Authentication

Prof. Dr. Helene Dörksen

helene.doerksen@th-owl.de

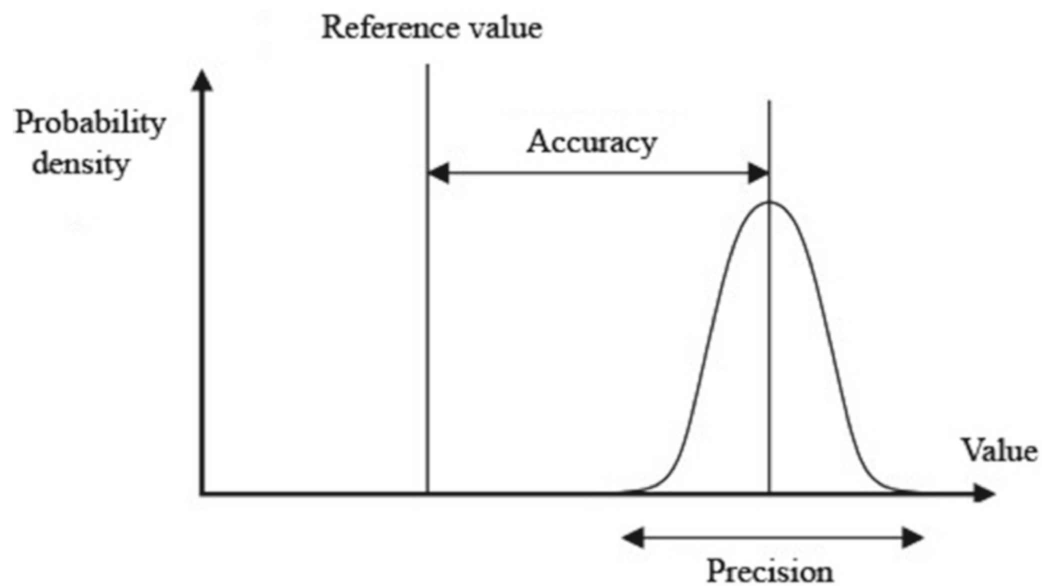
Lecture 3:

