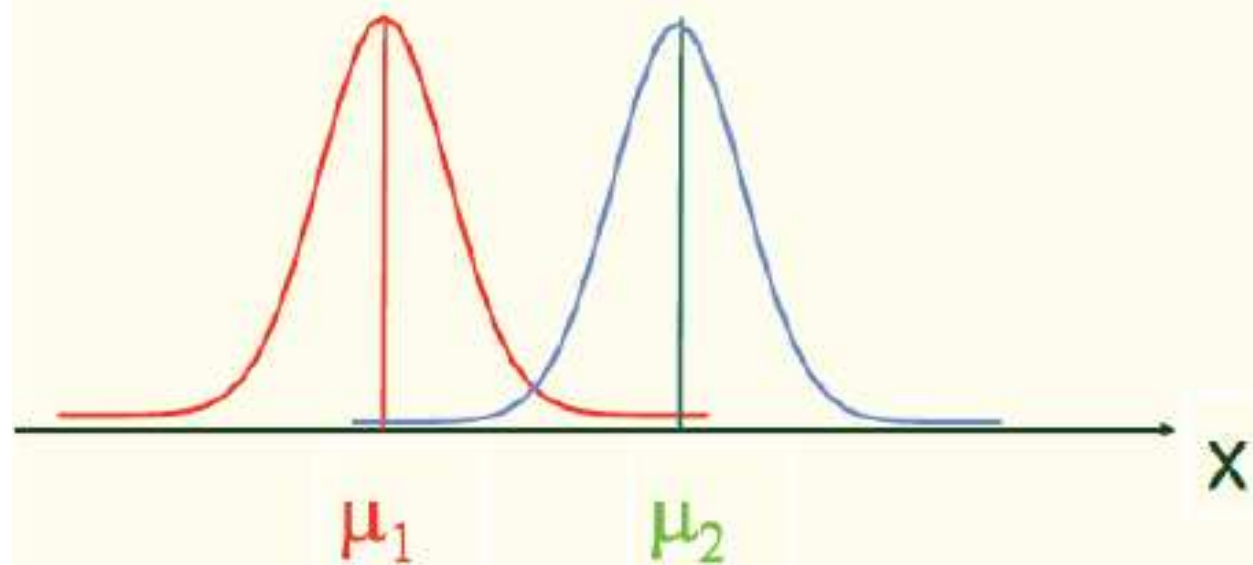


A probabilidade é a área sob a curva de densidade. Portanto, para qualquer  $P(x)$ :

