

Classificação por Similaridade

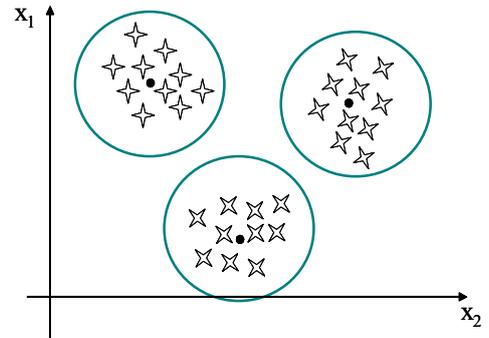
Crterios: máxima similaridade + similaridade mínima

Domínio da classe: esfera centrada no padrão da classe

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

Ruído

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

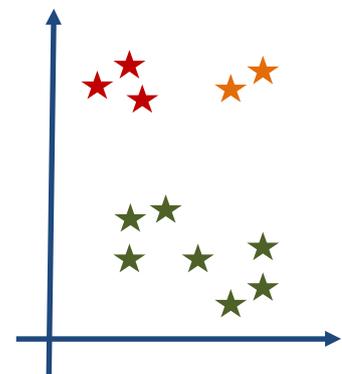


Ruído

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



Classificação por similaridade



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}

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

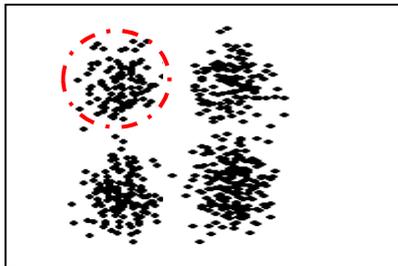
$$\vec{r} = \begin{bmatrix} r_1 \\ \dots \\ r_n \end{bmatrix}$$


2 - O ruído r_i em cada dimensão i é estocástico com média zero e distribuição aproximadamente gaussiana.

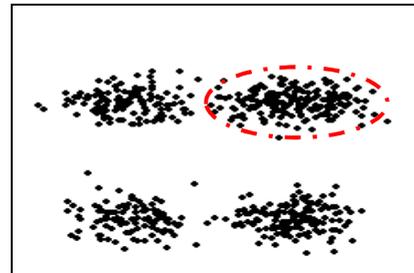
3 - As características do ruído independem da classe mas podem variar para cada componente x_i de \vec{x} .

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

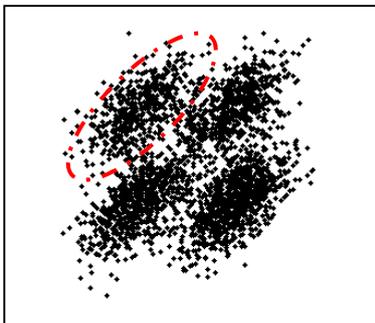
Classes e domínios



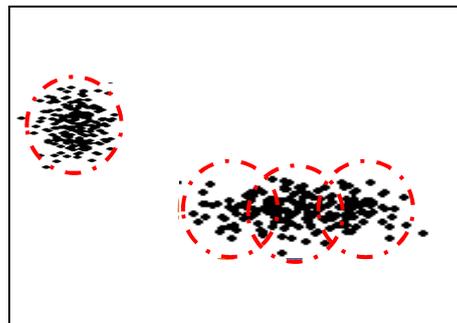
Ruído não correlato, $\sigma_1 \cong \sigma_2$



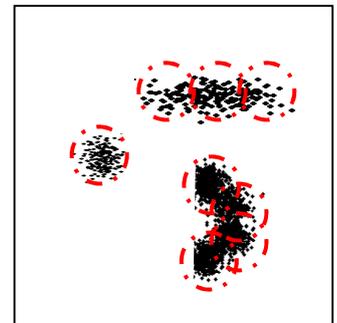
Ruído não correlato, $\sigma_1 \neq \sigma_2$



Ruído correlato, $r_1 = f(r_2)$



Hipótese 3 não atendida



Classes compostas

Caracterização do ruído

1 - Caso supervisionado

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

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

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

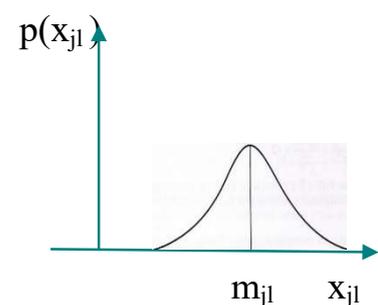
A caracterização do ruído das classes em cada dimensão l , r_l , pode ser realizada a partir da extração de \vec{r} das entradas pertencentes à classe

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

A análise do conjunto das variáveis r_l permite conhecer sua distribuição, seu valor eficaz σ_l , sua correlação com o ruído nas demais dimensões

$r_k \mid k \neq l$ (se houver), 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.



Um caso clássico: Íris

- Comprimento da pétala
- Comprimento da sépala
- Largura da pétala
- Largura da sépala



Compr Pétala	Compr Sépala	Largura Pétala	Largura Sépala	Tipo de Iris
5,1	3,5	1,4	0,2	setosa
4,9	3	1,4	0,2	setosa
4,7	3,2	1,3	0,2	setosa
...
7	3,2	4,7	1,4	versicolor
6,4	3,2	4,5	1,5	versicolor
6,9	3,1	4,9	1,5	versicolor
...
6,3	3,3	6	2,5	virginica
5,8	2,7	5,1	1,9	virginica
7,1	3	5,9	2,1	virginica
...

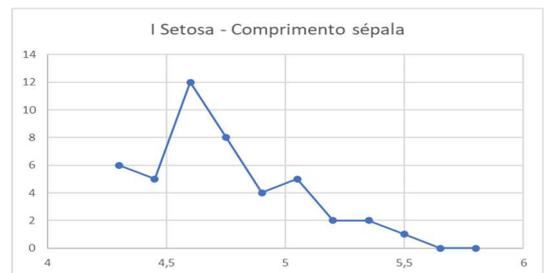
50 casos de cada tipo de íris

Íris Setosa



CP	CS	LP	LS
5,1	3,5	1,4	0,2
4,9	3	1,4	0,2
4,7	3,2	1,3	0,2
...

