

This paper explains the research conducted by Minitab statisticians to develop the methods and data checks used in the Assistant in Minitab Statistical Software.

# Attribute Agreement Analysis

## Overview

Attribute Agreement Analysis is used to assess the agreement between the ratings made by appraisers and the known standards. You can use Attribute Agreement Analysis to determine the accuracy of the assessments made by appraisers and to identify which items have the highest misclassification rates.

Because most applications classify items into two categories (for example, good/bad or pass/fail), the Assistant analyzes only binary ratings. To evaluate ratings with more than two categories, you can use the standard Attribute Agreement Analysis in Minitab (**Stat > Quality Tools > Attribute Agreement Analysis**).

In this paper, we explain how we determined what statistics to display in the Assistant reports for Attribute Agreement Analysis and how these statistics are calculated.

**Note** No special guidelines were developed for data checks displayed in the Assistant reports.

# Output

There are two primary ways to assess attribute agreement:

- The percentage of the agreement between the appraisals and the standard
- The percentage of the agreement between the appraisals and the standard after removing the effect of percentage of agreement by random chance (known as the kappa statistics)

The analyses in the Assistant were specifically designed for Green Belts. These practitioners are sometimes unsure about how to interpret kappa statistics. For example, 90% agreement between appraisals and the standard is more intuitive than a corresponding kappa value. Therefore, we decided to exclude kappa statistics from the Assistant reports. However, the disadvantage of only reporting the percent of agreement is that the value includes both the agreement due to using a common assessment standard and the agreement by chance; the kappa statistic removes agreement by chance in its calculation. For this reason, when you use the Assistant, we encourage you to select an equal number of good and bad products across evaluations so that the percentage of agreement by chance is approximately the same.

The Assistant report displays pairwise percentage agreement values, which is different than the results from **Stat > Quality Tools > Attribute Agreement Analysis**. For example, an appraiser collects 2 trials on each test item. In the Assistant report, if the Appraiser matches the standard for test item X on the first trial but not on the second trial, the Appraiser gets credit for 1 match. In the analysis from the Stat menu, the Appraiser only gets credit when his or her ratings for both trials match. See Methods and Formulas in Minitab Help for the detailed calculations used in the Stat menu analysis.

The Assistant reports show pairwise percentage agreements between appraisals and standard for appraisers, standard types, trials, and the confidence intervals for the percentages. The reports also display the most frequently misclassified items and appraiser misclassification ratings.

# Calculations

The pairwise percentage calculations are not included in the output in the standard Attribute agreement analysis in Minitab (**Stat > Quality Tools > Attribute Agreement Analysis**). In fact, kappa, which is the pairwise agreement adjusted for the agreement by chance, is used to represent the pairwise percent agreement in this output. We may add pairwise percentages as an option in the future if the Assistant results are well received by users.

We use the following data to illustrate how calculations are performed.

Appraisers	Trials	Test Items	Results	Standards
Appraiser 1	1	Item 3	Bad	Bad
Appraiser 1	1	Item 1	Good	Good
Appraiser 1	1	Item 2	Good	Bad
Appraiser 2	1	Item 3	Good	Bad
Appraiser 2	1	Item 1	Good	Good
Appraiser 2	1	Item 2	Good	Bad
Appraiser 1	2	Item 1	Good	Good
Appraiser 1	2	Item 2	Bad	Bad
Appraiser 1	2	Item 3	Bad	Bad
Appraiser 2	2	Item 1	Bad	Good
Appraiser 2	2	Item 2	Bad	Bad
Appraiser 2	2	Item 3	Good	Bad

## Overall accuracy

The formula is

$$100 \times \frac{X}{N}$$

Where

- X is the number of appraisals that match the standard value
- N is the number of rows of valid data

Example

$$100 \times \frac{7}{12} = 58.3\%$$

## Accuracy for each appraiser

The formula is

$$100 \times \frac{\text{number of appraisals that match the standard value for the } i^{\text{th}} \text{ appraiser}}{N_i}$$

