Measurement 46 (2013) 3710–3714

Contents lists available at SciVerse ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Viewpoints to characterize precision evaluation methods in binary measurements

Tomomichi Suzuki [a,*], Yusuke Tsutsumi [b], Hironobu Kawamura [c]

[a] Tokyo University of Science, Chiba, Japan
[b] Mistubishi Tanabe Pharma Corporation, Tokyo, Japan
[c] Tsukuba University, Ibaraki, Japan

ARTICLE INFO

Article history:
Available online 12 June 2013

Keywords:
Precision
Repeatability
Reproducibility
Binary data
ISO 5725

ABSTRACT

Precision evaluation in quantitative measurements is a thoroughly discussed topic and the established methods are in use. Many methods are proposed for qualitative data including binary data, but their effectiveness and statistical properties are not so clear.

This paper first reviews and examines the current status of the existing methods for evaluating precision in binary measurements from statistical point of view. This paper then proposes the viewpoints to characterize those methods and the methods are characterized using the proposed viewpoints.

1. Introduction

Precision evaluation in quantitative measurements is a thoroughly discussed topic and the established methods are in use. ISO published ISO 5725 *accuracy* (*trueness and precision*) *of measurement methods and results* which deals with quantitative measurements. Now there is a movement to develop a document for non-quantitative measurements.

Many methods are proposed for qualitative data including binary data, but their effectiveness and statistical properties are not so clear. This paper first reviews and examines the current status of the existing methods for evaluating precision in binary measurements from statistical point of view. This paper then proposes the viewpoints to characterize those methods and the methods are characterized using the proposed viewpoints.

2. Precision for quantitative data

ISO 5725 [1] uses two terms trueness and precision to describe the accuracy of a measurement method. Trueness refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value. Precision refers to the closeness of agreement between test results.

ISO 5725 defines the precision using the model shown in the following equation:

$$y = m + B + e \tag{1}$$

where y is the measurement result, m is the general mean (expectation), B is the laboratory component of bias under repeatability conditions, and e is the random error in every measurement under repeatability conditions. The term B is considered a random variable whose expectation equals zero and whose variance is expressed as σ_L^2. The term e is a random variable whose expectation equals zero and whose variance is expressed as σ_e^2.

ISO 5725 [1] also introduces two components of precision: repeatability and reproducibility.

Repeatability is the precision under repeatability conditions. Repeatability conditions are defined as "conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time". Repeatability indicates the smallest variation for a particular measurement method.

* Corresponding author. Address: 2641 Yamazaki, Noda, Chiba 2788510, Japan. Tel.: +81 4 7122 9791; fax: +81 4 7122 4566.
E-mail address: szk@rs.tus.ac.jp (T. Suzuki).

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.05.032

Reproducibility is the precision under reproducibility conditions. Reproducibility conditions are defined as "conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment". Reproducibility indicates the largest variation for a particular measurement method.

Using the notation shown in Eq. (1), the repeatability variance σ_r^2 and reproducibility variance σ_R^2 can be expressed as follows.

$$\sigma_r^2 = \sigma_e^2 \tag{2}$$

$$\sigma_R^2 = \sigma_L^2 + \sigma_r^2 = \sigma_L^2 + \sigma_e^2 \tag{3}$$

The estimates of repeatability variance and reproducibility variance are calculated from interlaboratory studies or collaborative assessment experiments [2].

3. Binary data

There are many types of non-quantitative measurements. The types of data vary accordingly, for example, binary data, categorical data, ordinal data, etc. In this paper, the methods to evaluate precision for binary data are considered. The data format of binary data is shown in Table 1. Every laboratory measures the identical test item number of times. The number of laboratories is denoted as p, and the number of replication is denoted as n. The data for each item or level can be expressed as Table 1. If the study is conducted on multiple items or levels, or in other words, when the number of items or levels is large, there will be many sets of data in the format of Table 1. In the binary data case, the value of y_{ik} is either 0 (negative, non-detect, fail, etc.) or 1 (positive, detect, pass, etc.).

The methods to evaluate precision for binary data are presented in the following chapter.

4. Binary methods

This chapter describes the methods proposed to evaluate precision for binary data.

4.1. ISO based method

This section describes the method proposed by Wilrich [3,4]. This method is based on the approach of ISO 5725 part 2. Treating the measurement values (zero and one) as the same as ISO 5725 in analysis (ANOVA, etc.).

Table 1
Data format for binary data.

Laboratory	Run 1	...	Run k	...	Run n
1	y_{11}		y_{1k}		y_{1n}
2	y_{21}		y_{2k}		y_{2n}
$\vdots$					
$\vdots$					
i	y_{i1}		y_{ik}		y_{in}
$\vdots$					
$\vdots$					
p	y_{p1}		y_{pk}		y_{pn}

The model adopted by Wilrich is shown below. It is similar to the ISO 5725 model.

$$y_{ij} = \pi + (\pi_i - \pi) + e_{ij} \tag{4}$$

where π_i denotes the probability of obtaining the positive result (value = "1") for laboratory number i, and π denotes the overall probability of obtaining the positive result, which can be shown as

$$\pi = \sum_{i=1}^{p} \pi_i / p \tag{5}$$

In the model given by Eq. (4), π_i is a random variable which takes a specific value in each laboratory and its expectation is equal to π. y_{ij} are independently Bernoulli distributed with expectation π_i.

Eq. (4) is analogous to Eq. (1) as the first term expresses the general mean, the second term the laboratory component of bias, and the third term the error. One big difference between Eqs. (1) and (4) is that, values of the variables in Eq. (4) have constraints. In particular, y_{ij}s take value of either 0 or 1, and π_i and π only take value between 0 and 1. Another big difference is the distribution for the error term. It is adequate to assume normal distribution in Eq. (1), but it is difficult to assume a reasonable distribution in Eq. (4).

The basic idea behind the method is ISO 5725 part 2. However, it is not totally clear that the same procedure is equally effective in the binary data. Uhlig et al. [5] point out that Wilrich's estimate for reproducibility variance for binary data is not adequate because the estimated value does not depend on the variability between laboratories. This is quite the opposite of the ISO definition for reproducibility variance for quantitative data which highly depends on variability between laboratories. Horie et al. [6] showed that the same variances as above are obtained when beta-binomial distribution (which is one of reasonable assumptions) is assumed, whose results support the ISO method.

4.2. Accordance and concordance

This section describes the method proposed by Langton et al. [7]. This method introduces two new concepts of precision, accordance and concordance. Accordance is a concept similar to repeatability and concordance is a concept similar to reproducibility.

Accordance is the percentage of pairs of measurements that give the same result within a laboratory. Concordance is the percentage of pairs of measurements that give the same between different laboratories.

Accordance for a specific laboratory i is calculated by

$$A_i = \frac{k_i(k_i - 1) + (n - k_i)(n - k_i - 1)}{n(n-1)}, \tag{6}$$

where n represents number of runs and k_i number of 1s for laboratory i.

Overall accordance (denoted as A) is the average of accordance of all laboratories.