Statistics calculated on confusion matrix

Confusion matrix

**Theoretical confusion matrix**

<table>
<thead>
<tr>
<th></th>
<th>Automatic classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Manual classification</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>TP</td>
</tr>
<tr>
<td>Negative</td>
<td>FP</td>
</tr>
</tbody>
</table>

**Example**

Automatic classification of three groups A, B, C

<table>
<thead>
<tr>
<th></th>
<th>Automatic classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Manual classification</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
</tr>
</tbody>
</table>

For group A, the matrix can be reduced as:

<table>
<thead>
<tr>
<th></th>
<th>Automatic classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group A</td>
</tr>
<tr>
<td>Manual classification</td>
<td></td>
</tr>
<tr>
<td>Group A</td>
<td>TP = 1</td>
</tr>
<tr>
<td>Not group A</td>
<td>FP = 1</td>
</tr>
</tbody>
</table>

For group B, the matrix can be reduced as:

<table>
<thead>
<tr>
<th></th>
<th>Automatic classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group B</td>
</tr>
<tr>
<td>Manual classification</td>
<td></td>
</tr>
<tr>
<td>Group B</td>
<td>TP = 3</td>
</tr>
<tr>
<td>Not group B</td>
<td>FP = 1</td>
</tr>
</tbody>
</table>

For group C, the matrix can be reduced as:

<table>
<thead>
<tr>
<th></th>
<th>Automatic classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group C</td>
</tr>
<tr>
<td>Manual classification</td>
<td></td>
</tr>
<tr>
<td>Group C</td>
<td>TP = 3</td>
</tr>
<tr>
<td>Not group C</td>
<td>FP = 1</td>
</tr>
</tbody>
</table>
**Explanation of main parameters**

**TP : True Positive**
Number of particles of the group of interest correctly classified.

**TN : True Negative**
Number of particles of all the other groups classified as other groups.

**FP : False Positive**
Number of particles of other groups classified in the group of interest.

**FN : False Negative**
Number of particles of the group of interest classified in the other groups.

**General statistics**

**Accuracy**
\[
\frac{TP + TN}{(TP + TN + FP + FN)}
\]

**Error**
1 – Accuracy
The eight basic ratios

**Recall**
Also called: Sensitivity, TPR : True Positive Rate, Power, Probability of detection

\[
\frac{TP}{TP+FN} = 1 - FNR
\]

**Specificity**
Also called: TNR : True Negative Rate, Selectivity

\[
\frac{TN}{TN+FP} = 1 - FPR
\]

**Precision**
Also called: PPV : Positive Predicted Value, Reproducibility, Repeatability

\[
\frac{TP}{TP+FP} = 1 - FDR
\]

**NPV : Negative Predicted Value**

\[
\frac{TN}{TN+FN} = 1 - FOR
\]

**FPR : False Positive Rate**
Also called: alpha, Type I Error, p-Value

\[
\frac{FP}{FP+TN} = 1 - \text{Specificity}
\]

**FNR : False Negative Rate**
Also called: beta, Type II Error

\[
\frac{FN}{TP+FN} = 1 - \text{Recall}
\]

**FDR : False Discovery Rate**
Also called: q-Value

\[
\frac{FP}{TP+FP} = 1 - \text{Precision}
\]
FOR : False Omission Rate

\[ \frac{FN}{(FN + TN)} = 1 - NPV \]
The four ratios of ratios

**LRPT : Likelihood Ratio for Positive Tests**

\[
\frac{TP}{FP} \frac{(TP + FN)}{(FP + TN)} = \frac{Recall}{(1 - Specificity)} = \frac{Recall}{FPR}
\]

**LRNT : Likelihood Ratio for Negative Tests**

\[
\frac{FN}{TN} \frac{(FN + TP)}{(TN + FP)} = \frac{(1 - Recall)}{Specificity} = \frac{FNR}{Specificity}
\]

**LRPS : Likelihood Ratio for Positive Subjects**

\[
\frac{TP}{FP} \frac{(TP + FN)}{(TN + FP)} = \frac{Precision}{(1 - NPV)} = \frac{Precision}{FOR}
\]

**LRNS : Likelihood Ratio for Negative Subjects**

\[
\frac{FP}{TN} \frac{(FP + TP)}{(TN + FN)} = \frac{(1 - Precision)}{(1 - FOR)} = \frac{FDR}{NPV}
\]
Additional statistics

**Fmes : F-measure**
Also called: F1-score, harmonic mean of precision and recall

\[ 2 \times \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \]

**BalAcc : Balanced accuracy**

\[ \frac{\text{Specificity} + \text{Recall}}{2} \]

**MCC : Matthews Correlation Coefficient**
Attention: If any sum of the denominator is 0, the total denominator can be set to 1

\[ \frac{(TP \times TN) - (FP \times FN)}{\sqrt{((TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN))}} \]

**Chisq : χ²**
Also called : Significance

\[ \frac{[(TP \times TN) - (FP \times FN)]^2 \times (TP + TN + FP + FN)}{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)} \]

**Auto_Manu : Difference between Automatic and Manual classification**

\[ (TP + FP) - (TP + FN) \]

**Dissimilarity Index of Bray Curtis**

\[ \frac{\| \text{Auto-Manu} \|}{\sum (TP + FP) + \sum (TP + FN)} \]