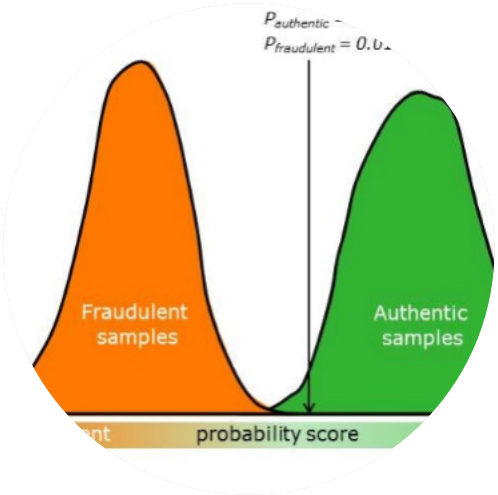
Validation of non-targeted methods in food fraud area

Martin Alewijn

RIKILT – Wageningen University and Research (NL)



WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

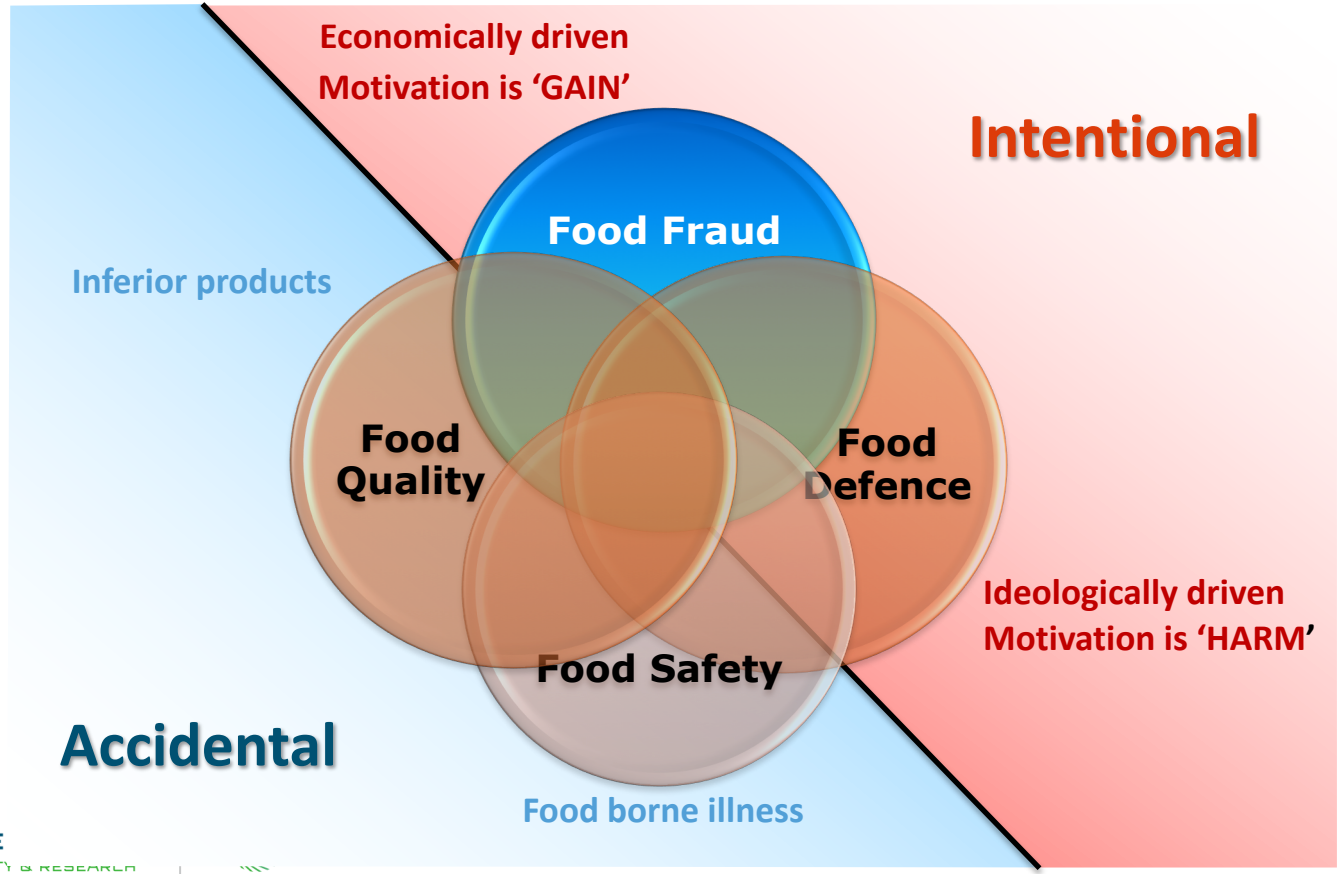
RIKILT

- Wageningen Food Safety Research, WFSR
- > 200 staff
- Clients: government, national/EU scientific funding bodies, industry, NGOs
- Research themes:
  - Natural and chemical contaminants
  - New risks
  - Residues
  - Feed
  - **Product composition / quality / authenticity**

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