Statistical Pattern Recognition

Measured Data and Measurement Errors

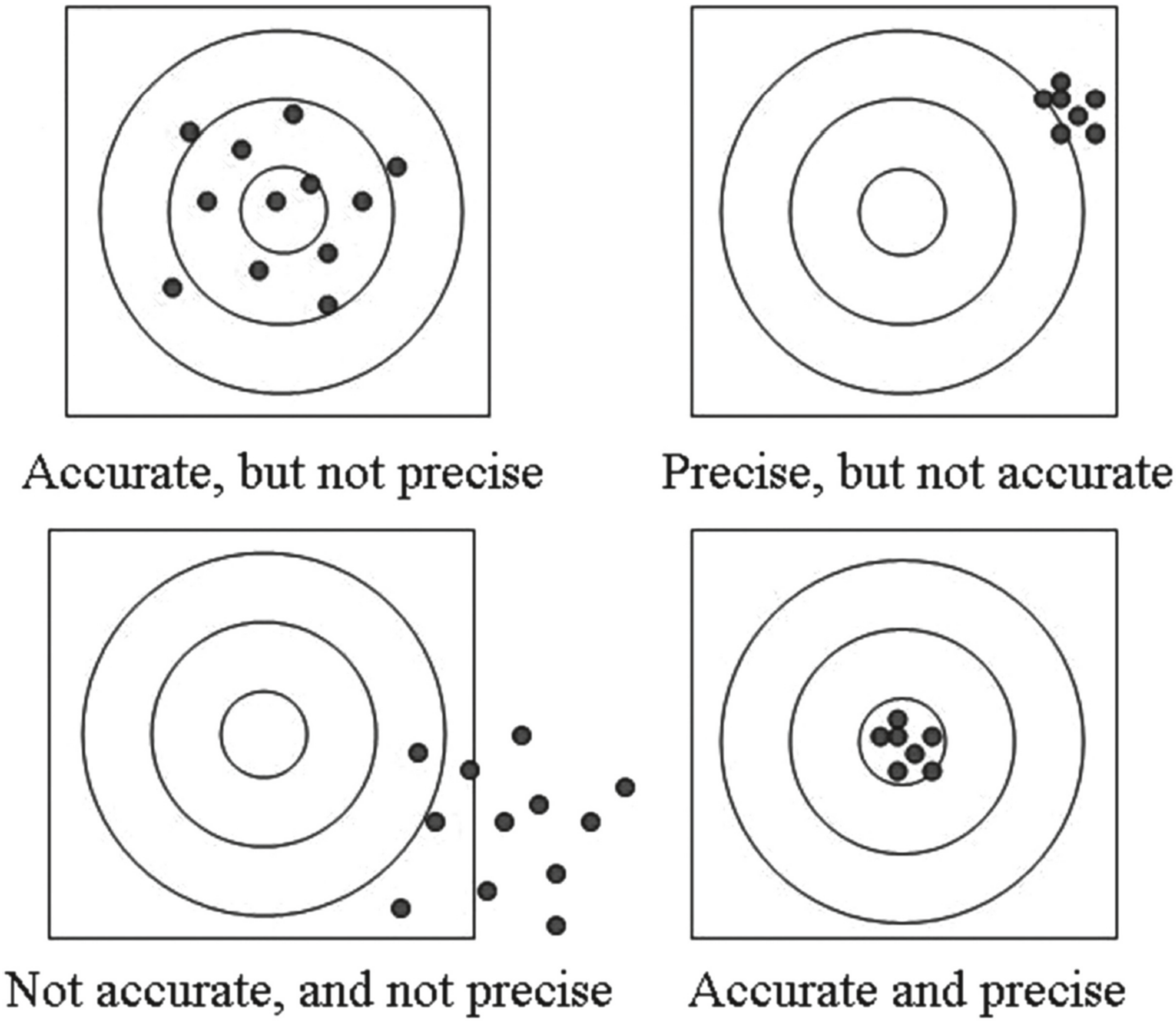
Fact is:

- data is not perfect
- data comes with a margin of error or uncertainty



© book: Dougherty, G. Pattern Recognition and Classification. Springer, 2013

Measured Data and Measurement Errors



Measured Data and Measurement Errors

How to handle errors in data?

Probability theory helps us to model random error and is therefore a solid basis for classifier design

Simple Probability Theory

If $\boldsymbol{A, B, C, \ldots}$ are events, the probability of these events can be denoted by a real number between $\mathbf{0}$ and $\mathbf{1}$, i.e.

$$\boldsymbol{P(A), P(B), P(C)}$$

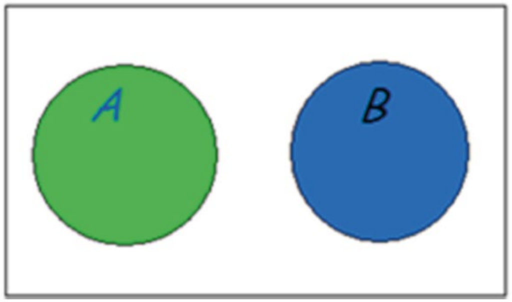
The probability is linked to the relative frequency of that event happening, i.e., an experiment is observed a large number of times $\boldsymbol{N}$, and if event $\boldsymbol{A}$ occurs $\boldsymbol{M}$ times then

$$\boldsymbol{P(A) = \frac{M}{N}}$$

Simple Probability Theory

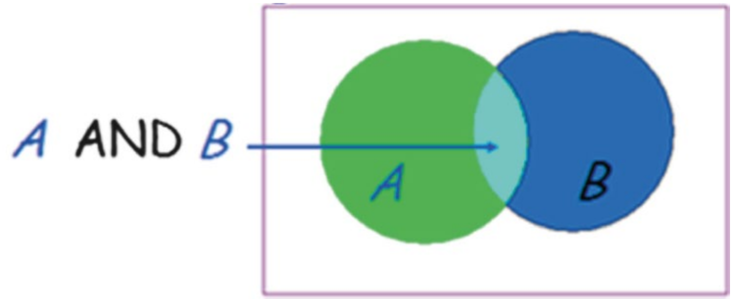
Probability rules can be easy illustrated by set-theoretical fundamentals:

A and B are (a) mutually exclusive (nonoverlapping)



$$P(A \text{ or } B) = P(A) + P(B)$$

A and B are (b) not mutually exclusive (overlapping)



$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Complement of the event: $$P(\bar{A}) = 1 - P(A)$$

Events, A and B, are **independent** if: $$P(A \text{ and } B) = P(A) \cdot P(B)$$

Simple Probability Theory

Probability information can be collected in

Contingency Table

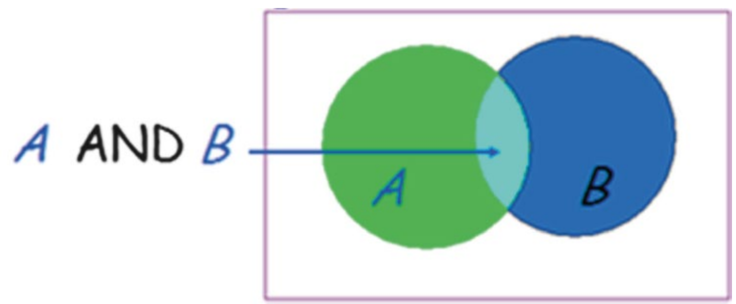
	Age (years)			
Sex	<30	30–45	>45	Total
Male (M)	60	20	40	120
Female (F)	40	30	10	80
Total	100	50	50	200

Simple Probability Theory

Conditional probability is the probability of some event A, given the occurrence of some other event B:

Main rule:

$$P(A|B) = P(A \text{ and } B)/P(B)$$



Other rules:

$$P(B|A) = P(A \text{ and } B)/P(A)$$

If the events A and B are statistically independent:

$$P(A|B) = P(A) \qquad P(B|A) = P(B)$$

Simple Probability Theory

Multiplicative rule:

$$P(A \text{ and } B) = P(A|B) \cdot P(B)$$
$$= P(B|A) \cdot P(A)$$

Bayes' Rule:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

P(A|B) is known as the posterior probability

$$\text{Posterior(probability)} = \frac{\text{likelihood} \times \text{prior(probability)}}{\text{evidence}}$$

Simple Probability Theory

$$\mathrm{P}\left(\textit{robot says defective} \mid \textit{defective}\right) = 0.98$$

$$\mathrm{P}\left(\textit{robot says defective} \mid \textit{not defective}\right) = 1 - 0.99 = 0.01$$

