

06 - Ruído - noções

Por que os sinais de uma classe são apenas semelhantes, e não iguais ?

Hipóteses:

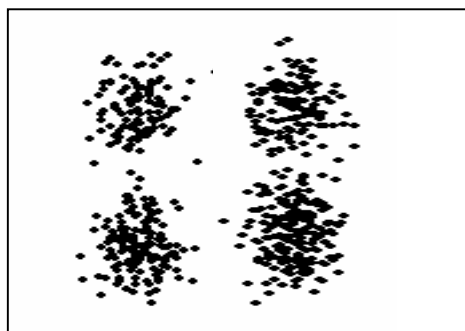
1 - Os elementos $\underline{x}_j$ de uma classe C_j são constituídos pelo padrão $\underline{m}_j$ da classe, ao qual foi adicionado um ruído $\underline{r}_j$ característico da classe C_j .

$$\vec{x} \in C_j \quad \therefore \quad \vec{x} = \vec{m}_j + \vec{r}_j$$

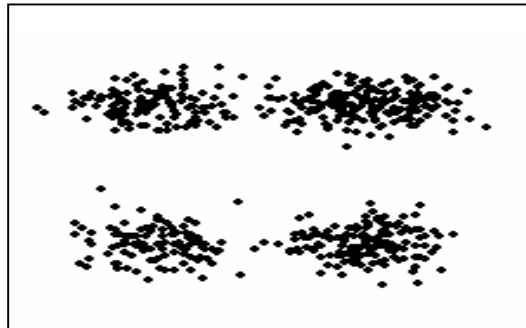
2 - O ruído $\underline{r}_j$ é estocástico com média zero e distribuição aproximadamente gaussiana.

3 - As características do ruído independem da classe (ou não).

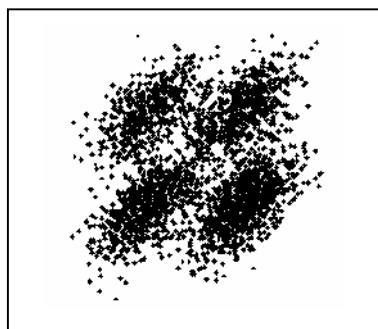
4 - Uma classe pode ser constituída por subclasses, às quais se aplicam as hipóteses acima.



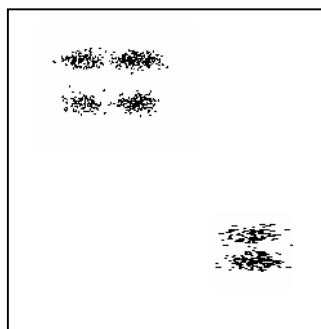
Ruído não correlato, $\sigma_1 \cong \sigma_2$



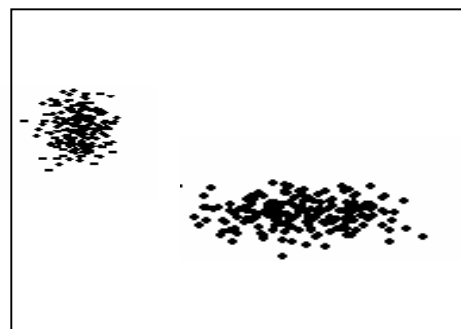
Ruído não correlato, $\sigma_1 \neq \sigma_2$



Ruído correlato, $x_1 = f(x_2)$



Classes compostas



Hipótese 3 não atendida



Classificação por similaridade



Caracterização do ruído

1 - Caso supervisionado

As classes de cada entrada (elemento) $\underline{x}$ são conhecidas. Neste caso seu baricentro $\vec{m}_j$ também é conhecido

$$\vec{m}_j = E_{\forall \vec{x}_i \in C_j} \vec{x}_i = \frac{1}{n_j} \sum_{\forall \vec{x}_i \in C_j} \vec{x}_i$$

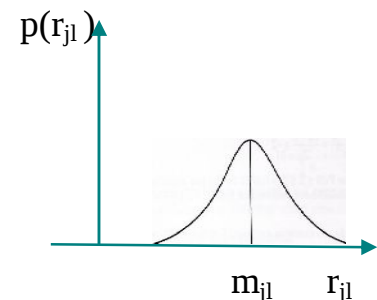
Pares entrada-saída $[\vec{x}, C(\vec{x})]$

