OUTLIER SCREENING IN TEST OF AUTOMOTIVE SEMICONDUCTORS

USE OF JMP 12 PRO ‘MULTIVARIATE ANALYSIS’ PLATFORMS AND ‘EXPLORE OUTLIERS’ UTILITY

CORINNE BERGÈS
BLACK BELT FOR ANALOG & SENSOR QUALITY

15, FEB, 2016
Introduction

- Hundreds of parametric tests per part

- Parts whose test results are far from normal process variability → Likely-to-fail parts / Outliers

- Test → Search for outliers

- Outlier detection means:
  - Univariate screening
  - Multivariate screening
    - methods without learning
    - methods with learning
Univariate analysis
Univariate outlier
Multivariate analysis with jmp
1. Presentation on some multivariate techniques that are possible to be run with jmp
   → Case study for an automotive valve driver

2. Considerations about space size

3. Efficiency and yield loss

4. ‘Explore Outliers’ jmp platform
1. Some multivariate technics with jmp

- Methods without learning step
- Methods needing a learning step
Types of multivariate analysis

- **Methods without learning step**
  - Based on a detection threshold that is directly linked with yield loss
  - Challenge: setting of a threshold that detects returns with the lowest yield loss
  - Examples: Mahalanobis distance estimation, k-means clustering method, deviation estimation from a linear regression

**Mahalanobis distance estimation**
Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

**K-near neighbors and clustering methods**
Distance estimation from each observation to the K-near neighbors
Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last centroids by new ones including the last observation assigned

**Deviation estimation from a linear regression**
Bivariate method (2 tests) on tests highly correlated
Distance estimation from each point to the linear regression line between the 2 tests
Types of multivariate analysis

- **Methods needing a learning step**
  - Implementation step:
    - Learning on first well-known customer returns
    - Running of the first built model to detect outliers among the following manufactured parts
    - Improvement of the first model by new potential returns
  - Challenge: building of a model that does not stick to the part sample but that could be used to detect outliers and returns on other following part samples (overfitting risk)
  - Examples: discriminant analysis, partial least squares (pls)

**Discriminant analysis**

Membership prediction in a category (failed/not-failed) from observed values

Search for a test combination that provides a maximal Mahalanobis distance between the two groups

Entropy R-Square measures model fit

**Partial Least Squares (PLS)**

Trade-off between two purposes: to maximise:

- Explained variance of the predictors
- Correlation between variables and response

Main method advantage: to be run even if number of tests > number of parts

Two available algorithms: NIPALS and SIMPLS
Types of multivariate analysis

• **Methods without learning step**
  - Based on a detection threshold that is directly linked with yield loss
  - Challenge: setting of a threshold that detects returns with the lowest yield loss
  - Examples: Mahalanobis distance estimation, k-means clustering method, deviation estimation from a linear regression

• **Methods needing a learning step**
  - Implementation step:
    - Learning on first well-known customer returns
    - Running of the first built model to detect outliers among the following manufactured parts
    - Improvement of the first model by new potential returns
  - Challenge: building of a model that does not stick to the part sample but that could be used to detect outliers and returns on other following part samples (overfitting risk)
  - Examples: discriminant analysis, partial least squares (pls)

→ **Success criteria: low yield loss**
How to do with jmp?
Case study for an automotive valve driver
Case study

- Product: automotive valve driver

- Failing of univariate detection for several customer returns

- Test of several multivariate analysis methods:
  - 745 tests (where standard deviation is not null)
  - 13,000 parts → only one is failing = the customer return
  - Around 20 wafers tested at final test (after part assembly)

- Space size: $n = 13,000$ parts; $p = 745$ tests

- Question: what is the best multivariate method to detect the customer returns with jmp?
Methods without learning (1/3)

Mahalanobis distance estimation
Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

File: ‘Multivariate analysis jmp’ → 3 tests
Methods without learning (2/3)

K-near neighbors and clustering methods

Distance estimation from each observation to the K-near neighbors

Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last centroids by new ones including the last observation assigned

File: ‘Multivariate analysis.jmp’ → 3 tests
Methods without learning (3/3)

Deviation estimation from a linear regression
Bivariate method (2 tests) on tests highly correlated
Distance estimation from each point to the linear regression line between the 2 tests

Distance \((P(x, y), \text{regression line})\)

\[ = \sqrt{[\text{residuals(test2 vs test1)}]^2 + [\text{residuals(test1 vs test2)}]^2}/2 \]

File: ‘Bivariate and PCA.jmp’
Methods with learning (1/2)

**Discriminant analysis**

Membership prediction in a category (failed/not-failed) from observed values

Search for a test combination that provides a maximal Mahalanobis distance between the two groups

Entropy R-Square measures model fit

File: ‘Bivariate and PCA.jmp’
Methods with learning (2/2)