$$\mathrm{P}\left(\textit{defective}\right) = 0.001$$

$$\mathrm{P}\left(\textit{not defective}\right) = 1 - 0.001 = 0.999$$

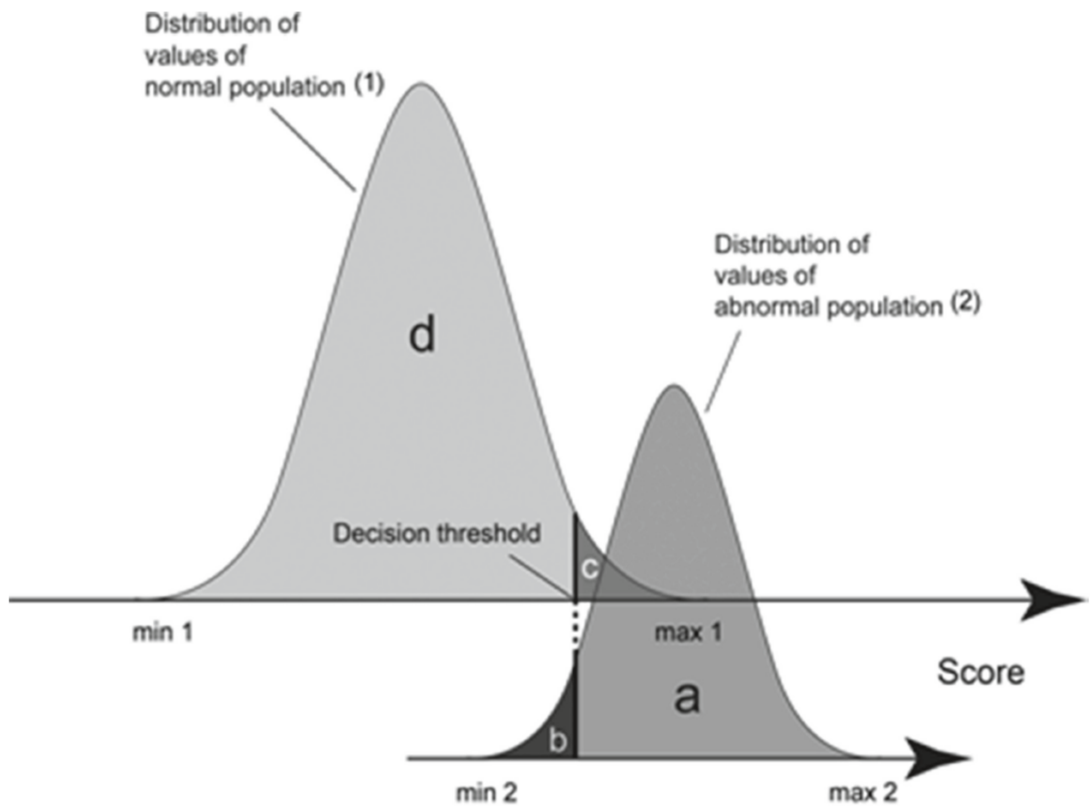
$$\begin{aligned}
&\mathrm{P}\left(\textit{robot says defective}\right) \\
&= \mathrm{P}\left(\textit{robot says defective} \mid \textit{defective}\right)\mathrm{P}\left(\textit{defective}\right) \\
&\quad + \mathrm{P}\left(\textit{robot says defective} \mid \textit{not defective}\right)\mathrm{P}\left(\textit{not defective}\right) \\
&= 0.98 \cdot 0.001 + 0.01 \cdot 0.999 \\
&= 0.00098 + 0.00999 \\
&= 0.01097
\end{aligned}$$

Bayes' Rule:

$$\left\{\begin{aligned}
&\mathrm{P}\left(\textit{defective} \mid \textit{robot says defective}\right) \\
&= \frac{\mathrm{P}\left(\textit{robot says defective} \mid \textit{defective}\right)\mathrm{P}\left(\textit{defective}\right)}{\mathrm{P}\left(\textit{robot says defective}\right)} \\
&= \frac{0.98 \cdot 0.001}{0.01097} \\
&= \frac{0.00098}{0.01097} \simeq 0.08933
\end{aligned}\right.$$

Simple Probability Theory

Diagnostic test **score** distributions for normal (1) and abnormal (2) population samples:



Corresponding contingency table:

	B	$\bar{B}$
A	a (TP)	b (FN)
$\bar{A}$	c (FP)	d (TN)

Simple Probability Theory

Actual	Predicted			
		B	$\bar{B}$	Sum
	A	$P(A \text{ and } B)$	$P(A \text{ and } \bar{B})$	$P(A)$
	$\bar{A}$	$P(\bar{A} \text{ and } B)$	$P(\bar{A} \text{ and } \bar{B})$	$P(\bar{A})$
		$P(B)$	$P(\bar{B})$	1

Simple Probability Theory

$$\text{sensitivity}, P(B|A) = \text{TP}/(\text{TP} + \text{FN}) = a/(a + b)$$

$$\text{specificity}, P(\bar{B}|\bar{A}) = \text{TN}/(\text{TN} + \text{FP}) = d/(d + c)$$

Naïve Bayes' Classifier

Assuming that the features (f_1, f_2, f_3, . . ., f_n) are independent:

Bayes' Classifier: $$P(C|f_1,f_2,f_3,\ldots,f_n) = \frac{P(C)P(f_1,f_2,f_3,\ldots,f_n|C)}{P(f_1,f_2,f_3,\ldots,f_n)}$$

Class, C, is dependent on several features (f_1, f_2, f_3, . . ., f_n)

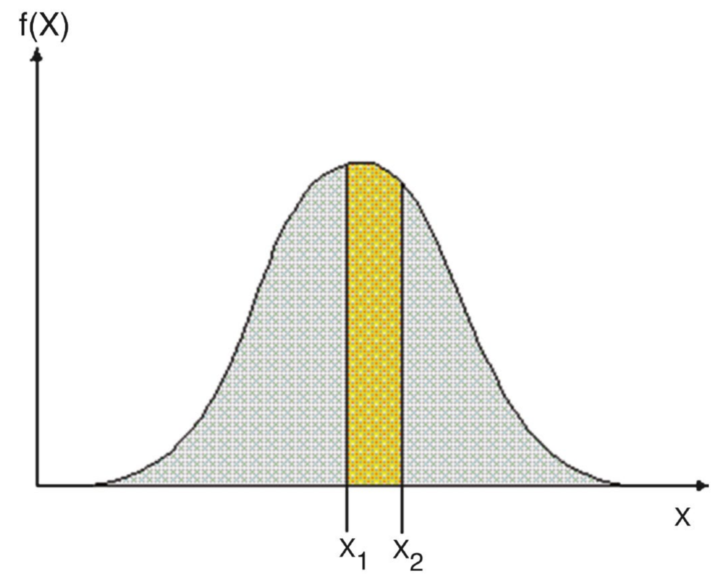
If the features are all independent, then

$$P(C|f_1,f_2,f_3,\ldots,f_n) \propto P(C)P(f_1|C)P(f_2|C)P(f_3|C)\ldots P(f_n|C)$$
$$\propto P(C)\prod_{i=1}^{n} P(f_i|C)$$

Advanced Probability Theory

Probability density function (**PDF**) of X:

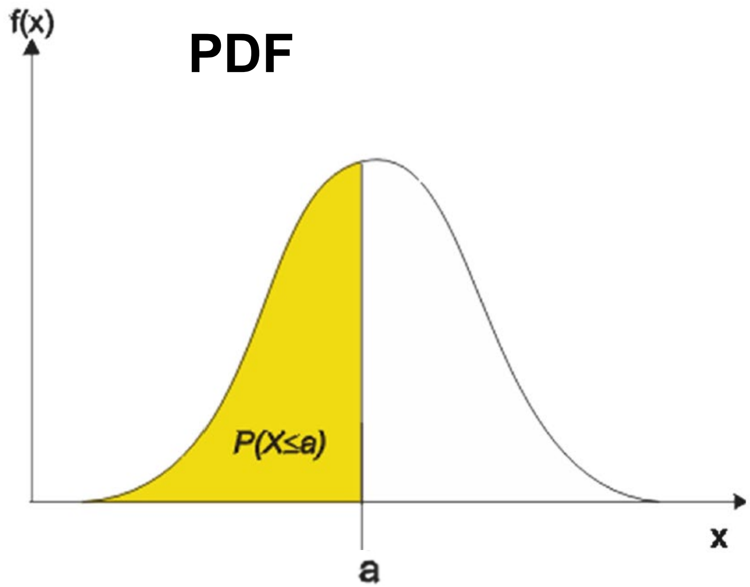
$$P(x_1 \leq X \leq x_2) = \int_a^b f(x)\mathrm{d}x$$