A caracterização do ruído de cada classe C_j , $\vec{r}_j$, em cada dimensão l , r_{jl} , pode ser realizada a partir da extração de $\vec{r}_j$ das entradas pertencentes à classe

$$\forall \vec{x} \in C_j \quad \vec{x} = \vec{m}_j + \vec{r}_j \quad \therefore \quad \vec{r}_j = \vec{x} - \vec{m}_j \quad e \quad r_{jl} = x_l - m_{jl}$$

A análise do conjunto das variáveis r_{jl} permite conhecer sua distribuição, seu valor eficaz σ , sua correlação com o ruído nas demais dimensões $r_{jk} \mid k \neq l$, etc.

Isto permitirá descorrelacionar os ruídos das diferentes dimensões, normalizar o valor eficaz por dimensão, etc., obtendo um conjunto de dados muito mais simples de classificar, como será visto mais tarde.



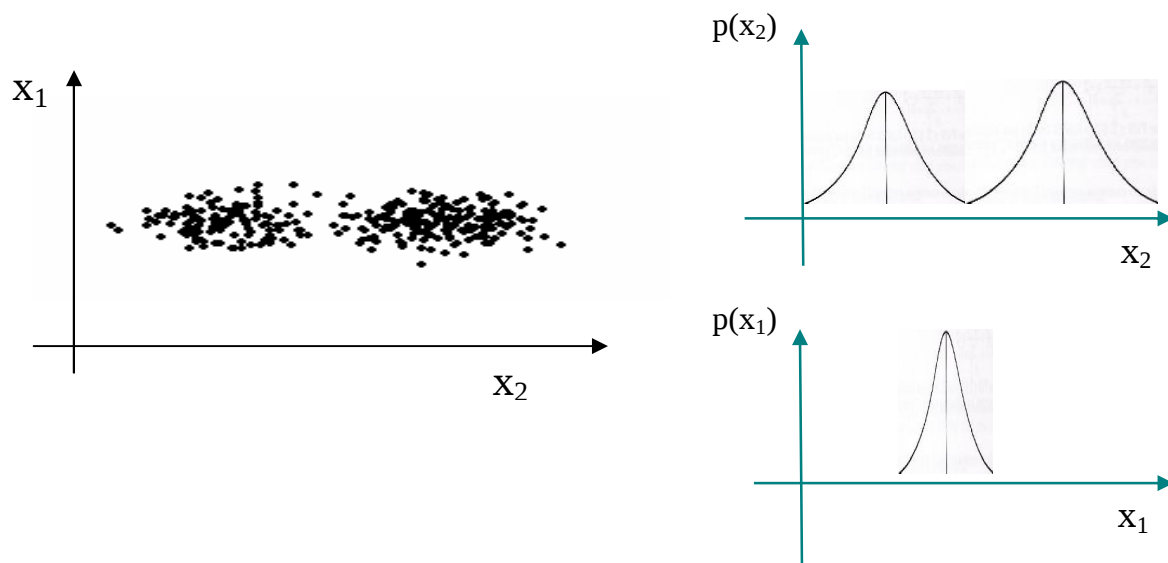
2 - Caso não supervisionado

As classes de cada entrada (elemento) $\underline{x}$ não são conhecidas

Se as classes apresentarem uma dispersão inter classes (distância) razoavelmente maior que intra classe (diâmetros) a análise de $p(x)$ em cada dimensão pode (eventualmente) indicar as características de cada classe em cada dimensão (centro, ruído, etc.).

Para uma classe
$$p(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - m_i)^2}{2\sigma_i^2}\right)$$

Para N classes
$$p(x_i) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - m_i)^2}{2\sigma_i^2}\right)$$



?

Tentativa e erro por clusterização experimental.