$$P(x) \geq 0$$



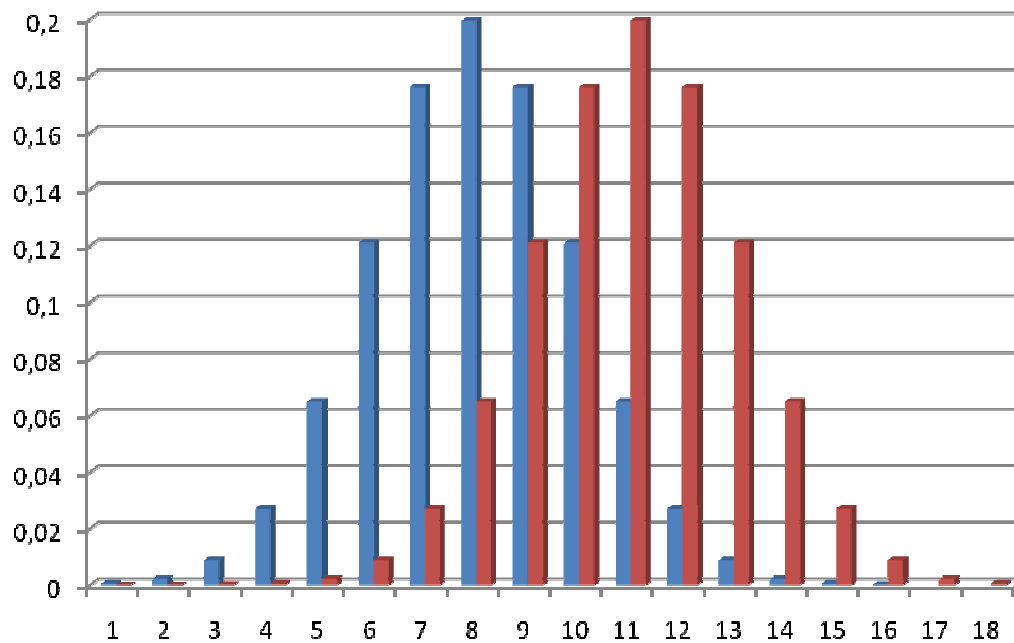
mesmo  $\sigma$  e diferentes  $\mu$



mesmo  $\mu$  e diferentes  $\sigma$



- Cada par de parâmetros ( $\mu$ ;  $\sigma$ ) define uma distribuição normal distinta;



- Convencionalmente, a variável aleatória descrita por uma distribuição Gaussiana padrão é denominada de  $Z$ .
- Sua densidade de probabilidade se simplifica a:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z)^2}{2}\right]$$

Onde  $Z$  é

$$Z = \frac{X - \mu}{\sigma}$$

- Quando estimamos os coeficientes da Gaussiana pelos dados, então utilizamos a seguinte notação (notem que agora a variável transformada é denotada como “z”):

$$z = \frac{x - \bar{x}}{s}$$

# Exemplo 1:

- Suponha que uma distribuição Gaussiana para o mês de janeiro em uma certa localidade seja caracterizada por  $\mu=22.2^{\circ}\text{C}$  e  $\sigma=4.4^{\circ}\text{C}$ .
- Suponha que você esteja interessado em avaliar a probabilidade de que um certo mês de janeiro tenha temperaturas menores ou iguais a  $21.4^{\circ}\text{C}$ .
- O primeiro passo para a solução desse problema é calcular o valor padronizado  $z$ .

$$z = (21.4^{\circ}\text{C} - 22.2^{\circ}\text{C})/4.4^{\circ}\text{C} = -0.18.$$

Assim, a probabilidade de uma temperatura igual ou mais fria que  $21.4^{\circ}\text{C}$  é a mesma que a probabilidade de um valor de  $Z$  igual ou menor que  $-0.18$ :

$$\Pr\{X \leq 21.4^{\circ}\text{C}\} = \Pr\{Z \leq -0.18\}.$$

# Como encontrar a probabilidade

- Identifique na tabela que contenha probabilidades acumuladas para a distribuição Gaussiana o valor  $z$ .
- Encontre a linha que indica  $-0.1$  e a coluna que mostra  $0.08$
- Você encontra a probabilidade de  $0.4286$ .

$P(Z < z)$

$z$	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455

# Como funciona a tabela

Coluna  $P(Z < z)$

Linha

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455

$$z = z(\text{linha}) + z(\text{coluna})$$



## Exemplo 2:

Considere uma variável aleatória  $X$  com  $\mu=15$  e  $\sigma=25$ .

Qual a probabilidade de que  $X$  assumam valores entre  $16 \leq X \leq 20$  ?

- A probabilidade desejada pode ser obtida, a partir da variável transformada (Z).

$$\text{Para } X = 16 \Rightarrow z = \frac{(16 - 15)}{25} = 0,04$$

$$\text{Para } X = 20 \Rightarrow z = \frac{(20 - 15)}{25} = 0,20$$

$$P(0,04 \leq Z \leq 0,20) = P(Z \leq 0,2) - P(Z \leq 0,04)$$

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8187	0,8214	0,8241	0,8267	0,8293	0,8318	0,8344	0,8369	0,8394

$$P(Z \leq 0,2) =$$

$$P(Z \leq 0,04) =$$

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133

$$P(Z \leq 0,2) = 0,5793$$

$$P(Z \leq 0,04) = 0,5160$$

$$P(0,04 \leq Z \leq 0,20) = P(Z \leq 0,2) - P(Z \leq 0,04)$$

$$P(0,04 \leq Z \leq 0,20) = 0,5793 - 0,5160 = 0,0633 \text{ ou } 6,33\%$$

## Exemplo 3

Queremos saber agora qual a temperatura que corresponde ao primeiro decil, sendo que esta distribuição de temperatura tem  $\mu=22.2^{\circ}\text{C}$  e  $\sigma=4.4^{\circ}\text{C}$ .