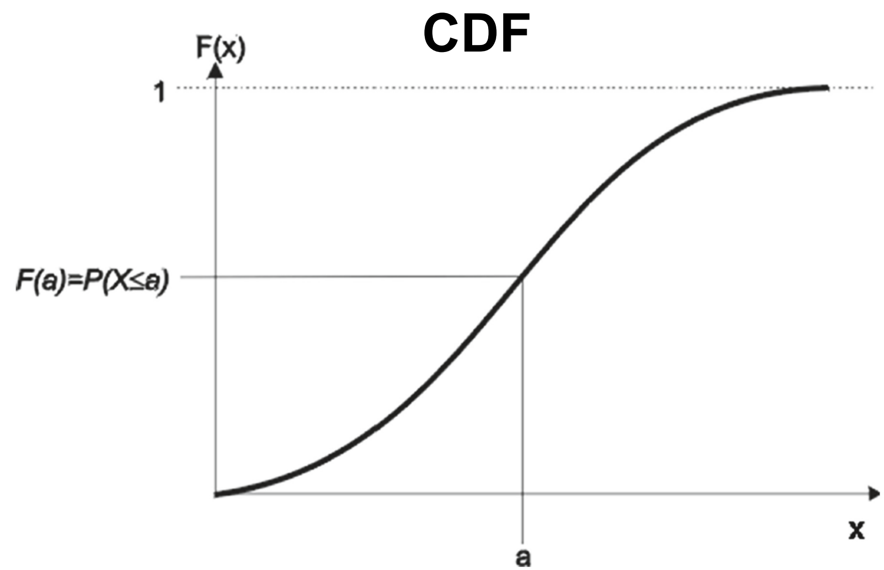


Cumulative distribution function (**CDF**) of X:

$$F(x) = P(X \leq x) = \int_0^x f(u)\mathrm{d}u$$

Advanced Probability Theory



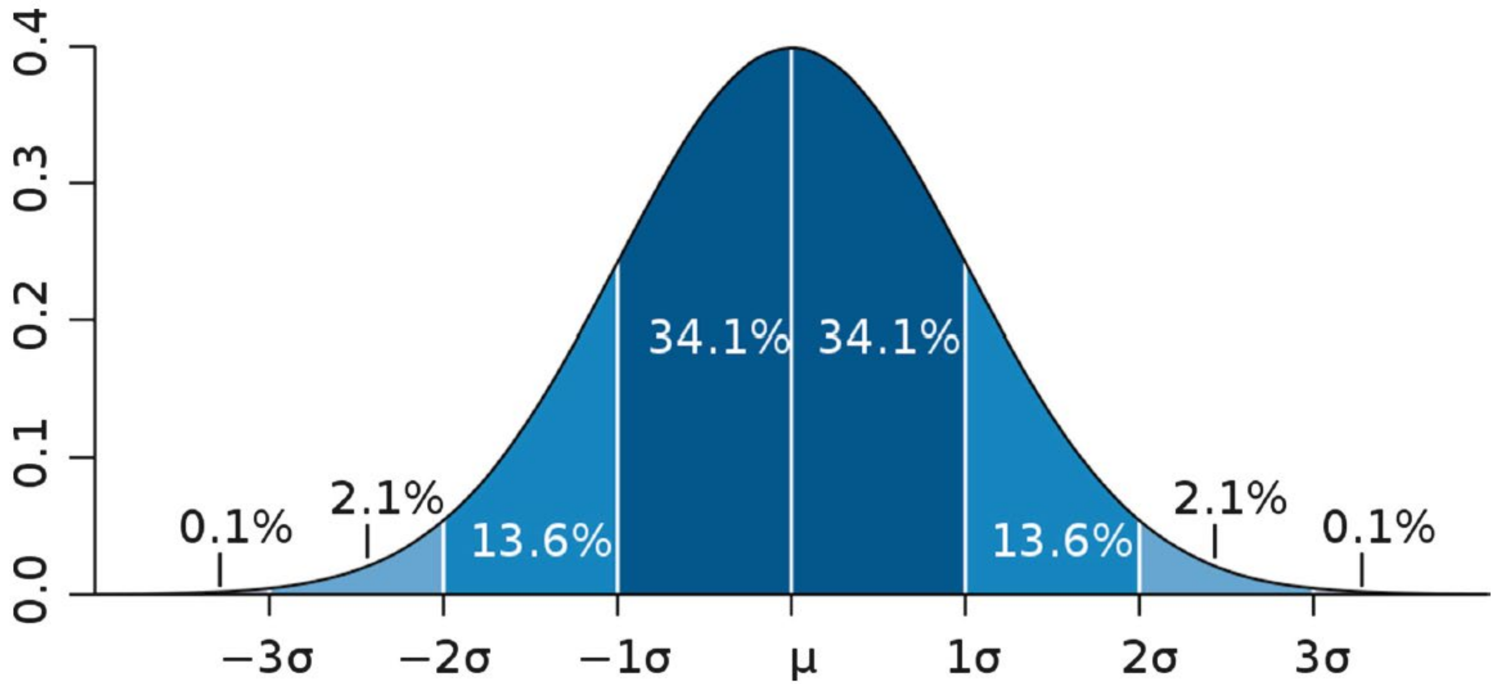


example of a PDF is the well-known normal (or Gaussian) distribution

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

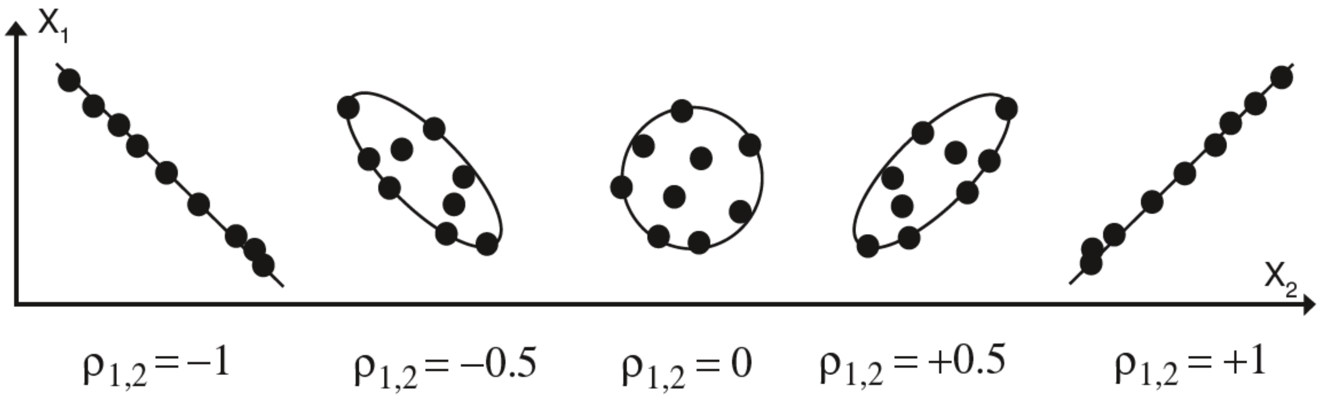
Advanced Probability Theory

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



normal (or Gaussian) distribution

Correlation Coefficients

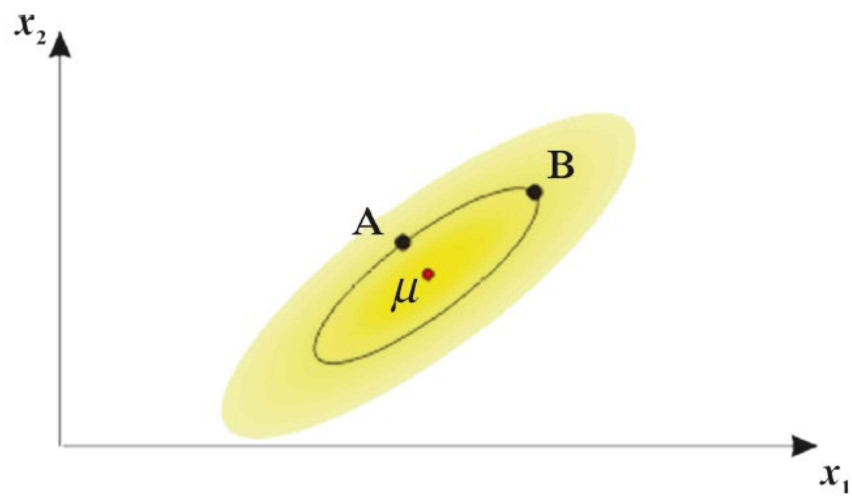


square of the correlation coefficient:
$\boldsymbol{r^2}$ called r-*squared* or *coefficient of determination*