Food fraud in perspective



Food fraud issues

- Ingredient authenticity:

  Replacing or diluting of the product (low value, species)

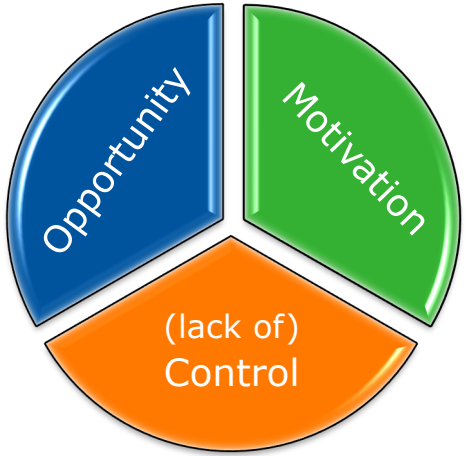
- Production systems:

  Not meeting corporate social responsibility issues (sustainable, organic, animal welfare, fair trade, etc.)

- Geographical origin
- Typicality

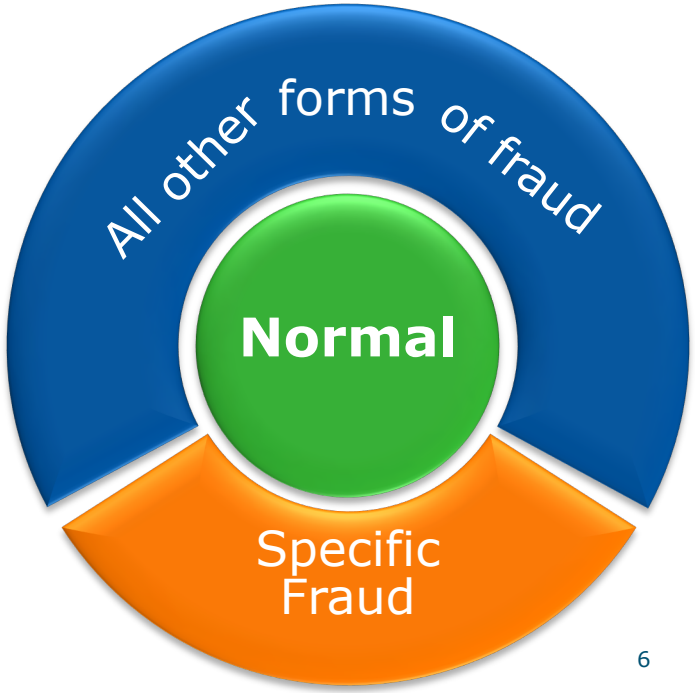
Food fraud detection tools

- Mass balance, tracking & tracing, blockchain,
- Databases (RASFF, Foodfraud.org,)
- Food fraud vulnerability self assessment
- Monitor global price changes, shortages
- Analyses!