Partial Least Squares (PLS)

Trade-off between two purposes: to maximise:
- Explained variance of the predictors
- Correlation between variables and response

Main method advantage: to be run even if number of tests > number of parts

Two available algorithms: NIPALS and SIMPLS

File: ‘Partial Least Square.jmp’
Methods without learning

Mahalanobis distance estimation
Spatial distance based on the inverse of the variance-covariance matrix for the p-tests

K-near neighbors and clustering methods
Distance estimation from each observation to the K-near neighbors
Clustering: Iterative algorithm that assigns each observation to the nearest cluster centroid and replaces the last

Deviation estimation from a linear regression
Bivariate method (2 tests) on tests highly correlated
Distance estimation from each point to the linear regression line between the 2 tests

Case study: best multivariate method → Mahalanobis distance with a yield loss = 0.36%
Methods with learning

• **Discriminant analysis**
  Membership prediction in a category (failed/not-failed) from observed values
  Search for a test combination that provides a maximal Mahalanobis distance between the two groups
  Entropy R-Square measures model fit
  Mahalanobis distance plot in two-dimensional space
  Case study: low entropy R-Square (0.39)

• **Partial Least Squares (PLS)**
  Trade-off between two purposes: to maximise:
  - Explained variance of the predictors
  - Correlation between variables and response
  Main method advantage: to be run even if number of tests > number of parts
  Two available algorithms: NIPALS and SIMPLS

Case study: Mahalanobis distance without learning stays the best multivariate method
2. Space size
Considerations about space size

• **Size reduction motivation:**
  - Test reduction:
    • Better results on a reduced space and on correlated tests
    • For the analysis with learning, overfitting risk reduction
  - Part reduction: noise reduction on homogeneous data

• **Size reduction means:**
  - Statistical analysis: Principal Component Analysis (PCA)
  - Other selection criteria: functionality criteria
  - Part reduction: run on wafer lot
Principal Component Analysis (PCA)

- Case study:
  After PCA, return detected with a higher yield loss (1%)

File: ‘Bivariate and PCA.jmp’
3. Efficiency and yield loss
Efficiency and yield loss (1/5)

• Efficiency: outlier detection ability with minimal yield loss

• One mean to increase efficiency: noise reduction

• Case study: test performed on four sites → multivariate analysis to visualize and understand additional noise due to sites:
  – K-means clustering method
  – Contingency analysis
  – ANOVA
**Efficiency and yield loss (2/5)**

- **K-means clustering method**

Two clusters observed:
- one for one site (blue) → *Cluster #10 for the following study*
- one gathering data from three sites → *Cluster #20 for the following study*

File: ‘Noise analysis.jmp’
Efficiency and yield loss (3/5)

- Contingency analysis

→ Cluster #10 contains site 2 data
→ Cluster #20 contains data from sites 0, 1 and 3

File: ‘Noise analysis.jmp’
Efficiency and yield loss (4/5)

- **ANOVA** → **ANOVA of one test distribution by site**

Statistical test to compare means

For this test, data from the site 1 are significantly different from the other sites

File: ‘Noise analysis.jmp’
Efficiency and yield loss (5/5)

- Efficiency: outlier detection ability with minimal yield loss
- One mean to increase efficiency: noise reduction
- Case study: test performed on four sites → multivariate analysis to visualize and understand additional noise

**K-means clustering method**

Two clusters:
- one for one site (blue)
- one gathering data from three sites

**Contingency analysis**

Vizualisation of clustering method results

**ANOVA**

ANOVA of one test distribution by site

- Noise elimination after part test: possibility to shift and align means of each site → *Yield loss decrease*
4. ‘Explore Outliers’ jmp platform
‘Explore Outliers’ jmp platform

File: ‘Multivariate analysis.jmp’ → 3 tests
Conclusion
Conclusion

- Outlier detection in univariate analysis \(\rightarrow\) Robust PAT for a better detection (real outliers) and a lower yield loss

- Outlier detection in multivariate analysis:
  - Many multivariate analysis based on the spatial Mahalanobis distance
  - Method without learning:
    - Useful data diluted in multidimensional space
    - High computation time and cost
    - In a reduced space, higher yield loss
  - Method with learning: reduced space but detection failing risk increase
  - One of the easiest method to be implemented: ‘distance-to-regression-line estimation’ method: Python used in the model design step / EWM in probe
  - Many other methods have to tested, in Python or in jmp, above all when a CQI happens

- In order to improve detection with a lower yield loss, a preliminary step has to be gage study / noise reduction and elimination \(\rightarrow\) will benefit also the univariate analysis
Thank you for your attention

Any question ?

→ Please, feel free to contact Corinne Bergès:
  corinne.berges@nxp.com