	Média	Desvio Padrão, Ruído
Comprimento da pétala	5.0	.35
Comprimento da sépala	3.4	.38
Largura da pétala	1.5	.17
Largura da sépala	.24	.11
Módulo	6.25	.48

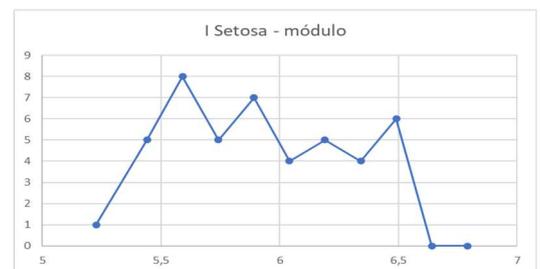


Centro da classe Setosa

$$\vec{m} = [5.0 \ 3.4 \ 1.5 \ .24]^t$$

Íris Versicolor

Íris Virgínica



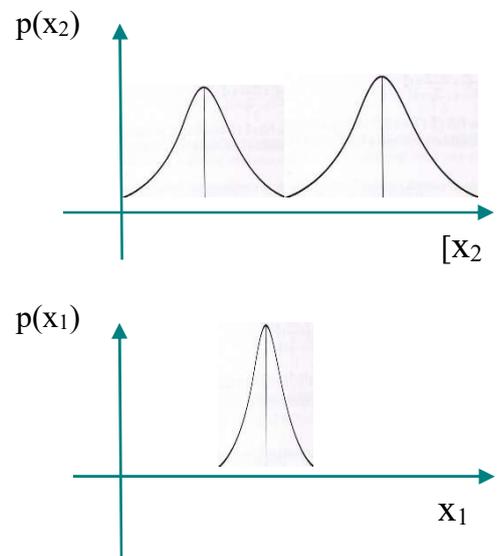
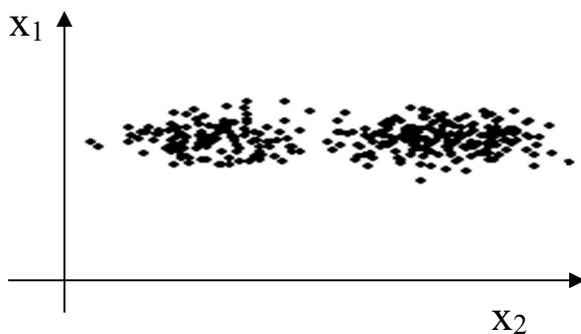
2 - Caso não supervisionado

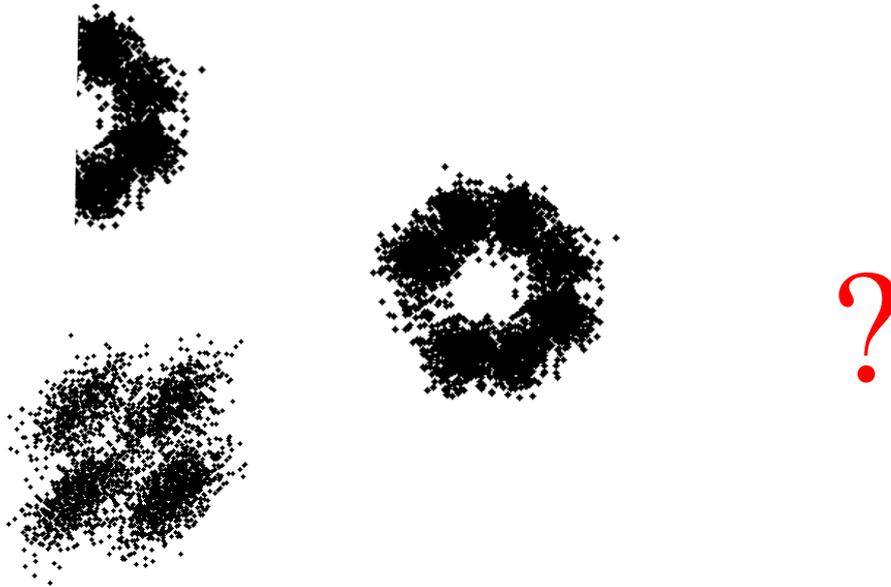
As classes de cada entrada (elemento) \underline{x} **não** são conhecidas

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.