Para isso, utilizamos a tabela para achar o valor de probabilidade 0.10.

$$z = -1.28$$

Mas  $Z = \frac{X - \mu}{\sigma}$

Logo  $X = Z\sigma + \mu$

$$X = (-1.28) \times (4.4) + 22.2 = 16.6^{\circ} \text{ C}$$

# DISTRIBUIÇÃO GAMA

- Muitas variáveis atmosféricas são distintamente assimétricas, e possuem uma assimetria para a direita.
- *Um exemplo clássico é a precipitação. Sabemos que não existem precipitações negativas, certo?*
- Então suponha que uma certa localidade tenha uma média de precipitação diária de 1.96 in e desvio-padrão de 1.12 in.
- **Utilizando a tabela de distribuições acumuladas Gaussiana podemos calcular a probabilidade de precipitações negativas como**

$$\Pr \{ Z \leq (0.00 - 1.96)/1.12 \} = \Pr\{Z \leq -1.75\} = 0.040.$$

- Esta probabilidade calculada não é especialmente grande, mas por outro lado não pode ser considerada zero.
- **AGORA, SABEMOS PELA NATUREZA QUE PRECIPITAÇÕES NEGATIVAS SÃO IMPOSSÍVEIS!**

# DISTRIBUIÇÃO GAMA

$$f(x) = \frac{(x / \beta)^{\alpha-1} \exp(-x / \beta)}{\beta \Gamma(\alpha)} \quad \text{onde } x, \alpha, \beta > 0$$

$\alpha$  PARÂMETRO DE FORMA ,

$\beta$  PARÂMETRO DE ESCALA.

$\Gamma(\alpha)$  FUNÇÃO GAMMA, definida pela integral

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$



$\Gamma(\alpha)$  FUNÇÃO GAMMA

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

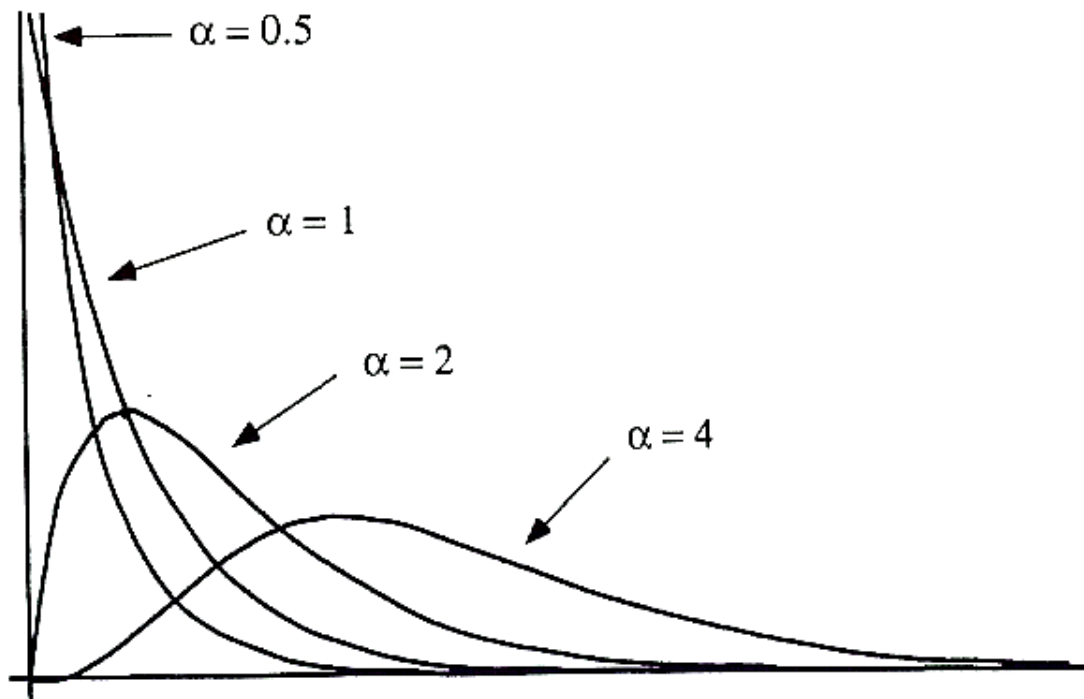
$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \quad \text{ou} \quad \Gamma(\alpha + 1) = \alpha!$$