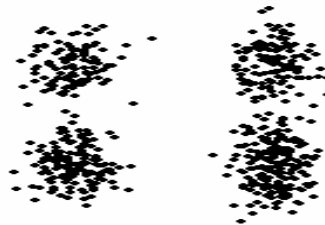
Outras técnicas

3 – Ruído - caso ideal

Classes com ruídos por componente com mesma variância e não correlatos - Classes esféricas

Considerando o caso em que:

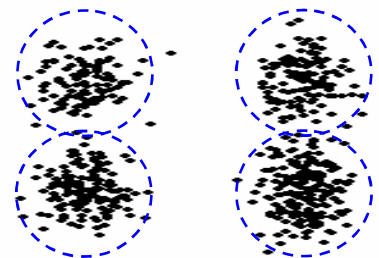
- As classes foram geradas a partir de padrões adicionados de ruído aproximadamente gaussiano com média nula.
- que o valor rms (desvio padrão) do ruído de cada componente x_j é σ , independente da componente e da classe.



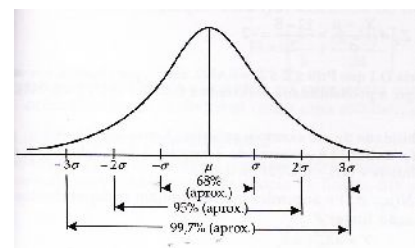
então

o separador ideal para as classes são esferas, que são também o separador natural por similaridade dos processos não supervisionados.

Se o ruído é gaussiano então as classes tem distribuição gaussiana multidimensional com desvio padrão unitário em todas as direções. Em consequência em cada dimensão o raio r_0



$r_0 = \sigma$ captura 68 % dos elementos da classe,
 $r_0 = 2\sigma$ captura 95 % dos elementos da classe,
 $r_0 = 3\sigma$ captura 99,7 % dos elementos da classe, etc.



Tentaremos chegar próximo a esta condição através do pré-processamento

4 - Branqueamento do Ruído – gerando classes esféricas *classes sphereing*

Se os ruídos das componentes forem correlatos os domínios das classes serão elipsoides e estarão contidos por esferas maiores, que incluem espaços da não classe.

O branqueamento do ruído leva a classes esféricas, que permitirá o uso de um classificador muito mais simples. O branqueamento do ruído dos dados é obtido em duas etapas.



Inicialmente a decorrelação dos ruídos r_i deve ser obtida representando os dados $\underline{x}$ (padrão da classe + ruído) em uma base $\underline{B}$ do espaço das componentes principais (PCA) do ruído $\underline{r}$ de cada dado de entrada (**PCA dos ruídos, e não dos sinais!**). Calcular as PCA de $\underline{r}$ e a matriz de mapeamento $\underline{B}$ do ruído $\underline{r}$ em sua PCA $\underline{p}$

$$\vec{x} \in C_j \quad \vec{x} = \vec{m}_j + \vec{r}_j \quad \therefore \quad \vec{r} = \vec{x} - \vec{m}_j$$

Utilizando $\underline{B}$ determinamos o mapeamento de cada ruído $\underline{r}$ em sua PCA $\underline{p}$

$$\vec{p} = B \vec{r} \quad \text{onde} \quad [B] = [\vec{b}_1^t \vec{b}_2^t \vec{b}_3^t \dots \vec{b}_m^t]$$

e a variância de cada componente p_i

$$\sigma_i^2 = E[p_i]^2 = E[\vec{b}_i^t \vec{r}]^2$$

A segunda etapa deve ser a equalização da potência do ruído nas diversas dimensões. A matriz de pesos $\underline{P}$ para normalização das potências dos ruídos será

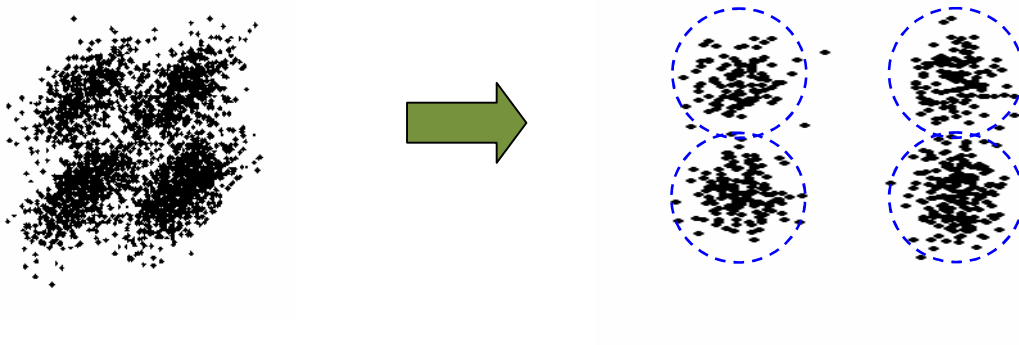
$$P = \text{diag} \left(\frac{1}{\sigma_i} \right)$$