Where

- $N_i$  is the number of appraisals for the  $i^{\text{th}}$  appraiser

Example (accuracy for appraiser 1)

$$100 \times \frac{5}{6} = 83.3\%$$

## Accuracy by standard

The formula is

$$100 \times \frac{\text{number of appraisals that match the } i^{\text{th}} \text{ standard value}}{N_i}$$

Where

$N_i$  is the number of appraisals for the  $i^{\text{th}}$  standard value

Example (accuracy for "good" items)

$$100 \times \frac{3}{4} = 75\%$$

## Accuracy by trial

The formula is

$$100 \times \frac{\text{number of appraisals that match the standard value for the } i^{\text{th}} \text{ trial}}{N_i}$$

Where

$N_i$  is the number of appraisals for the  $i^{\text{th}}$  trial

Example (trial 1)

$$100 \times \frac{3}{6} = 50\%$$

## Accuracy by appraiser and standard

The formula is

$$100 \times \frac{\text{number of appraisals for the } i^{\text{th}} \text{ appraiser that match the } i^{\text{th}} \text{ standard}}{N_i}$$

Where

$N_i$  is the number of appraisals for the  $i^{\text{th}}$  appraiser for the  $i^{\text{th}}$  standard

Example (appraiser 2, standard "bad")

$$100 \times \frac{1}{4} = 25\%$$

## Misclassification rates

The overall error rate is

$$100 - \text{overall accuracy}$$

Example

$$100 - 58.3\% = 41.7\%$$

If appraisers rate a "good" item as "bad", the misclassification rate is

$$100 \times \frac{\text{Number of "good" items rated as "bad"}}{\text{Total number of "good" items}}$$

Example

$$100 \times \frac{1}{4} = 25\%$$

If appraisers rate a "bad" item as "good", the misclassification rate is

$$100 \times \frac{\text{Number of "bad" items rated as "good"}}{\text{Total number of "bad" items}}$$

Example

$$100 \times \frac{4}{8} = 50\%$$

If appraisers rate the same item both ways across multiple trials, the misclassification rate is

$$100 \times \frac{\text{Number of items rated inconsistently across trials}}{\text{Number of items} \times \text{Number of appraisers}}$$

Example

$$100 \times \frac{3}{3 \times 2} = 50\%$$

## Appraiser misclassification rates

If appraiser  $i$  rates a "good" item as "bad", the misclassification rate is

$$100 \times \frac{\text{Number of "good" items rated as "bad" by appraiser } i}{\text{Total number of "good" items rated by appraiser } i}$$

Example (for appraiser 1)

$$100 \times \frac{0}{2} = 0\%$$

If appraiser  $i$  rates a "bad" item as "good", the misclassification rate is

$$100 \times \frac{\text{Number of "bad" items rated as "good" by appraiser } i}{\text{Total number of "bad" items by appraiser } i}$$

Example (for appraiser 1)

$$100 \times \frac{1}{4} = 25\%$$

If appraiser  $i$  rates the same item both ways across multiple trials, the misclassification rate is

$$100 \times \frac{\text{Number of items rated inconsistently across trials}}{\text{Total number of items rated by appraiser } i}$$

Example (for appraiser 1)

$$100 \times \frac{1}{3} = 33.3\%$$

## Most frequently misclassified items

%good rated "bad" for the  $i^{\text{th}}$  "good" item is

$$100 \times \frac{\text{Number of appraisals for the } i^{\text{th}} \text{ "good" item that were rated "bad"}}{\text{Number of appraisals for the } i^{\text{th}} \text{ "good" item}}$$

Example (item 1)

$$100 \times \frac{1}{4} = 25\%$$

%bad rated "good" for the  $i^{\text{th}}$  "bad" item is

$$100 \times \frac{\text{Number of appraisals for the } i^{\text{th}} \text{ "bad" item that were rated "good"}}{\text{Number of appraisals for the } i^{\text{th}} \text{ "bad" item}}$$

Example (item 2)

$$100 \times \frac{2}{4} = 50\%$$

© 2015, 2017 Minitab Inc. All rights reserved.

Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See [minitab.com/legal/trademarks](http://minitab.com/legal/trademarks) for more information.