What are Mandel’s $k/h$ test statistics?
Objectives of Mandel’s \( k/h \) consistency test statistics

- \( k \) and \( h \) test statistics are measures for data consistency, particularly useful for inter-laboratory studies.
- By studying the collated data deviations and accuracy, the performance of a laboratory in terms of its reliability and errors can be established.
- The laboratory with poor performance can then do its own in-house investigation and make corrective actions for such deficiencies.
Mandel’s $k$ and $h$ consistency test statistics are discussed in ASTM E691 standards for interlaboratory analysis:

“Standard practice for conducting an interlaboratory study to determine the precision of a test method”
Inter-laboratory cross-checks and proficiency testing programs

• Inter-laboratory comparison of test results is an efficiency way to validate the precision of a test method and also to compare the technical competence of the laboratory personnel in terms of precision and accuracy

• Many participating laboratories will carry out series of analyses on one or more given similar samples at about the same period. The data collated are statistically analyzed
Evaluating $k$ test statistic

- $k$ value is a measure of *within*-laboratory consistency in repeatability

- If there are $p$ number of participating laboratories $(j)$, and $n$ is the number of repeats in a laboratory $(x_1, x_2, x_3, \ldots x_i \ldots, x_{n-1}, x_n)$

- The $k$ value of lab $(j)$ is:

\[
 k_j = \frac{s_j}{s_r}
\]

where

\[
 s_j = \sqrt{\frac{\sum_{i=1}^{n} (x_{i,j} - \bar{x}_j)^2}{n-1}}
\]

and

\[
 s_r = \sqrt{\frac{\sum_{j=1}^{p} s_j^2}{p}}
\]
**Interpretation of $k$ value**

- The $k$ value compares the repeatability standard deviation of a laboratory data set with the average of the repeatability standard deviations of all other laboratories.

- From the $k$ value, we can evaluate the spread of the data set and its precision.

- This test statistic reflects the *single* lab’s repeatability against the average repeatability of all participating laboratories.

- The larger the $k$ value, the bigger is the data deviation, indicating the poorer the precision.
\( k \) critical value for consistency \((k\text{-crit})\)

- \( k\text{-crit} \) value is the critical value of seriousness for data deviation at a given probability.
- \( k\text{-crit} \) defines as:
  \[
  k\text{-crit} = \sqrt{\frac{p}{1+(p-1)/F}}
  \]
- Where: \( F \) value is from the \( F - F \) distribution, \( p \) is the number of participating laboratories.
- When the \( k \) value is higher than the \( k\text{-crit} \), it can be concluded that the test result deviation is serious with poor precision and unacceptable.
How to obtain the $F$ test statistic value?

- $F(v_1, v_2)$ is the $F - F$ distribution value
- Degree of freedom $v_1 = (n-1)$, $n$ is the number of repeats in a single laboratory
- Degree of freedom $v_2 = (p-1)(n-1)$
- Upon knowing the degrees of freedom, we can obtain the $F$ value from the $F$–$F$ table
- Or use the Excel function “=FINV(0.05,v_1,v_2)”
Evaluating the $h$ test statistic

• The $h$ test statistic is used to examine the consistency of inter-laboratory data, confirming if any laboratory data is an outlier.

• In other words, it is to indicate the accuracy of a lab results against the others reported.

• Let $p$ be the number of participating labs with the lab mean results as follows:

$$\left( x_1, x_2, \ldots, x_j, \ldots, x_p \right)$$

• The overall mean result of this interlaboratory study is:

$$x = \frac{\sum_{j=1}^{p} x_j}{p}$$
Evaluating the $h$ test statistic

- The deviation of mean result of a lab ($j$) from the overall mean is:
  \[ d_j = \bar{x}_j - \bar{x} \]

- The standard deviation of these comparison is:
  \[ s_x = \sqrt{\frac{\sum_{j=1}^{p} d_j^2}{p - 1}} \]

- The $h$ value of lab ($j$) is:
  \[ h_j = \frac{d_j}{s_x} \]
Interpretation of \( h \) test value

- \( h \) test statistic value reflects the deviation of a single laboratory’s mean test results from the overall mean results obtained from all participating laboratories.
- The larger the \( h \) value, the bigger the deviation, the poorer is the accuracy of that single laboratory.
**h critical value for consistency (h-crit)**

- **h-crit** is a measure of seriousness in a lab’s inaccuracy

- **h-crit defines as**:

  \[ h - crit = \pm \frac{t(p-1)}{\sqrt{p(t^2 + p-2)}} \]

- where: \( t \) is the Student’s distribution with degree of freedom \( v = p - 2 \), and \( \alpha = 0.05 \);

  \( p \) is the number of participating laboratories

When the \( h \) value is larger than the \( h\text{-crit} \), it is concluded that the mean result given by the laboratory concerned is not accurate and reliable.
### Mandel's k/h 统计测验表

#### 原始数据

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>总体值</td>
<td>-</td>
<td>11.2</td>
<td>9.88</td>
<td>11.64</td>
<td>8.86</td>
</tr>
<tr>
<td>偏差值</td>
<td>5.55</td>
<td>7.06</td>
<td>3.92</td>
<td>5.34</td>
<td>5.88</td>
</tr>
<tr>
<td>方差</td>
<td>0.435</td>
<td>0.590</td>
<td>0.305</td>
<td>0.477</td>
<td>0.554</td>
</tr>
<tr>
<td>总方差</td>
<td>1.377</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>实验室数目</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>均方误差</td>
<td>0.375</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>均方标准差</td>
<td>0.813</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$k$ 测验

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>偏差值</td>
<td>-</td>
<td>5.44</td>
<td>1.42</td>
<td>5.77</td>
<td>0.27</td>
</tr>
<tr>
<td>方差</td>
<td>0.435</td>
<td>0.349</td>
<td>0.827</td>
<td>0.227</td>
<td>0.265</td>
</tr>
<tr>
<td>总方差</td>
<td>1.877</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### h 测验

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>偏差值</td>
<td>0.813</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>方差</td>
<td>0.813</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>总方差</td>
<td>9.567</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>实验室数目</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>均方误差</td>
<td>0.313</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>均方标准差</td>
<td>0.235</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$h$ 测验

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>偏差值</td>
<td>0.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>方差</td>
<td>0.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>总方差</td>
<td>1.526</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ k - crit = \frac{p}{1 + \frac{(p - 1)}{k}} \]

\[ h - crit = \frac{r(p - 1)}{\sqrt{p(\sum r^2 + p - 2)}} \]