Os novos dados à classificar serão

$$\vec{z} = P B \vec{x} = P B \vec{m}_j + P B \vec{r} = \vec{m}_{j_{novo}} + \vec{r}_{novo}$$

As classes passam a ter um novo padrão $\vec{m}_{j_{novo}}$ adicionado de um ruído branco $\vec{r}_{novo}$ com desvio padrão unitário.

$$\vec{m}_{j_{novo}} = P B \vec{m}_j \quad \vec{r}_{novo} = P B \vec{r} = \vec{1} = [1 \ 1 \ 1 \dots 1]^t$$



4.1 Algumas observações:

- A multiplicação por B representa uma rotação e por isto não altera a posição relativa das classes, como será visto adiante.
- O uso de PCA sugere a eliminação de componentes pouco relevantes, i.e., com pequenos

$$\sigma_i^2 = E[p_i]^2 = E[\vec{b}_i^t \vec{r}]^2$$

e a consequente redução da dimensionalidade e simplificação do classificador. Entretanto, a eliminação de uma componente principal para $\vec{r}_j$ somente pode ser feita se for pouco relevante para $\vec{x}$, isto é, se

$$\sigma^2(z_i) = E\left[\frac{1}{\sigma_i} \vec{b}_i^t \vec{x}\right]^2 \text{ for muito pequeno comparado com seus pares.}$$

A relevância para $\vec{x}$ é que determina a eliminação ou não da componente !

- A multiplicação por P realiza um escalamento diferente por componente e pode alterar a classificação, principalmente se houver grande dispersão entre os σ_i , como será visto adiante,
- Caso os ruídos sejam não correlatos, a primeira etapa é obviamente desnecessária

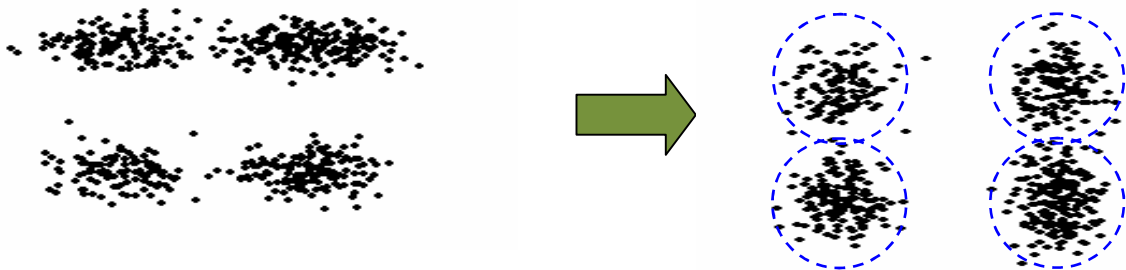
$$B = I$$

E a dispersão em cada dimensão i é dada por

$$\sigma_i^2 = E[r_i]^2 \quad \text{e} \quad P = \text{diag}\left(\frac{1}{\sigma_i}\right)$$

Os novos dados à classificar serão

$$\vec{z} = P \vec{x} = P \vec{m}_j + P \vec{r} = \vec{m}_{j_{\text{novo}}} + \vec{r}_{\text{novo}}$$



5. Algumas distribuições importantes:

5.1 Distribuição Normal ou Gaussiana

A distribuição Normal ou Gaussiana é uma das distribuições de probabilidade mais utilizadas para modelar fenômenos da natureza porque um grande número destes fenômenos apresenta sua distribuição de probabilidade aproximadamente normal, por exemplo o ruído branco.