Exemplo 1

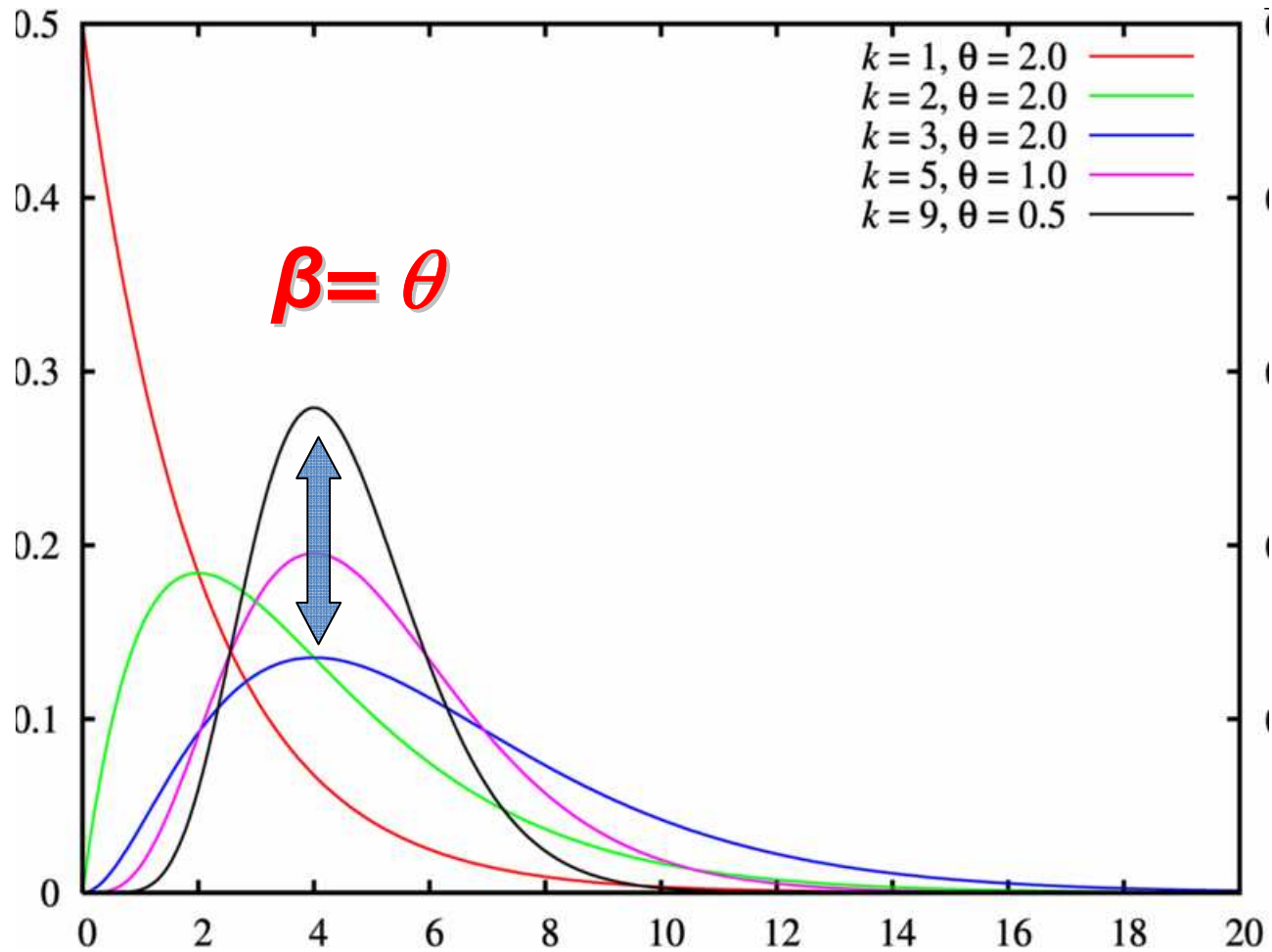
$$\begin{aligned} \Gamma(3.5) &= 2.5 \times \Gamma(2.5) \\ &= 2.5 \times 1.5 \times \Gamma(1.5) \\ &= 2.5 \times 1.5 \times 0.8862 = 3.323. \end{aligned}$$

