EPPO Workshop on Accreditation for plant pest diagnostic laboratories

Statistical analysis (in interlaboratory comparison and method validation)

Aude Chabirand, Laurent Folcher, Vincent Hérau
ANSES Plant Health Laboratory

February 2014
Introduction:
Basic concepts in statistics
Life-cycle of the analytical methods

Definition:
The validation is the confirmation by examination and the provision of objective evidence that the particular requirements for a specific intended use are fulfilled (ISO17025 § 5.4.5.1)

Need to monitor the analytical method and its implementation:
- Internal quality controls:
  - Use of reference materials
  - Compliance testing
  - Quality control charts...
- External quality controls: proficiency testing ...

DEVELOPMENT

SELECTION

VALIDATION

ROUTINE

Optimization

Transfer

DATA

Increasing need for statistics
Why statistics in plant pathology?

Statistics are an essential tool to be able to look critically at numerical information and not to be misled.

Statistics help to make good use of such information to make sound decisions.

In plant pathology, statistics play a crucial role in particular for designing and analyzing method validations and interlaboratory comparisons. They allow to quantify the evidence in data about a particular truth/relationship and can help to choose the most appropriate method according to the epidemiological context (modeling).
Why statistics in plant pathology?

Example in method validation

Example: the criterion analytical specificity (ASP) obtained from 60 target and 60 nontarget plant samples is equal to 92.5%.

- What confidence can I put in this observed result? I have analyzed only 60 target samples and 60 nontarget samples!

- Is this observed result compatible with an expected value for this criterion equal to 99%?
Why statistics in plant pathology?

Example in proficiency testing

Example: I produced, from the same test matrix, 25 test samples supposed to be identical.

- How can I be sure that these samples are sufficiently homogeneous and stable to be used for the participants’ performance assessment?

It is essential to identify the nature of the variables studied to determine the appropriate data processing and apply the adequate statistical tests.
A variable is any characteristic, number or quantity that can be measured or counted.

- **Quantitative**
  - Discrete
  - Continuous

- **Qualitative**
  - Nominal
  - Ordinal
Understanding the different types of variables

**QUANTITATIVE VARIABLE (possible measurement)**

- **DISCRETE VARIABLE**
  - Is restricted to particular values (e.g. integer value) in an interval
  - Based on a count
  - *E.g. : number of bacteria*

- **CONTINUOUS VARIABLE**
  - Can take any value between a certain set of real numbers (in an interval)
  - *E.g. : blood glucose, absorbance value*

**QUALITATIVE VARIABLE (no possible measurement)**

- **ORDINAL VARIABLE - SEMI-QUANTITATIVE VARIABLE**
  - Can be logically ordered or ranked
  - *E.g. : disease severity rated on a scale (none, mild, moderate, severe), or with numerical values (0,1,2,3) . The different points of the scale are not necessarily equivalent.*

- **NOMINAL (OR CATEGORICAL) VARIABLE – binary if only two modalities**
  - *Not able to be organized in a logical sequence*
  - *E.g. : test result (positive, negative, indeterminate)*
A **statistical unit** is a unit of observation or measurement for which data are collected.

The variable of interest is generally studied from a group of statistical units (the **statistical sample**) selected from a larger group (**the population**). For each population there are many possible samples. **The statistical sample must NOT be confused with the laboratory sample.**
Statistical Inference makes use of information from a sample to draw conclusions (inferences) about the population from which the sample was taken.

A population is any entire collection of people, animals, plants or things from which we may collect data. It is the entire group we are interested in, which we wish to describe or draw conclusions about.

In order to make any generalizations about a population, a sample, that is meant to be representative of the population, is often studied.

By studying the sample (descriptive statistics), it is hoped to draw valid conclusions about the population (inferential statistics).

The representativeness of the statistical sample determine the scope of the inference and of its conclusion.
Descriptive and inferential statistics

**Population**
(Size: \(N_{\text{pop}}\))

- Entire collection of possible laboratory samples (\(N_{\text{pop}}\))

**Sampling and survey**
\(N_{\text{sam}} < N_{\text{pop}}\)

**Sample**
(Size: \(N_{\text{sam}}\))

- Group of units (laboratory samples) selected from the population (\(N_{\text{sam}}\))

**Deduction**

- Inferential statistics: estimate, hypothesis tests...

---

**Study of the sample characteristics**

**Descriptive statistics**: proportion, mean, standard deviation...; graphical representation...
In method validation or interlaboratory comparison, a laboratory sample = a unit (an individual).