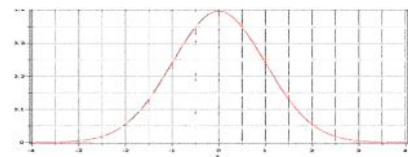
Em nosso caso consideramos que os elementos $\underline{x}$ de uma classe C são compostos pelo padrão $\underline{m}$ da classe adicionado de ruído branco $\underline{r}$.

$$\forall \vec{x} \in C_j \quad \vec{x} = \vec{m}_j + \vec{r} \quad \therefore$$

$$\vec{r} = \vec{x} - \vec{m}_j \quad e \quad r_i = x_i - m_{ij}$$

Função de distribuição de probabilidade, f.d.p.

$$p(r_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{r_i^2}{2\sigma^2}}$$

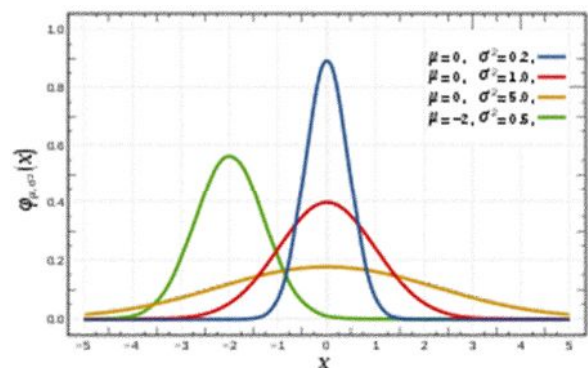


Média = mediana = moda

$$\mu = 0$$

Desvio padrão

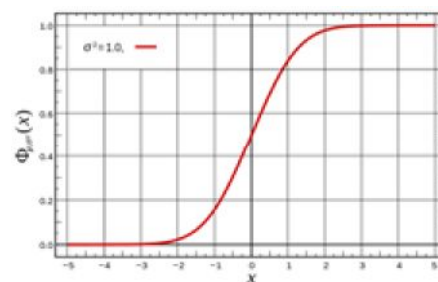
$$\sigma$$



Função erro (uma das definições)