Exemplo 2

$$\begin{aligned} \Gamma(4) &= \Gamma(3 + 1) \\ &= 3! \\ &= 3 \times 2 = 6 \end{aligned}$$



$\alpha$  PARÂMETRO DE FORMA e  $\beta$  PARÂMETRO DE ESCALA



$\alpha=k$  PARÂMETRO DE FORMA e  $\beta= \theta$  PARÂMETRO DE ESCALA

A PDF da distribuição Gamma pode apresentar uma grande variedade de formas, dependendo, portanto, do parâmetro de forma  $\alpha$ .

Para valores de  $\alpha$  muito altos, a distribuição gamma tende à Gaussiana

O parâmetro de escala  $\beta$ , tem a função de ESTICAR OU ENCOLHER (isto é escalonar) a função de densidade gamma para a direita ou esquerda, dependendo das magnitudes gerais dos valores dos dados representados.

## Calculando os fatores da função gama

- Existem duas maneiras:
  - 1) Estimadores de Thom – Thom (1958))
  - 2) Aproximação polinomial

# Estimadores de Thom – Thom

$$\hat{\alpha} = \frac{1 + \sqrt{1 + 4D/3}}{4D}$$

parâmetro de forma  $\alpha$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}$$

parâmetro de escala  $\beta$

$$D = \ln(\bar{x}) - \frac{1}{n} \sum_{i=1}^n \ln(x_i)$$

# Aproximação polinomial

parâmetro de forma  $\alpha$

Para  $0 \leq D \leq 0.5772$

$$\hat{\alpha} = \frac{0.5000876 + 0.1648852D - 0.0544274D^2}{D}$$

Para  $0.5772 \leq D \leq 17$

$$\hat{\alpha} = \frac{8.898919 + 9.059950D + 0.9775373D^2}{17.79728D + 11.968477D^2 + D^3}$$

parâmetro de escala  $\beta$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}$$

$$D = \ln(\bar{x}) - \frac{1}{n} \sum_{i=1}^n \ln(x_i)$$

# Para calcular a Probabilidade

é sempre necessário fazer uma transformação para re-escalonar a variável  $X$  de interesse (caracterizada por uma gama com parâmetro de escala arbitrário  $\beta$ ) para a variável

$$\xi = \frac{x}{\beta}$$



# Exemplo 1

Suponha que a a média aritmética da ppt é igual a 1.96 polegadas.

***Sera que a precipitação de janeiro de 1987, 3.15 polegadas é usual?***

- 1) Calculando o valor da média dos logaritmos dos totais mensais = 0.5346
- 2) Obtemos o valor de  $D = 0.139$
- 3) De acordo com o método de Thom  $\alpha=3.76$  e  $\beta=0.52$  polegadas.

$$\zeta = (X/\beta) = 3.15 / 0.52 = 6.06.$$

O passo seguinte é encontrar no corpo da tabela onde se encontra a probabilidade 6.06 para  $\alpha=3.76$ . **Tabela 2 - Wilks**

- Este valor se encontra entre os valores tabulados  $F(5.214)=0.90$  e  $F(6.354)=0.90$ .
- Logo uma interpolação nos dá  $F(6.06)=0.874$ .
- Portanto a probabilidade de chover em janeiro menos ou igual a 3.15 polegadas em Ithaca é de 0.874.
- O complementar  $(1 - 0.874) = 0.126$  é a probabilidade de chover mais do que esse valor (a qual é equivalente a aproximadamente 1 chance em 8 ( $1/8$ )).

# Lista de Exercício 3

## Entrega: 17 de Junho

- Utilizando como referência a temperatura média do ar e o respectivo desvio padrão calculado na lista de exercício 1 e que a distribuição de temperatura do ar pode ser representada por uma distribuição normal, calcule.
- 1) Qual a probabilidade de termos temperaturas acima de 25,30 e 35oC ?
- 2) Qual a probabilidade de a temperatura do ar estiver entre o intervalo de 25 a 30oC ?
- 3) Qual a probabilidade de termos temperaturas abaixo de 10,15 e 20oC ?
- 4) Qual é a temperatura que representa um decil?