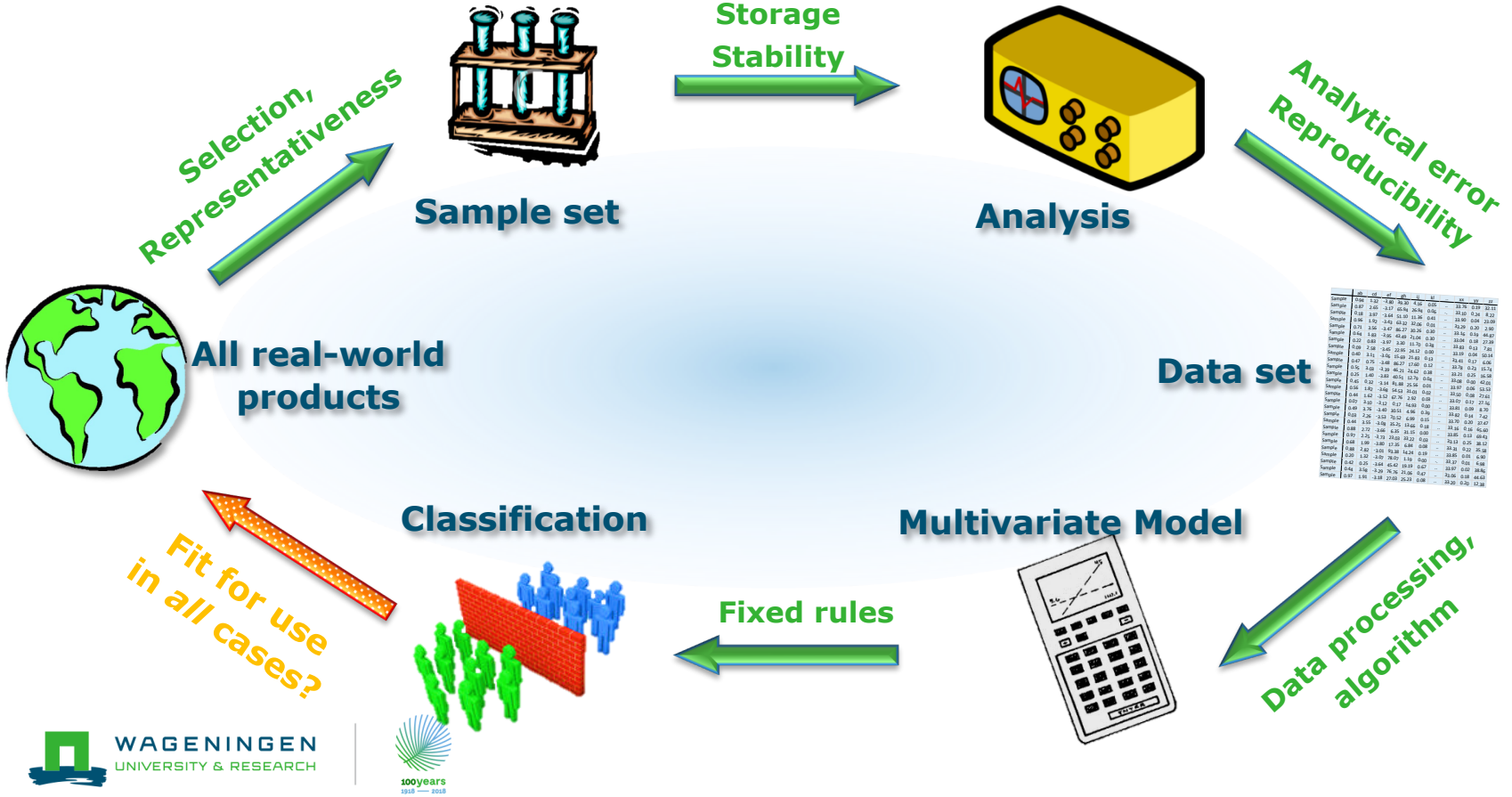


Food fraud analytical methods

- Targeted (moisture in meat, melamine in milk products, ...)
- Untargeted – fingerprinting
    - Chromatography
    - Non-chromatographic MS
    - Spectroscopy
    - Molecular techniques
    - etc.