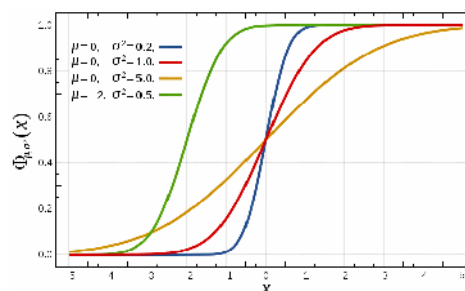
Função de distribuição acumulada

$$erf(r_0) = p(r_i | r_i < r_0) = \int_{-\infty}^{r_0} p(t) dt$$

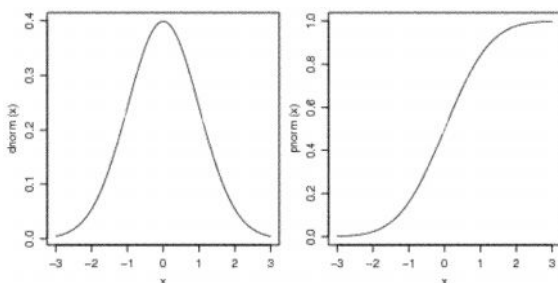


Sem expressão analítica,

tabelada para $\sigma = 1$



Gaussiana e erf para $\mu=0$ e $\sigma=1$



47

AREAS UNDER THE
STANDARD NORMAL CURVE

From $-\infty$ to z

$$\text{erf}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$$

z	0	1	2	3	4	5	6	7	8
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319
0.1	.5060	.5100	.5140	.5179	.5219	.5259	.5299	.5339	.5379
0.2	.5120	.5160	.5199	.5239	.5279	.5319	.5359	.5399	.5439
0.3	.5179	.5219	.5259	.5299	.5339	.5379	.5419	.5459	.5499
0.4	.5219	.5259	.5299	.5339	.5379	.5419	.5459	.5499	.5539
0.5	.5259	.5299	.5339	.5379	.5419	.5459	.5499	.5539	.5579
0.6	.5299	.5339	.5379	.5419	.5459	.5499	.5539	.5579	.5619
0.7	.5339	.5379	.5419	.5459	.5499	.5539	.5579	.5619	.5659
0.8	.5379	.5419	.5459	.5499	.5539	.5579	.5619	.5659	.5699
0.9	.5419	.5459	.5499	.5539	.5579	.5619	.5659	.5699	.5739
1.0	.5459	.5499	.5539	.5579	.5619	.5659	.5699	.5739	.5779
1.1	.5499	.5539	.5579	.5619	.5659	.5699	.5739	.5779	.5819
1.2	.5539	.5579	.5619	.5659	.5699	.5739	.5779	.5819	.5859
1.3	.5579	.5619	.5659	.5699	.5739	.5779	.5819	.5859	.5899
1.4	.5619	.5659	.5699	.5739	.5779	.5819	.5859	.5899	.5939
1.5	.5659	.5699	.5739	.5779	.5819	.5859	.5899	.5939	.5979
1.6	.5699	.5739	.5779	.5819	.5859	.5899	.5939	.5979	.6019
1.7	.5739	.5779	.5819	.5859	.5899	.5939	.5979	.6019	.6059
1.8	.5779	.5819	.5859	.5899	.5939	.5979	.6019	.6059	.6099
1.9	.5819	.5859	.5899	.5939	.5979	.6019	.6059	.6099	.6139
2.0	.5859	.5899	.5939	.5979	.6019	.6059	.6099	.6139	.6179
2.1	.5899	.5939	.5979	.6019	.6059	.6099	.6139	.6179	.6219
2.2	.5939	.5979	.6019	.6059	.6099	.6139	.6179	.6219	.6259
2.3	.5979	.6019	.6059	.6099	.6139	.6179	.6219	.6259	.6299
2.4	.6019	.6059	.6099	.6139	.6179	.6219	.6259	.6299	.6339
2.5	.6059	.6099	.6139	.6179	.6219	.6259	.6299	.6339	.6379
2.6	.6099	.6139	.6179	.6219	.6259	.6299	.6339	.6379	.6419
2.7	.6139	.6179	.6219	.6259	.6299	.6339	.6379	.6419	.6459
2.8	.6179	.6219	.6259	.6299	.6339	.6379	.6419	.6459	.6499
2.9	.6219	.6259	.6299	.6339	.6379	.6419	.6459	.6499	.6539
3.0	.6259	.6299	.6339	.6379	.6419	.6459	.6499	.6539	.6579
3.1	.6299	.6339	.6379	.6419	.6459	.6499	.6539	.6579	.6619
3.2	.6339	.6379	.6419	.6459	.6499	.6539	.6579	.6619	.6659
3.3	.6379	.6419	.6459	.6499	.6539	.6579	.6619	.6659	.6699
3.4	.6419	.6459	.6499	.6539	.6579	.6619	.6659	.6699	.6739
3.5	.6459	.6499	.6539	.6579	.6619	.6659	.6699	.6739	.6779
3.6	.6499	.6539	.6579	.6619	.6659	.6699	.6739	.6779	.6819
3.7	.6539	.6579	.6619	.6659	.6699	.6739	.6779	.6819	.6859
3.8	.6579	.6619	.6659	.6699	.6739	.6779	.6819	.6859	.6899
3.9	.6619	.6659	.6699	.6739	.6779	.6819	.6859	.6899	.6939
4.0	.6659	.6699	.6739	.6779	.6819	.6859	.6899	.6939	.6979
4.1	.6699	.6739	.6779	.6819	.6859	.6899	.6939	.6979	.7019
4.2	.6739	.6779	.6819	.6859	.6899	.6939	.6979	.7019	.7059
4.3	.6779	.6819	.6859	.6899	.6939	.6979	.7019	.7059	.7099
4.4	.6819	.6859	.6899	.6939	.6979	.7019	.7059	.7099	.7139
4.5	.6859	.6899	.6939	.6979	.7019	.7059	.7099	.7139	.7179
4.6	.6899	.6939	.6979	.7019	.7059	.7099	.7139	.7179	.7219