All the laboratory samples used to evaluate for example the analytical specificity of a method constitute the statistical sample used for the evaluation of this method specificity.

By studying the analytical specificity of the method in the sample, it is hoped to draw valid conclusions about the analytical specificity of the method in the population i.e. in the entire collection of possible laboratory samples.

When possible, it is preferable to use laboratory samples with a known reference result to make more precise comparisons with fewer laboratory samples.

However, when no reference result is available, it is possible to use statistical models appropriate for such situations.
Contingency table

A contingency table is a type of table in a matrix format that displays the frequency distribution of the qualitative variables.

<table>
<thead>
<tr>
<th>Test result</th>
<th>Reference result</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td></td>
<td>PA</td>
<td>PD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Positive agreement)</td>
<td>(Positive deviation)</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Negative deviation)</td>
<td>(Negative agreement)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>N⁺</td>
<td>N⁻</td>
</tr>
</tbody>
</table>

Analytical specificity (ASP) (Diagnostic accuracy)

EPPO PM7/98

Diagnostic sensitivity (SE) = number of PA / N⁺
Diagnostic specificity (SP) = number of NA / N⁻
Application in plant pathology

Which variable? Which distribution? Which statistical test?

Test result:
- positive, negative
- or, if applicable indeterminate

Agreement:
- Yes (agreement)
- vs. No (deviation)

Number of agreements (count):
- 14

Qualitative nominal variable
- Qualitative binary variable
- Quantitative discrete variable

Approximation
- \( n \to \infty, p \neq 0 \) and \( p \neq 1 \)
  - (in practice: \( n \geq 30, np \geq 5 \) and \( n(1-p) \geq 5 \))
- \( n \to \infty, p \to 0 \)
  - (in practice: \( n \geq 30, p \leq 0.1 \) and \( np < 5 \))

Bernoulli distribution
- with the parameter \( p \)

Binomial distribution
- with the parameters \( n \) and \( p \)
  - (\( p \) constant for each trial)

Normal distribution

Poisson distribution

\( n = \) size of the statistical sample
\( p = \) probability of successes (proba of agreement)
Application in plant pathology

Example:
Validation requirement: to be validated the candidate method B must provide 99% of agreement with reference results \((SE_B = 99\% \ SE_R \text{ and } SP_B = 99\% \ SP_R)\).

Can we validate the candidate method B, according to the following results?

<table>
<thead>
<tr>
<th>Reference result R</th>
<th>R+</th>
<th>R-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate method B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B+</td>
<td>PA=58</td>
<td>PD=7</td>
</tr>
<tr>
<td>B-</td>
<td>ND=2</td>
<td>NA=53</td>
</tr>
<tr>
<td>Total</td>
<td>N+=60</td>
<td>N-=60</td>
</tr>
</tbody>
</table>

Diagn. sensitivity \((SE_B) = 58/60 = 96.7\%\)  
Diagn. specificity \((SP_B) = 53/60 = 88.3\%\)

\(ASP_B\) Compatible ?

Statistics Generalization, Modeling

Sample

Observed \(SE_B = 96.7\% \ SE_R\)
Observed \(SP_B = 88.3\% \ SP_R\)

Population

True \(SE_B = 99\% \ SE_R\)
True \(SP_B = 99\% \ SP_R\)
Synoptic of statistical tests used to compare proportions

Comparison of proportions (or distributions)

1 observed proportion versus 1 theoretical proportion
- Normal approximation
  - $n \geq 30$, $np \geq 5$ and $n(1-p) \geq 5$
  - Other cases
    - $np_i < 5$
    - Chi-square goodness of fit test
    - Exact binomial test

2 observed proportions
- Independent
  - Normal approximation
    - $n_1 \geq 30$, $n_2 \geq 30$, $n_1 p' \geq 5$ and $n_2 p' \geq 5$
  - Other cases
    - $np_i < 5$
    - Z-test
    - Chi-square test of homogeneity
- Paired
  - Normal approximation
    - Number of discordances $\geq 10$
  - Other cases

Comparison of more than 2 proportions
- Independent
  - Normal approximation
    - No theoretical numbers $np_i$ equal to zero and no fewer than 20% of the theoretical numbers $np_i < 5$
  - Other cases
- Paired
  - Normal approximation

With $p$ = observed proportion (e.g., observed diagnostic SE) ; $p'$ = mean of the two proportions to be compared
$n$ = size of the statistical sample (N+ for diagnostic SE)
SO...

To be continued in the practical session...

Thank you for your attention