Íris – não supervisionada

Comprimento da sépala

Comprimento da pétala

Largura da sépala

Largura da pétala

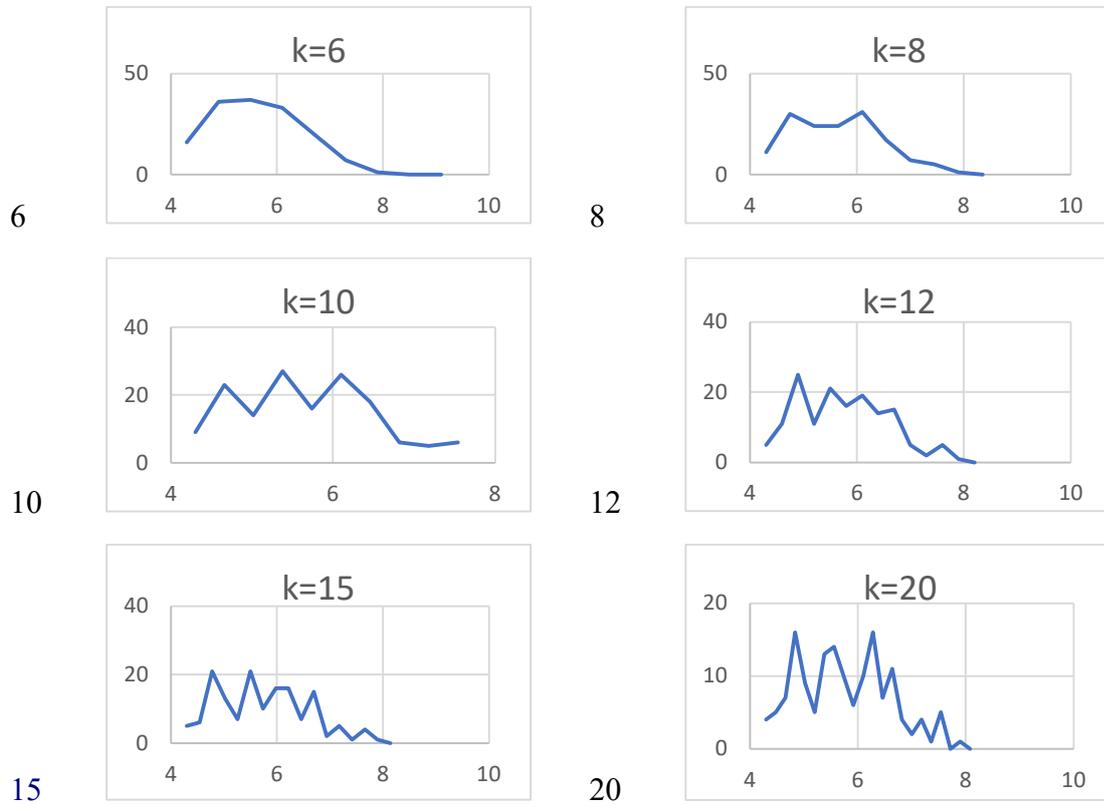
Compr Pétala	Compr Sépala	Largura Pétala	Largura Sépala	Tipo de Íris
5,1	3,5	1,4	0,2	?
4,9	3	1,4	0,2	?
4,7	3,2	1,3	0,2	?
7	3,2	4,7	1,4	?
6,4	3,2	4,5	1,5	?
6,9	3,1	4,9	1,5	?
6,3	3,3	6	2,5	?
5,8	2,7	5,1	1,9	?
7,1	3	5,9	2,1	?
...



150 casos (50 de cada tipo)

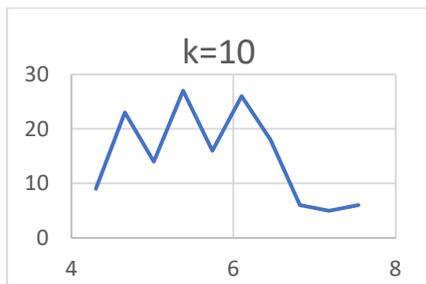
Histogramas do comprimento da pétala com 8, 10, 12, 15 e 20 intervalos

(ver seção 6 – Histogramas)



Histogramas das quatro variáveis k=10 intervalos

Comprimento da pétala



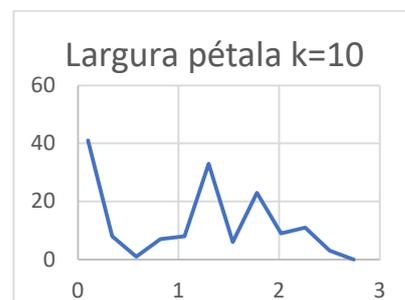
Largura da pétala



Comprimento da sépala

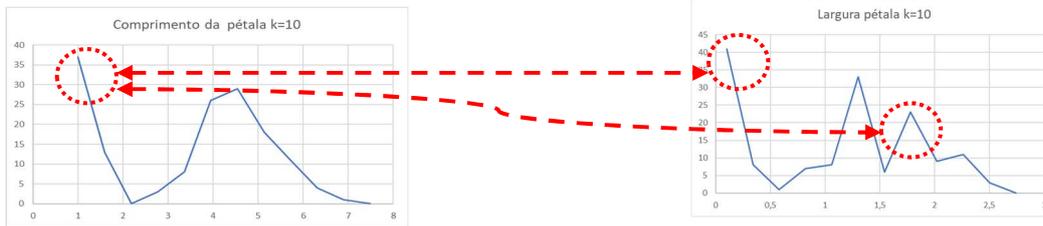


Largura da sépala

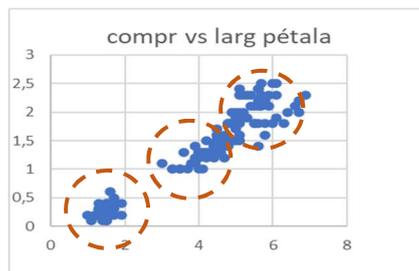


Correspondência entre modas (vetores baricentros) ?

Verificar onde são mapeados os valores de uma moda de uma componente:
na moda de outra componente ?



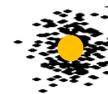
Plotar os valores das modas das variáveis duas à duas



3 – Ruído Branco – o caso ideal

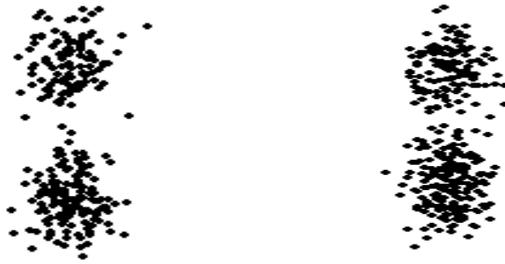
Os elementos \underline{x}_j de uma classe C_j são constituídos pelo padrão \underline{m}_i da classe, ao qual foi adicionado um ruído \underline{r}

$$\vec{x} \in C_j \quad \therefore \quad \vec{x} = \vec{m}_j + \vec{r} \quad \left. \begin{array}{l} r_1 \\ \vec{r} = \dots \\ r_n \end{array} \right\}$$



As componentes do ruído r_j em cada dimensão j

- são estocásticas,
- tem média zero,
- independem da classe
- tem distribuição gaussiana (ou aproximadamente gaussiana).
- tem a mesma variância
- são não correlatas

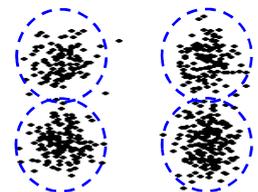


Obs 1: O tratamento de alguns casos mais genéricos (branqueamento do ruído) será visto logo adiante.