4.7	.6939	.6979	.7019	.7059	.7099	.7139	.7179	.7219	.7259
4.8	.6979	.7019	.7059	.7099	.7139	.7179	.7219	.7259	.7299
4.9	.7019	.7059	.7099	.7139	.7179	.7219	.7259	.7299	.7339
5.0	.7059	.7099	.7139	.7179	.7219	.7259	.7299	.7339	.7379
5.1	.7099	.7139	.7179	.7219	.7259	.7299	.7339	.7379	.7419
5.2	.7139	.7179	.7219	.7259	.7299	.7339	.7379	.7419	.7459
5.3	.7179	.7219	.7259	.7299	.7339	.7379	.7419	.7459	.7499
5.4	.7219	.7259	.7299	.7339	.7379	.7419	.7459	.7499	.7539
5.5	.7259	.7299	.7339	.7379	.7419	.7459	.7499	.7539	.7579
5.6	.7299	.7339	.7379	.7419	.7459	.7499	.7539	.7579	.7619
5.7	.7339	.7379	.7419	.7459	.7499	.7539	.7579	.7619	.7659
5.8	.7379	.7419	.7459	.7499	.7539	.7579	.7619	.7659	.7699
5.9	.7419	.7459	.7499	.7539	.7579	.7619	.7659	.7699	.7739
6.0	.7459	.7499	.7539	.7579	.7619	.7659	.7699	.7739	.7779
6.1	.7499	.7539	.7579	.7619	.7659	.7699	.7739	.7779	.7819
6.2	.7539	.7579	.7619	.7659	.7699	.7739	.7779	.7819	.7859
6.3	.7579	.7619	.7659	.7699	.7739	.7779	.7819	.7859	.7899
6.4	.7619	.7659	.7699	.7739	.7779	.7819	.7859	.7899	.7939
6.5	.7659	.7699	.7739	.7779	.7819	.7859	.7899	.7939	.7979
6.6	.7699	.7739	.7779	.7819	.7859	.7899	.7939	.7979	.8019
6.7	.7739	.7779	.7819	.7859	.7899	.7939	.7979	.8019	.8059
6.8	.7779	.7819	.7859	.7899	.7939	.7979	.8019	.8059	.8099
6.9	.7819	.7859	.7899	.7939	.7979	.8019	.8059	.8099	.8139
7.0	.7859	.7899	.7939	.7979	.8019	.8059	.8099	.8139	.8179
7.1	.7899	.7939	.7979	.8019	.8059	.8099	.8139	.8179	.8219
7.2	.7939	.7979	.8019	.8059	.8099	.8139	.8179	.8219	.8259
7.3	.7979	.8019	.8059	.8099	.8139	.8179	.8219	.8259	.8299
7.4	.8019	.8059	.8099	.8139	.8179	.8219	.8259	.8299	.8339
7.5	.8059	.8099	.8139	.8179	.8219	.8259	.8299	.8339	.8379
7.6	.8099	.8139	.8179	.8219	.8259	.8299	.8339	.8379	.8419
7.7	.8139	.8179	.8219	.8259	.8299	.8339	.8379	.8419	.8459
7.8	.8179	.8219	.8259	.8299	.8339	.8379	.8419	.8459	.8499
7.9	.8219	.8259	.8299	.8339	.8379	.8419	.8459	.8499	.8539
8.0	.8259	.8299	.8339	.8379	.8419	.8459	.8499	.8539	.8579
8.1	.8299	.8339	.8379	.8419	.8459	.8499	.8539	.8579	.8619
8.2	.8339	.8379	.8419	.8459	.8499	.8539	.8579	.8619	.8659
8.3	.8379	.8419	.8459	.8499	.8539	.8579	.8619	.8659	.8699
8.4	.8419	.8459	.8499	.8539	.8579	.8619	.8659	.8699	.8739
8.5	.8459	.8499	.8539	.8579	.8619	.8659	.8699	.8739	.8779
8.6	.8499	.8539	.8579	.8619	.8659	.8699	.8739	.8779	.8819
8.7	.8539	.8579	.8619	.8659	.8699	.8739	.8779	.8819	.8859
8.8	.8579	.8619	.8659	.8699	.8739	.8779	.8819	.8859	.8899
8.9	.8619	.8659	.8699	.8739	.8779	.8819	.8859	.8899	.8939
9.0	.8659	.8699	.8739	.8779	.8819	.8859	.8899	.8939	.8979
9.1	.8699	.8739	.8779	.8819	.8859	.8899	.8939	.8979	.9019
9.2	.8739	.8779	.8819	.8859	.8899	.8939	.8979	.9019	.9059
9.3	.8779	.8819	.8859	.8899	.8939	.8979	.9019	.9059	.9099
9.4	.8819	.8859	.8899	.8939	.8979	.9019	.9059	.9099	.9139
9.5	.8859	.8899	.8939	.8979	.9019	.9059	.9099	.9139	.9179
9.6	.8899	.8939	.8979	.9019	.9059	.9099	.9139	.9179	.9219
9.7	.8939	.8979	.9019	.9059	.9099	.9139	.9179	.9219	.9259
9.8	.8979	.9019	.9059	.9099	.9139	.9179	.9219	.9259	.9299
9.9	.9019	.9059	.9099	.9139	.9179	.9219	.9259	.9299	.9339
10.0	.9059	.9099	.9139	.9179	.9219	.9259	.9299	.9339	.9379