Mahalanobis Distance

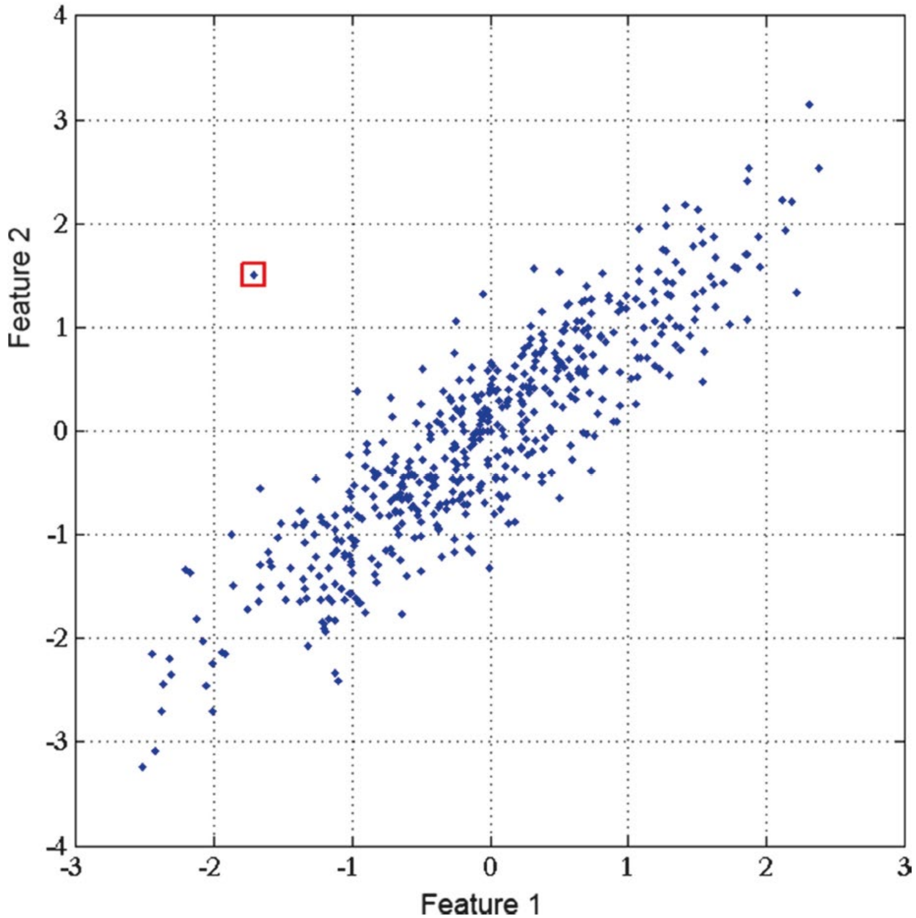
$$D_{\mathrm{m}}(\mathbf{x}, \mathbf{y}) = \mathrm{sqrt}((\mathbf{x} - \mathbf{y})\Sigma^{-1}(\mathbf{x} - \mathbf{y})^{\mathrm{T}})$$

$\boldsymbol{\Sigma}^{-1}$ is the inverse of the covariance matrix



A and B at the same Mahalanobis distance from the centroid μ

Mahalanobis Distance



Summary

- data is not perfect
- **probability theory** helps us to model error in the data

Bayes' Classifier

Probability density function (**PDF**)
Cumulative distribution function (**CDF**)
normal (or Gaussian) distribution

- Covariance matrices are related to correlation coefficients
- Mahalanobis distance is not always equivalent to Euclidean

Homework: Exercises and Labs

for the next week prepare practical exercises and labs from
Exercises Lec 3 (you will find it in the donwload area)