Obs 2: Se necessário uma classe pode ser considerada como constituída por diversas subclasses, às quais se aplicam as hipóteses acima.



No caso de ruído branco o separador ideal para as classes são esferas, que são também o separador natural por similaridade dos processos não supervisionados.

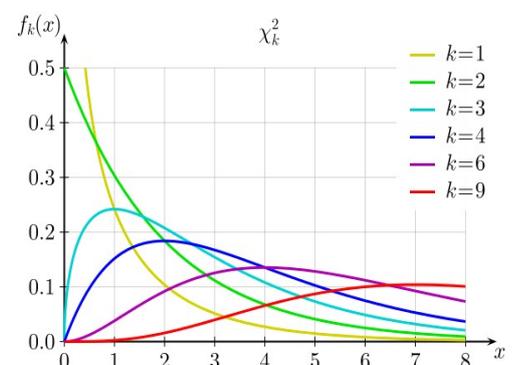


Se o ruído em cada dimensão é gaussiano com $\sigma = 1$ então módulo do ruído total tem distribuição chi-quadrado (ver seção 5 - Distribuições).

$$|\vec{r}|^2 = \sum_{i=1}^n r_i^2 = r_1^2 + r_2^2 + r_3^2 + \dots + r_n^2$$

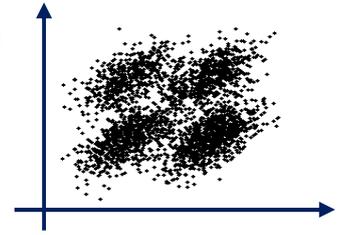
Em consequência o valor do raio r_0 da classe deve ser escolhido baseado nesta distribuição.

Tentaremos chegar próximo a esta condição através do branqueamento do ruído.



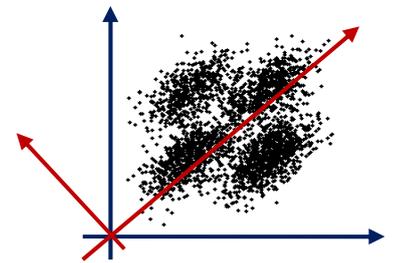
4 - Branqueamento do Ruído – gerando classes esféricas *classes sphereing*

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 é obtida por uma rotação na base dos dados. Isto pode ser obtido 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!).



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

Calculamos a matriz \underline{B} das PCA de \underline{r} . Utilizando \underline{B} determinamos o mapeamento de cada ruído \underline{r} em sua PCA \underline{p}

$$\vec{p} = \underline{B} \vec{r} \quad \text{onde} \quad [B] = \begin{bmatrix} \vec{b}_1 \\ \dots \\ \vec{b}_m \end{bmatrix} \quad e \quad p_i = \vec{b}_i \vec{r}$$

(obs: \vec{b}_i são vetores linha)

e a variância de cada componente p_i

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

A segunda etapa deve ser a equalização da potência do ruído nas diversas dimensões. A matriz de pesos $\underline{\sigma}^{-1}$ para normalização das potências dos ruídos será

$$\underline{\sigma}^{-1} = \text{diag}\left(\frac{1}{\sigma_i}\right)$$

Para a classe C_j os novos dados à classificar serão

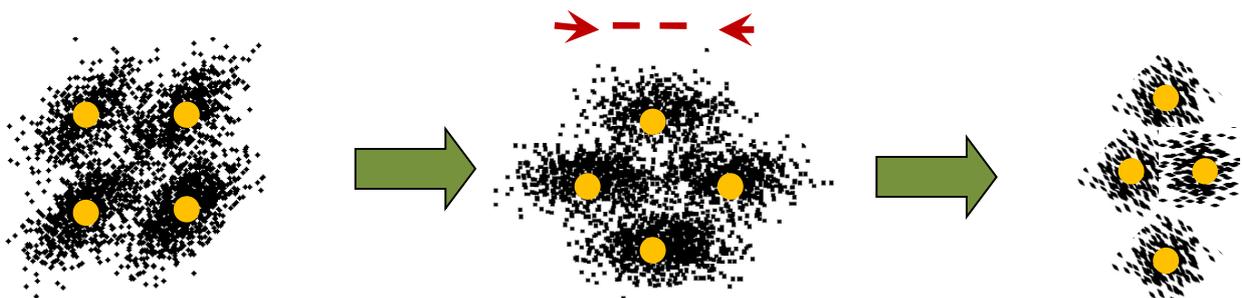
$$\vec{z} = \underline{\sigma}^{-1} \underline{B} \vec{x} = \underline{N} \vec{x} = \underline{N} \vec{m}_j + \underline{N} \vec{r} = \vec{m}_{j\text{ novo}} + \vec{r}_{\text{ novo}}$$

$$\text{onde } [N] = \underline{\sigma}^{-1} \underline{B} = \dots = \dots \begin{bmatrix} \vec{n}_1 & \frac{1}{\sigma_1} \vec{b}_1 \\ \vdots & \vdots \\ \vec{n}_m & \frac{1}{\sigma_m} \vec{b}_m \end{bmatrix} \quad e \quad r_{\text{ novo}} = \vec{n}_i^t \vec{r} \quad e \quad m_{\text{ novo}} = \vec{n}_i^t \vec{m}_i$$

(obs: \vec{n}_i são vetores linha)

As classes passam a ter um novo padrão $\vec{m}_{j\text{ novo}}$ adicionado de um ruído branco $\vec{r}_{\text{ novo}}$ cujas componentes $r_{j\text{ novo}}$ são não correlatas e apresentam média nula e desvio padrão unitário.