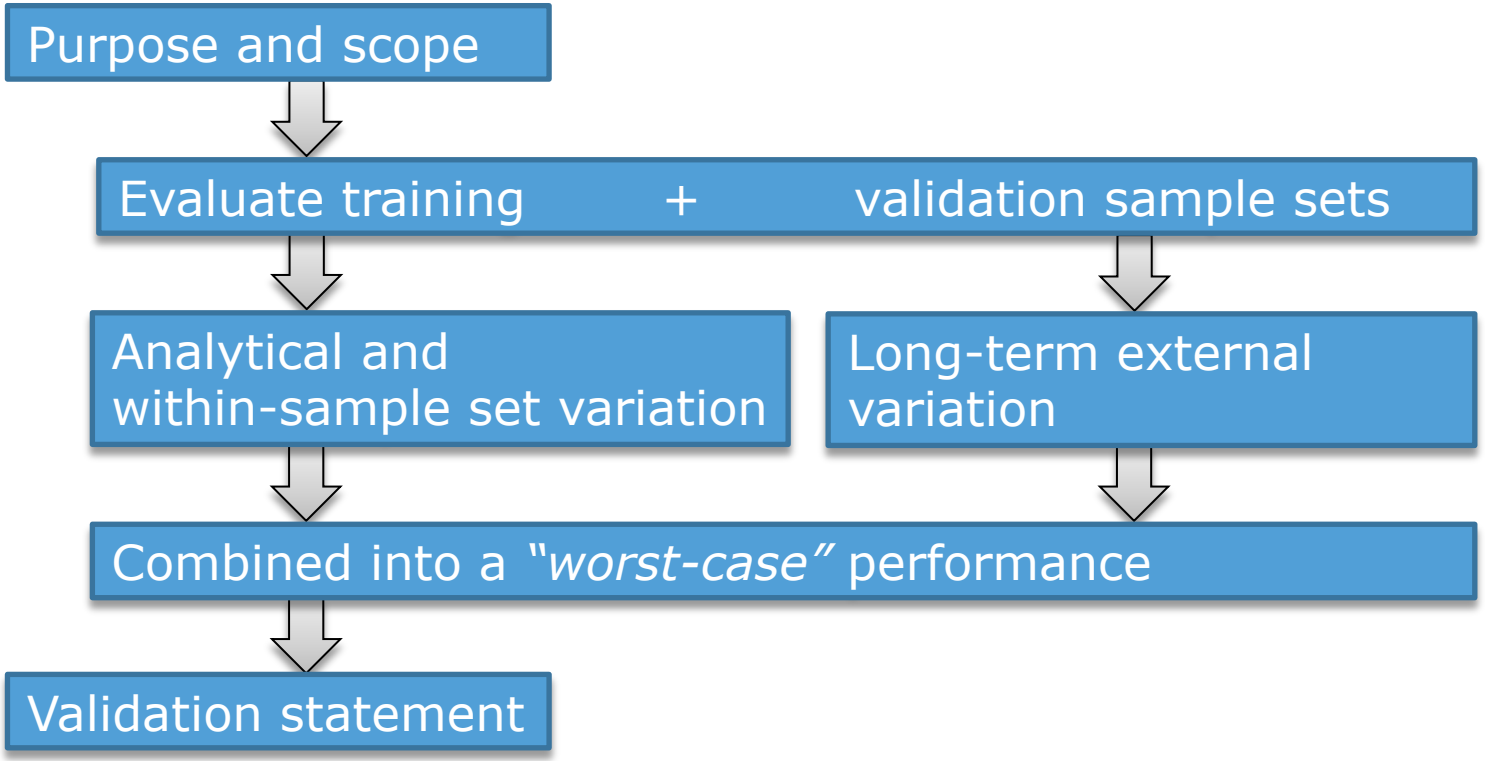
Building a classification model



Issues in validation

- Classification mechanism is usually indirect
- Use of a database of samples to predict future –*unseen*- samples!
- When is a database ‘sufficient’?
- How to quantify the certainty of a future result?

Suggested validation protocol *(binary classification)*



Sample sets: training and validation set

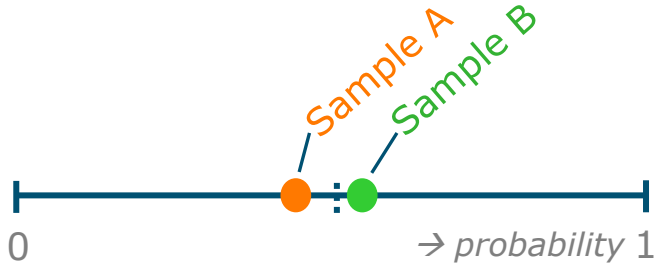
- Is sampling representative, true, and balanced for:
  - Target class
  - Relevant subgroups (season, variety, storage time, etc.)
- Is the sampling quantitatively sufficient?

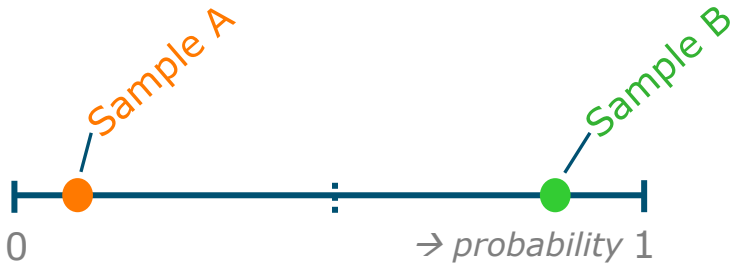
Validation set (system challenge) :

- Is all *additional* variation included?
  - Separate sample source, different time period, use of different equipment/technicians, etc.