5.2 Distribuição Chi-quadrado χ_k^2

A distribuição qui-quadrado é importante porque permite calcular a probabilidade de uma entrada $\underline{x}$ afetada de ruído branco $\underline{r}$ estar a uma distância $\underline{d}$ do padrão da classe $\underline{m}$, isto é, permite estimar a probabilidade da similaridade de uma entrada com seu padrão.

$$\forall \vec{x} \in C_j \quad \vec{x} = \vec{m}_j + \vec{r} \quad \therefore \quad \vec{r} = \vec{x} - \vec{m}_j$$

$$r_i = x_i - m_{ij} \quad e$$

$$d^2 = \left| \vec{x} - \vec{m}_j \right|^2 = \left| \vec{r} \right|^2 = \sum_{i=1}^k r_i^2 = \sum_{i=1}^k (x_i - m_{ij})^2 = -u$$

Para o caso ideal, com todos os r_i não correlatos e com $\sigma_i = 1$

em cada dimensão i de $\vec{x}$ $d_i^2 = r_i^2 = -u_i$

distribuição χ_1^2 com 1(um) grau de liberdade

em todas as k dimensões de $\vec{x}$ $d^2 = \sum_{j=1}^k r_j^2 = -u$

distribuição χ_k^2 com k graus de liberdade

Função densidade de probabilidade χ_k^2

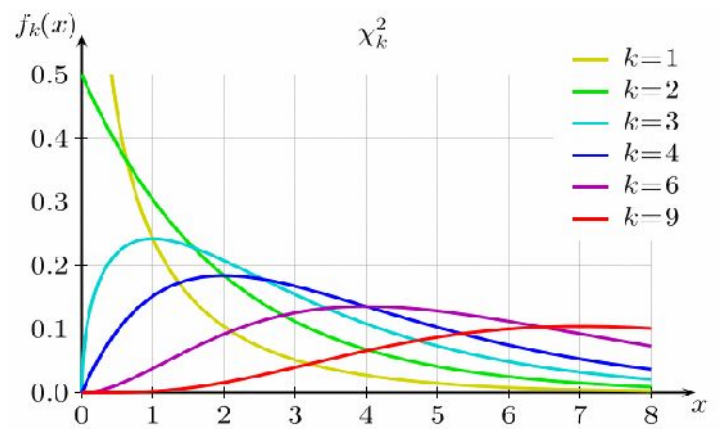
$$p(r) = \frac{1}{2^{k/2} \Gamma(k/2)} r^{k/2-1} e^{-r/2}$$

$$\text{média} = k$$

$$\text{mediana} = \approx k \left(1 - \frac{2}{9k} \right)^3$$

$$\text{moda} = \max[0, (k-2)]$$

$$\text{variância} = 2k$$



Função erro (uma das definições) de χ_k^2

Função de distribuição acumulada

$$\text{erf}(r_0) = p(r | r < r_0) = \int_{-\infty}^{r_0} p(t) dt = \frac{1}{\Gamma(k/2)} \gamma\left(\frac{k}{2}, \frac{r_0}{2}\right)$$

tabelada para diversos k