$$\vec{m}_{j\text{ novo}} = \underline{N} \vec{m}_j \quad \vec{r}_{\text{ novo}} = \underline{N} \vec{r} \quad \sigma(\vec{r}_{\text{ novo}}) = \vec{1} = [1 \ 1 \ 1 \dots \ 1]^t$$



4.1 Algumas observações:

- A multiplicação por $\underline{\mathbf{B}}$ representa uma rotação e por isto não altera a posição relativa das classes, como será visto adiante. Mas a multiplicação por $\underline{\sigma}^{-1}$ altera.
- Como $\underline{\mathbf{B}}$ é ortonormal (matriz rotação) e $\underline{\sigma}^{-1}$ é diagonal o conhecimento da matriz branqueadora $\underline{\mathbf{N}}$ permite determinar $\underline{\mathbf{B}}$ e $\underline{\sigma}^{-1}$

$$\underline{\mathbf{B}} = \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \dots \\ \vec{b}_n \end{bmatrix} \quad \vec{b}_i \vec{b}_j^t = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

$$\vec{n}_i \vec{n}_i^t = (\sigma_i^{-1} \vec{b}_i) (\sigma_i^{-1} \vec{b}_i^t) = \sigma_i^{-2} \quad e \quad \vec{b}_i = \sigma_i \vec{n}_i$$

- O uso de PCA sugere a eliminação de componentes pouco relevantes, i.e., com pequenos

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

e a conseqüente 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

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

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

- A multiplicação por \underline{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

$$\underline{B} = \underline{I}$$

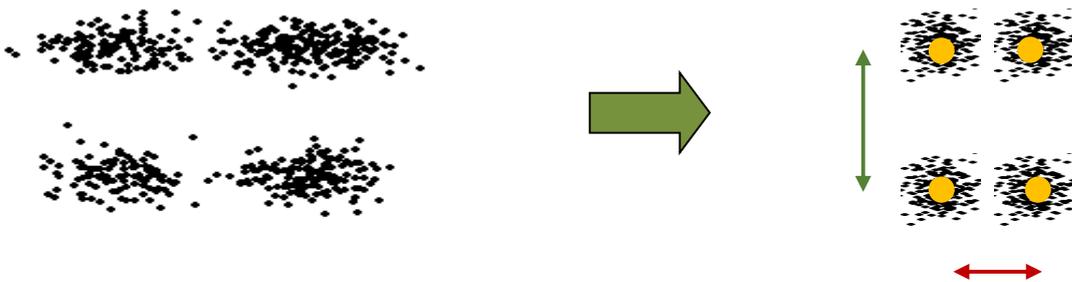


E a dispersão em cada dimensão i é dada por

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

Os novos dados à classificar serão

$$\vec{z} = \underline{\sigma}^{-1} \vec{x} = \underline{\sigma}^{-1} \vec{m}_j + \underline{\sigma}^{-1} \vec{r} = \vec{m}_{j_{novo}} + \vec{r}_{novo}$$



Branqueando o ruído: retornando ao estudo da Íris

Cálculo dos centros de classe e desvios padrão por componente

	X_1	X_2	X_3	X_4		σ_1	σ_2	σ_3	σ_4
	CP	CS	LP	LS		σ_{cp}	σ_{cs}	σ_{lp}	σ_{ls}
Setosa	5,0	3,4	1,5	0,24		0,35	0,38	0,17	0,11
Versicolor	5,9	2,8	4,3	1,33		0,51	0,31	0,47	0,20
Virginica	6,6	3,0	5,6	2,01		0,63	0,32	0,55	0,27
média						.53	.34	.40	.19



Hipóteses:

- A distribuição dos valores por componente pode ser aceita como gaussiana para todas componentes e para todas as classes.
- Consideraremos que não há correlação significativa entre o ruído nas diversas componentes.
- Consideraremos os desvios padrões do ruído como o valor médios dos desvios calculados para as classes, apresentados na tabela acima.

Assumindo o valor médio de σ_i para cada componente X_i e normalizando os dados (e os centros) para obter um desvio unitário por classe para cada componente. i.e. fazendo

$$x_i = X_i / \sigma_i$$

Os novos centros serão

	x_1	x_2	x_3	x_4
	CP	CS	LP	LS
Setosa	9.4	10.0	3.8	1.3
Versicolor	11.1	8.2	10.8	7.0
Virginica	12.5	8.8	14.0	10.6

E a distância entre os centros normalizados será

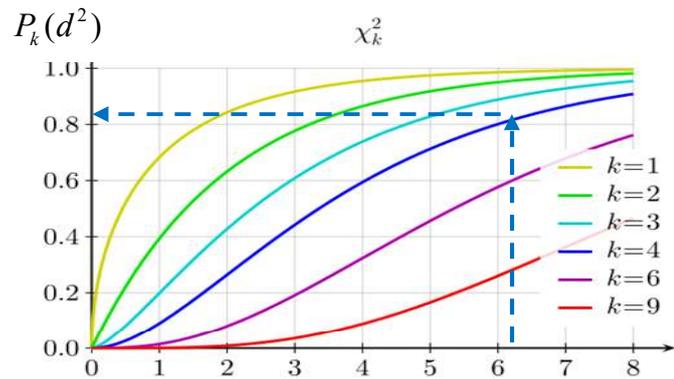
	Setosa	Versicolor	Virgínica
Setosa	0	9,4	14,2
Versicolor		0	5,0
Virgínica			0

Os dois centros mais próximos (Virgínica e Versicolor) distam 5,0.