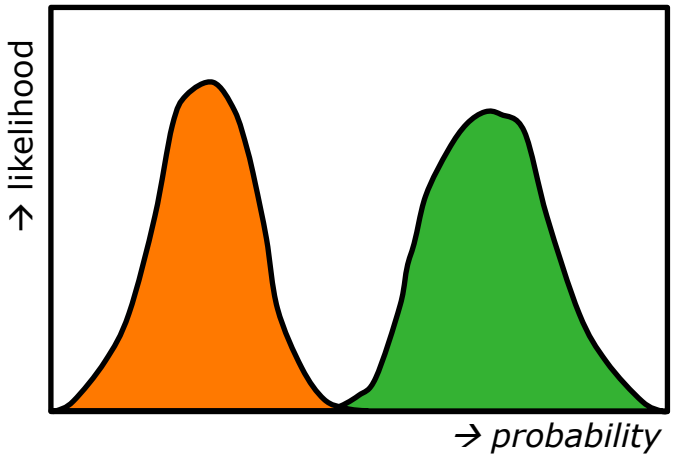
Performance evaluation

Two examples, the same classification result:



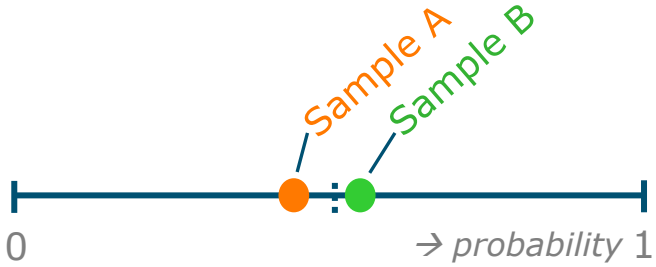


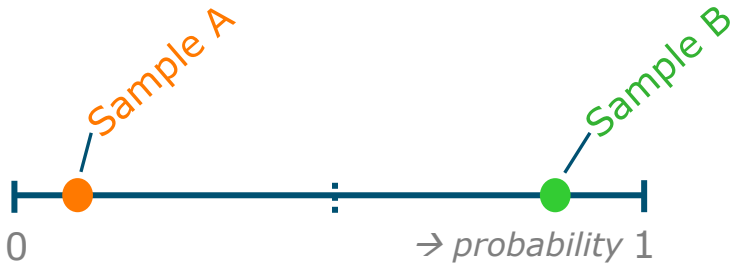
- During evaluation: **go probabilistic!**
- Every algorithm can be made probabilistic or class-distance-based
- Obtain model probability distributions



Performance evaluation

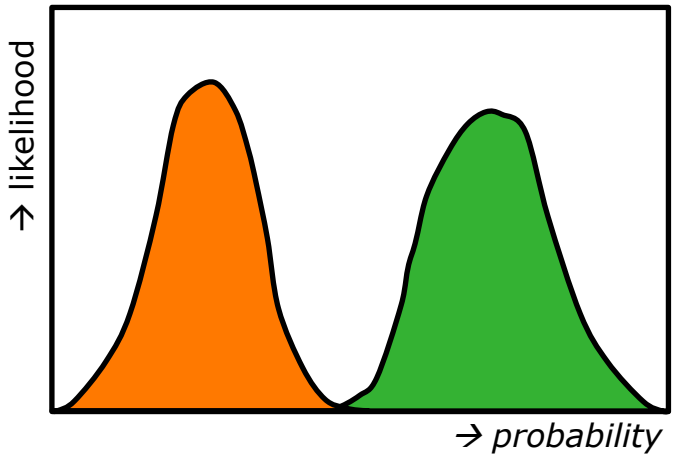
Two examples, the same classification result:





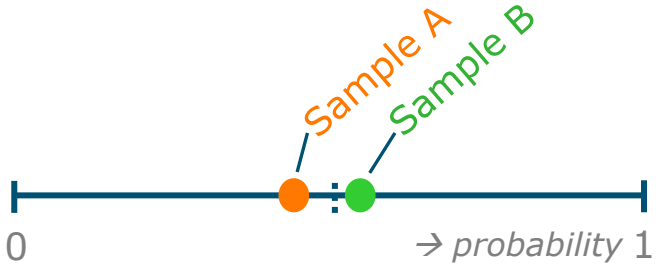
Probability score:

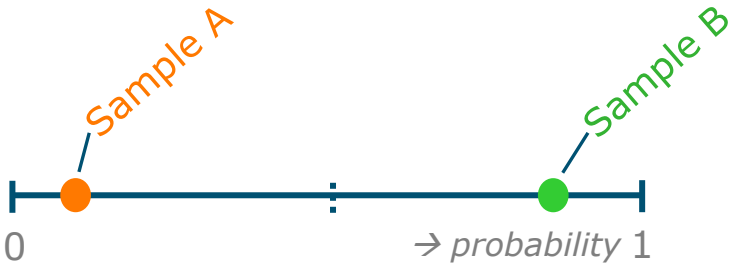
- a candidate QC for *multivariate* **repeatability** and **reproducibility**?



Performance evaluation

Two examples, the same classification result:

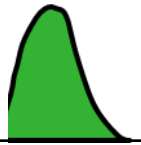
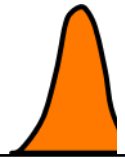




Density →

Confusion matrix

Model performance?

	Predicted: Authentic	Predicted: Fraud
Actual: Authentic		
Actual: Fraud		

Sources of error/variation: training set

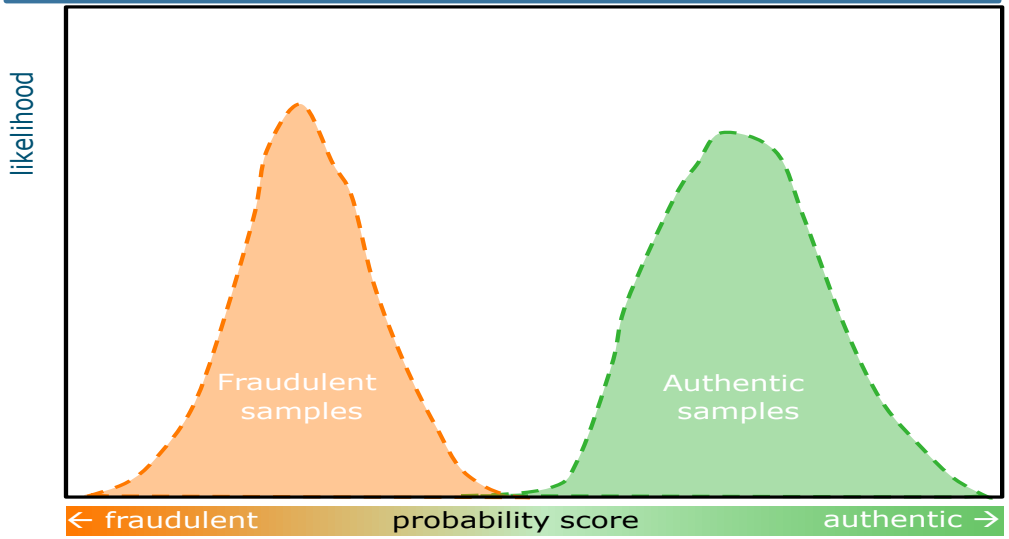
Error & Variation:

- Analytical variation:

➔ all usual analytical variation (for *each* of the variables!)

- Natural variability
- Storage
- Sample handling
- Species
- Regions
- ....

Probability distribution obtained using resampling techniques *(e.g. rCV)* on training set