Escolhendo um raio de similaridade mínima

$$r_0 \approx d_{\min} / 2 = 2.5 \quad d^2 = r_0^2 = 6.25$$

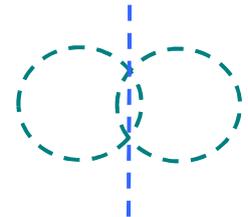
Para $k=4$ e $d^2 = 6.25$ das curvas (ou tabelas) verificamos que cerca de 85% dos elementos das classes serão capturados pelas esferas de similaridade mínima com $r_0 = 2.5$



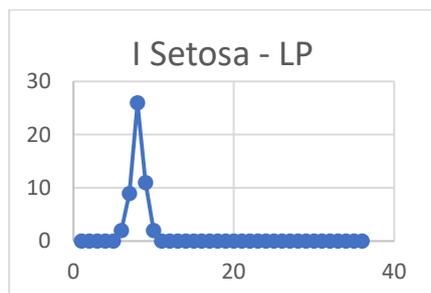
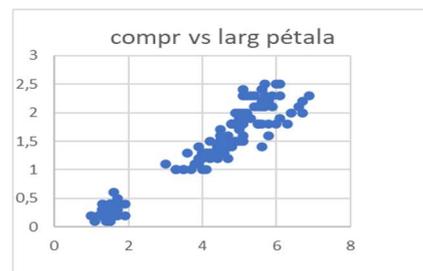
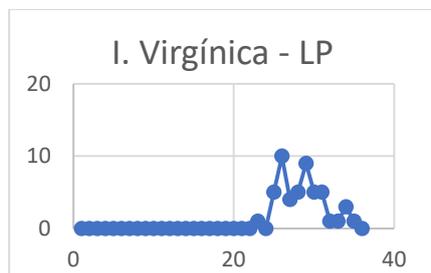
Obs:

1 - Podemos aumentar o raio de similaridade mínima para aumentar este percentual, mas isto aumentará a sobreposição do domínio das classes virgínica e versicolor

2 - Neste caso é necessário usar também o critério de máxima similaridade (centro mais próximo da entrada)



3 - Devido a grande dispersão entre os σ das diferentes classes é recomendável verificar a possibilidade de bipartir classes ou usar raios de similaridade diferentes por classe (a ser visto futuramente)



5. Algumas distribuições importantes:

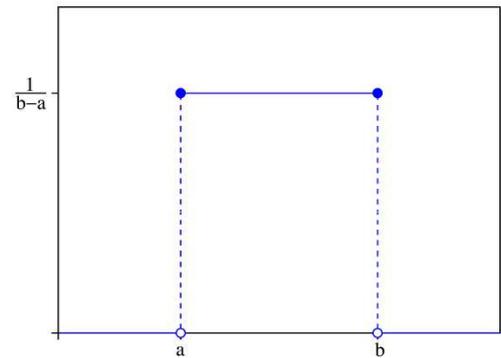
5.1 Distribuição uniforme

É a distribuição mais simples, a probabilidade de geração de um número em um intervalo $[a, b]$ é constante e independe do número. É bastante usada na geração de números aleatórios em programas computacionais, como por exemplo para a geração dos valores iniciais aleatórios das sinapses de uma rede neural

Função de distribuição (ou densidade) de probabilidade, f.d.p.

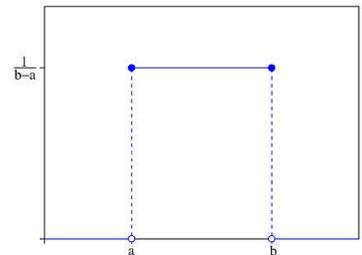
$$p(x) = \begin{cases} \frac{1}{b-a} & \text{se } a \leq x \leq b \\ 0 & \text{se } x < a \text{ ou } x > b \end{cases}$$

$$\mu = \frac{a+b}{2} \quad \sigma = \frac{b-a}{2\sqrt{3}}$$



No caso da geração de valores iniciais das sinapses

$$w \in [-a, +a] \quad \mu = 0 \quad \sigma = \frac{a}{\sqrt{3}}$$

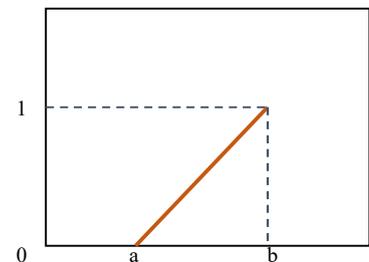


Função erro (uma das definições)

Função de distribuição de probabilidade acumulada

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

$$= \begin{cases} 0 & \text{se } r_0 < a \\ \frac{r_0 - a}{b - a} & \text{se } a \leq r_0 \leq b \\ 1 & \text{se } r_0 > b \end{cases}$$



5.2 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.

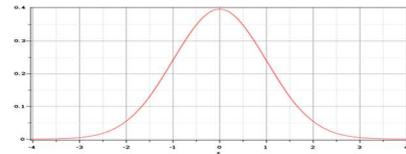
Em nosso caso consideramos que os elementos \underline{x} de uma classe C são compostos pelo padrão \underline{m} da classe adicionado de ruído branco \underline{r} .

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

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

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

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

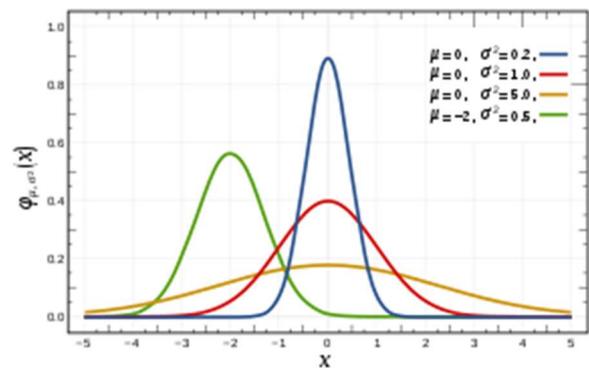


Média = mediana = moda

$$\mu = 0$$

Desvio padrão

$$\sigma$$



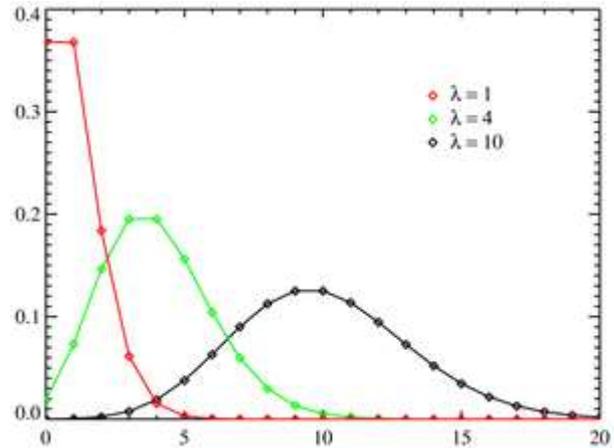
5.3 Distribuição de Poisson

Fenômenos naturais que podem assumir apenas valores inteiros e não negativos usualmente seguem a distribuição de Poisson. São exemplos o número de partículas emitidas por uma fonte radioativa por segundo, o número de pessoas que entram em uma determinada loja por dia, etc.

$$p(x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$\mu \quad \sigma = \sqrt{\mu}$$

As funções de densidade simples e acumulada existem tabeladas.



A distribuição de Poisson tende para a Gaussiana quando a média é elevada.

5.4 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 d do padrão da classe \underline{m} , isto é, permite estimar a probabilidade da similaridade de uma entrada com seu padrão.

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

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

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

r_i - distribuição Gaussiana com $\mu = 0$ e $\sigma = 1$, não correlatos $\forall i = 1, \dots, k$

$d^2 = \chi_k^2$ - distribuição Chi-quadrado com k graus de liberdade

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

em cada dimensão $i = 1, \dots, k$ de \vec{x} $d_i^2 = r_i^2$

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

Gaussiana (x^2)

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

distribuição χ_k^2 com k graus de liberdade

Função densidade de probabilidade χ_k^2

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

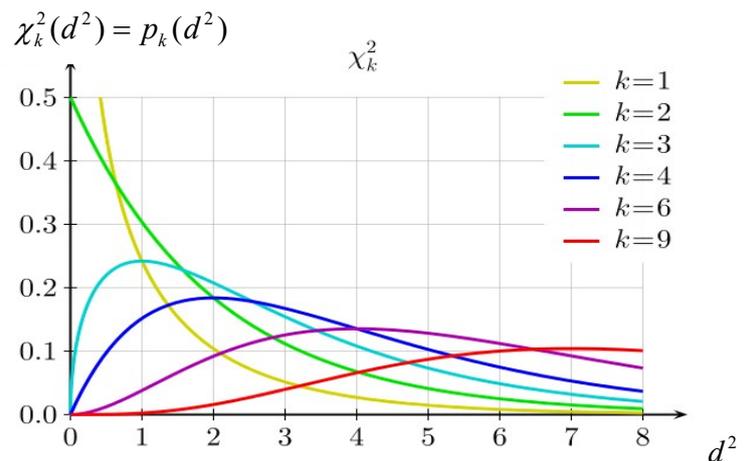
$$\Gamma(k/2) = \int_0^{\infty} x^{k/2-1} e^{-x} dx$$

média = k

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

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

variância = $2k$

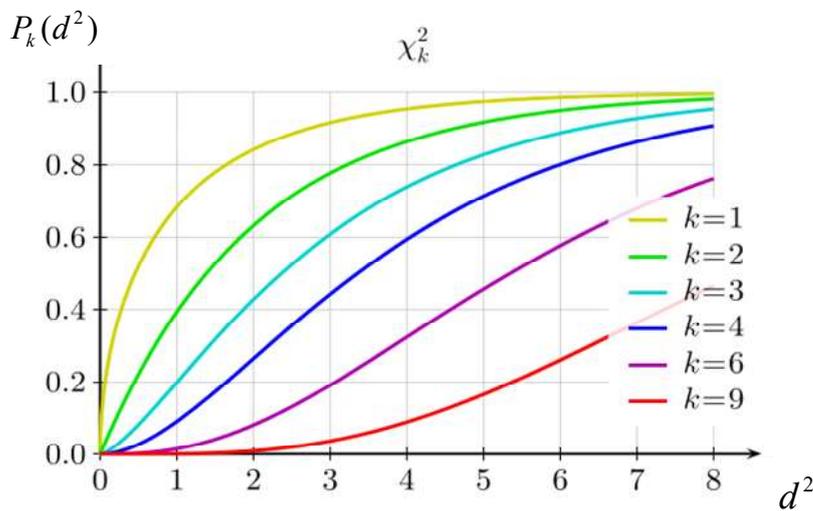


Função erro (uma das definições) de χ_k^2

Função de distribuição de probabilidade acumulada

$$P_k(d^2) = erf(d^2) = p(x | x < d^2) = \int_{-\infty}^{d^2} p_k(t) dt = \frac{1}{\Gamma(k/2)} \gamma\left(\frac{k}{2}, \frac{d^2}{2}\right)$$

tabelada para diversos k



Distribuição Chi quadrado acumulada

Tabela simplificada

Fornece o quantil $x = d^2$ em função da dimensão (g.l.) k (linhas) e de $P(d^2) = P(x < d^2)$ (colunas).