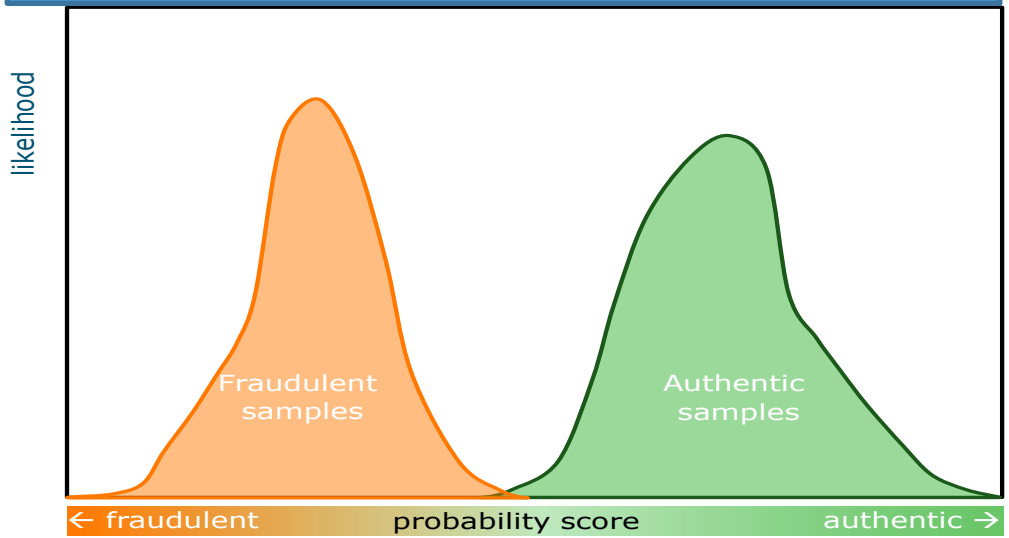


Sources of error/variation: validation set

Additional variation (within scope) due to obtained differences in:

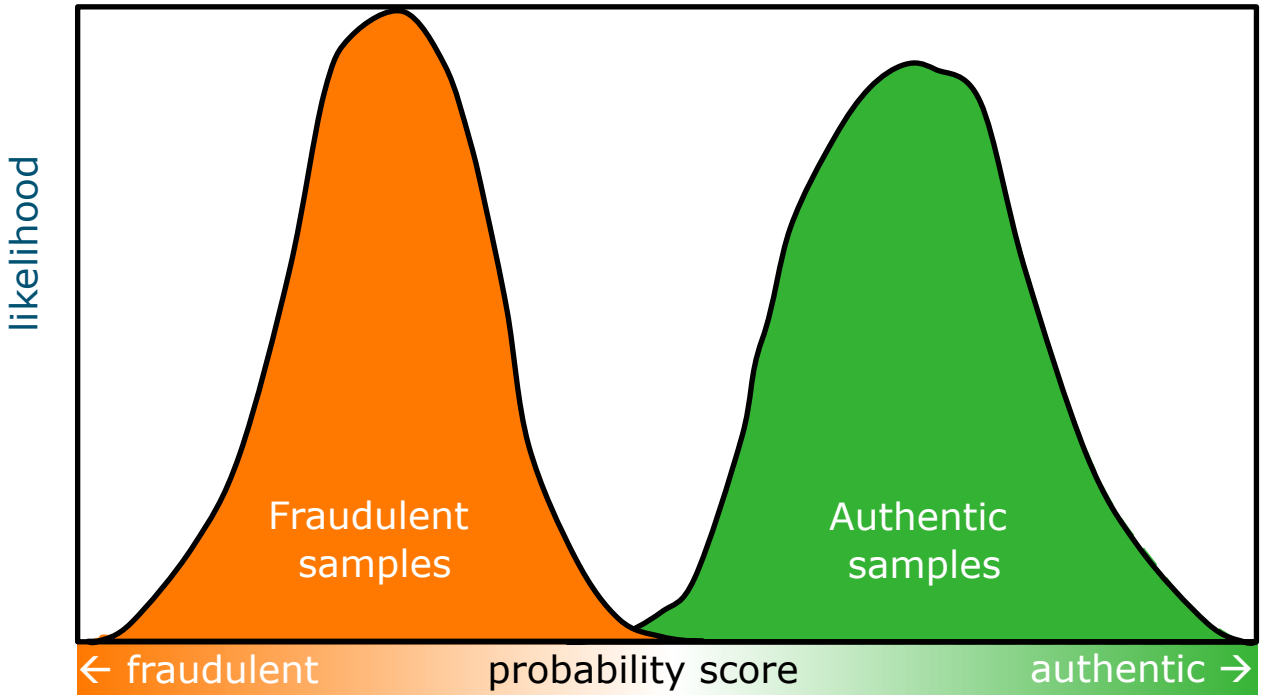
- Harvest
- Producers
- Storage conditions
- ....
- Analytical equipment
- Technician
- Sample preparation
- Solvents
- ...

Probability distribution obtained predicting the validation set samples by the model