$P(d^2)$.5	.75	.9	.95	.99
k					
1	.46	1.32	2.71	3.84	6.64
2	1.39	2.77	4.61	5.99	9.21
3	2.37	4.11	6.25	7.82	11.35
4	3.36	5.39	7.78	9.49	13.28
5	4.35	6.63	9.24	11.07	15.09
6	5.35	7.84	10.65	12.59	16.81
7	6.35	9.04	12.02	14.07	18.46
8	7.34	10.22	13.36	15.51	20.09
9	8.34	11.39	14.68	16.92	21.67
10	9.34	12.55	15.99	18.31	23.21

TABLE 49 PERCENTILE VALUES (χ_p^2) FOR THE CHI-SQUARE DISTRIBUTION with n degrees of freedom (shaded area = p)

n	$\chi_{.995}^2$	$\chi_{.99}^2$	$\chi_{.975}^2$	$\chi_{.95}^2$	$\chi_{.9}^2$	$\chi_{.8}^2$	$\chi_{.7}^2$	$\chi_{.6}^2$	$\chi_{.5}^2$	$\chi_{.4}^2$	$\chi_{.3}^2$	$\chi_{.2}^2$	$\chi_{.1}^2$	$\chi_{.05}^2$	$\chi_{.025}^2$	$\chi_{.01}^2$	$\chi_{.005}^2$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000				
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.735	.211	.103	.0506	.0201	.0100				
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.322	.216	.115	.072				
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207				
5	16.7	15.1	12.8	11.1	9.24	6.63	4.95	2.67	1.61	1.15	.881	.554	.412				
6	18.5	16.8	14.4	12.6	10.6	7.84	5.85	3.45	2.20	1.54	1.24	.872	.676				
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	1.77	1.24	.989	.789				
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.18	1.65	1.34	.917				
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73				
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16				
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60				
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07				
13	29.8	27.7	24.7	22.4	19.8	15.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57				
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07				
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60				
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14				
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70				
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26				
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84				
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43				
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03				
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64				
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26				
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89				
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5				
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2				
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8				
28	51.0	48.3	44.5	41.3	37.9	32.6	27.2	22.7	18.9	16.9	15.3	13.6	12.5				
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1				
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8				
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7				
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0				
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	45.5	43.2	40.5	37.5	35.5				
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3				
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2				
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2				
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3				

Source: Catherine M. Thompson, Table of percentage points of the χ^2 distribution, Biometrika, Vol. 32 (1941), by permission of the author and publisher.

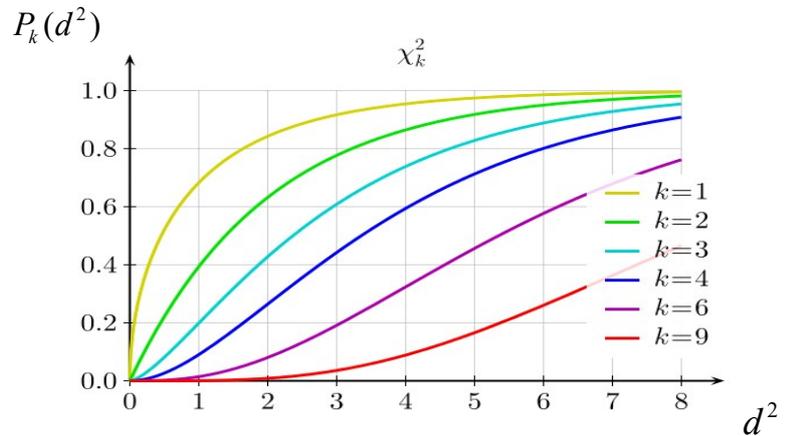
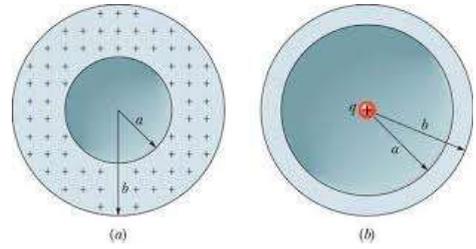
Densidade populacional local

% da população em uma casca de raio r / volume da casca

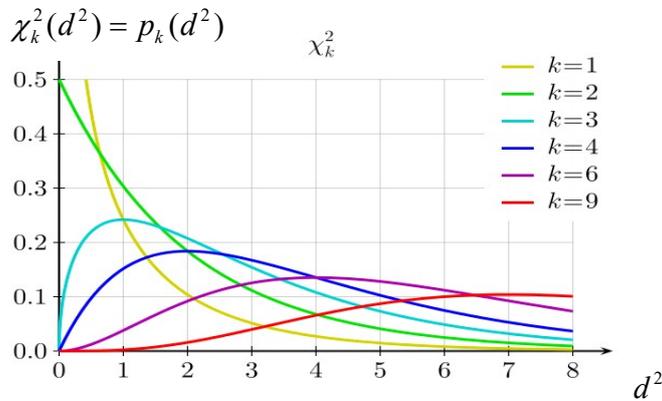
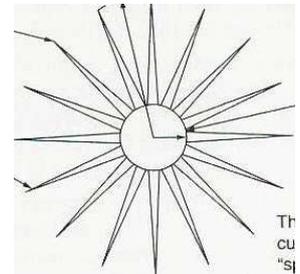
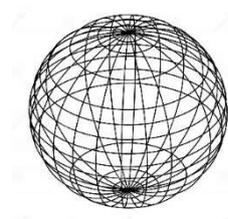
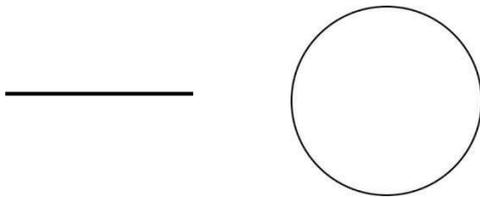
$$\frac{\Delta P_k(r)}{\Delta V_k(r)} = \frac{dP_k(r)}{dr} / \frac{dV_k(r)}{dr} =$$

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

$$= \frac{f(k)}{r^{k/2} e^{r/2}}$$



Geometria multidimensional



**o problema da multidimensionalidade:
quanto maior a dimensão mais difícil um agrupamento eficaz !**

6. Histogramas - heurísticas:

n – número de pontos (eventos)

k – número de barras (intervalos)

Algumas heurísticas para determinar o número k de intervalos à utilizar

$$k = \sqrt{n} \qquad n > 100$$

$$k = 1 + \lg_2 n \simeq 1 + 3.3 \lg_{10} n$$