"Worst-case" probability distribution

- Combining the sources of error: widening effect



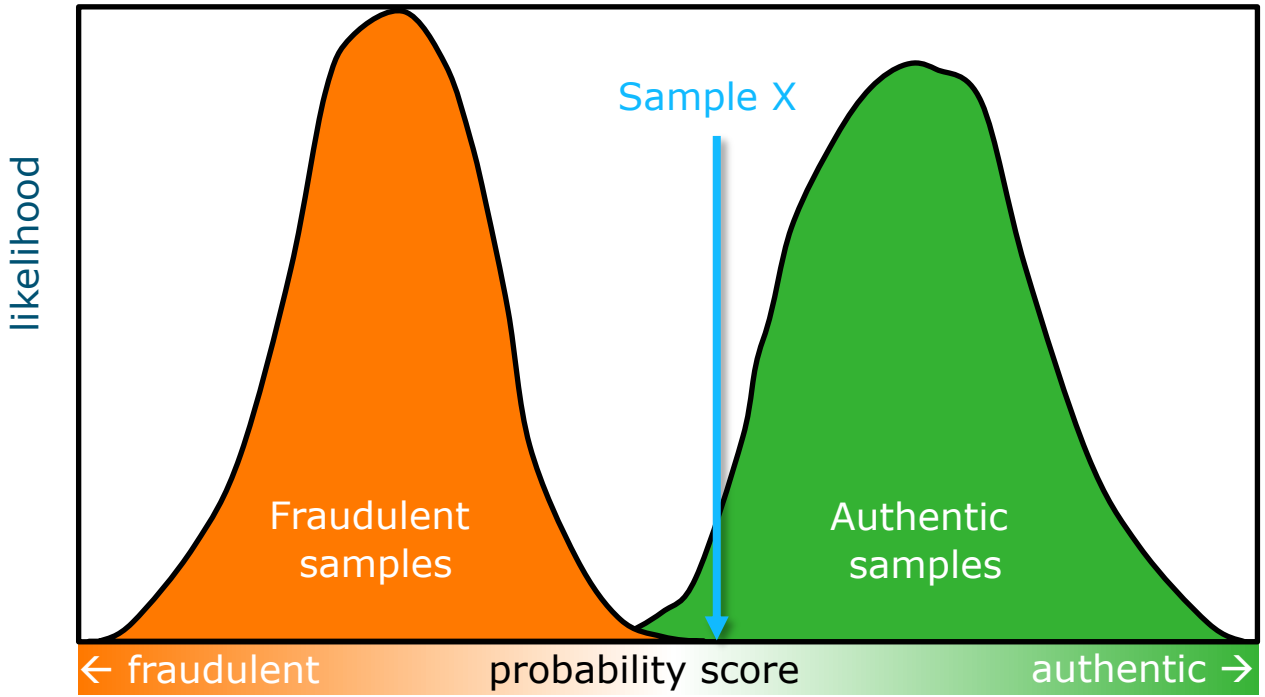
Training set error:
-natural variation
-analytical variation

External errors (validation set):
-population change
-analytical drift
-robustness

Expanded (widened) **distribution:**
-"worst-case" model performance

"Worst-case" probability distribution

- Allows providing certainty statements for future results



Example

Sample X:
score = 0.55

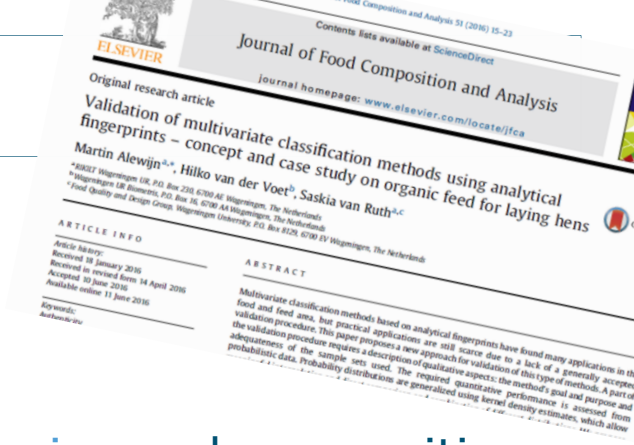
→ *interpretation:*

$P_{authentic}$ = 7%
P_{fraud} = 0.05%

Summary of validation approach:

- Evaluates the sum of sample set + analysis + model
- Bases performance on probabilities rather than binary results
- Combines different sources of error into the "worst-case" overall (un)certainty profile
- Allows adding certainty statements to future samples

Next steps


Journal of Food Composition and Analysis

Validation of multivariate classification methods using analytical fingerprints – concept and case study on organic feed for laying hens

Martin Alewijn, Hilko van der Voet, Saskia van Ruth

- ✓ Publication
- International input (CEN TC460 FA)

- Approach for quantitative validation of sample size and composition
- Add "expanded measurement uncertainty" to the final distribution?

- Get a method formally accredited – *in progress*

http://www.sciencedirect.com/science/article/pii/S088915751630076X

Thank you
for your attention

martin.alewijn@wur.nl