

Understanding PT statistics

Introduction

The Eurachem Guide on “Selection, Use and Interpretation of Proficiency Testing (PT) Schemes” [1] recommends participants to consider the statistical approach used by the PT provider when selecting a PT scheme. This leaflet is intended to help participants in quantitative PT schemes to better understand the statistical parameters that describe the data distribution in the PT report. The PT report should also provide information on how the assigned value (x_{pt}), the corresponding uncertainty ($u(x_{pt})$) and the standard deviation for proficiency assessment (σ_{pt}) were obtained [2].

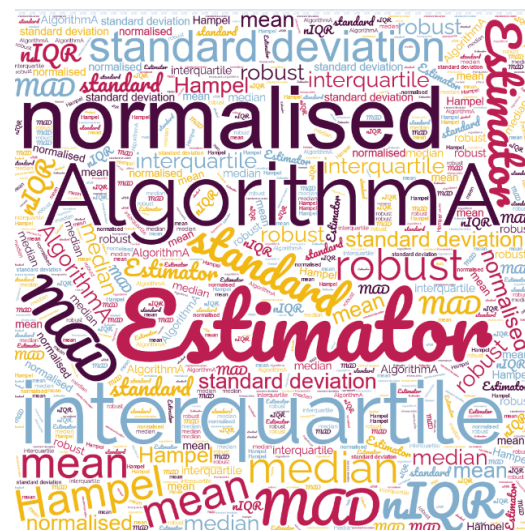
Summary statistics reported in PT schemes typically describe two features of a data set or of an underlying normal distribution of technically valid results. One feature is the **location**, usually described by some kind of average. The second is the **dispersion**, or spread of the data. A further important summary statistic is the (standard) **uncertainty associated with the location**, which can help to determine whether the chosen location estimator differs from (for example) an external reference value or from individual reported results. The different statistics used for these purposes are summarised below.

Location of the results reported by participants

An estimated location is usually reported even if it is not used as the assigned value. ISO 13528 [3] describes different ways to estimate the location using the arithmetic mean or median, or applying sophisticated robust statistical methods (Table 1).

Key points to note:

- When data sets contain outliers or when symmetry of the data set cannot be verified, ISO 13528 strongly recommends to use robust estimators.
- When using the arithmetic mean, the PT provider would normally have checked the data set and removed any outliers.
- All estimates of location can be unreliable for data sets with a small number of results.
- A significant difference between the estimated location and a reference value may indicate a bias of the analytical methods (maybe just one, some or for all methods) that have been used by the participants.
- If the participants used different analytical methods, the PT provider might report different estimated locations for each of these methods.



Dispersion of results reported by participants

The standard deviation is an estimate of dispersion of a data set. There are various possibilities to estimate the dispersion of data including MAD_e , $nIQR$, Algorithm A, Q_n or Q method (Table 1). Often, they are normalised to allow direct comparability to the standard deviation of a normally distributed data set.

Key points to note:

- The dispersion is expressed in the same units as the measurand or reported as a percentage.
- When estimating the classical standard deviation, the PT provider would normally have checked the data set and removed any outliers.
- All estimates of dispersion can be unreliable for data sets with a small number of results.



Eurachem

A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE

Table 1 - Pros (😊) & cons (😞) of the various statistical approaches

Location	Arithmetic mean	Median (robust mean)	--	--	Algorithm A (robust mean)	Hampel estimator (robust mean)
Dispersion	Classical standard deviation (s)	--	Scaled median absolute deviation (MADe)	Normalized interquartile range (nIQR)	Robust standard deviation (s*)	Robust standard deviation Q_n or $Q(s^*)$
Simple	😊 Yes	😊 Yes	😊 Yes	😊 Yes	😞 No	😞 No
Applies to	😞 Normally distributed data	😊 All data	😞 Normally distributed data (1)	😞 Normally distributed data (1)	😊 All data	😊 All data
Robust against outliers	😞 No	😊 Yes	😊 Yes	😊 Yes	😊 Yes	😊 Yes
Resistance to minor modes, RMM (2)	😞 Poor	😊 Good	😊/😊 Moderate/Good	😊 Moderate	😊 Moderate	😊/😊 Moderate/Good
Efficiency, Eff (3)	😊 High	😊 Moderate	😞 Low	😞 Low	😊 High	😊 High
Breakdown point, BP(4)	😞 Very low (5)	😊 High	😊 High	😊 Moderate	😊 Moderate	😊 High

(1) The factors used to convert MAD to MADe and IQR to nIQR assume a normally distributed dataset

(2) RMM - Ability of an estimator to resist the bias caused by a minority group of discrepant results.

(3) For normal distribution, Eff = estimator variance / variance of the mean;

Note: Higher efficiency usually means a lower uncertainty of the estimate.

(4) BP - Proportion of values that can be outliers without the estimate being adversely affected.

(5) Outlier test required.

Standard uncertainty of the location

The standard uncertainty of the location is a measure of the reliability of the mean that can be important for participants.

Key points to note:

- The standard uncertainty of the arithmetic mean is calculated as $u(\bar{x}) = s/\sqrt{p}$, where s is the classical standard deviation, and p the number of reported results.
- For other estimators of the mean ISO 13528 suggests to use an additional (conservative) multiplication factor set to 1.25, hence $u(\bar{x}) = 1.25 s^*/\sqrt{p}$, where s^* is the robust standard deviation from MADe, nIQR, Algorithm A, Q_n or Q method, and p the number of reported results.

More information / further reading

[1] B. Brookman and I. Mann (eds.) Eurachem Guide: Selection, Use and Interpretation of Proficiency Testing (PT) Schemes (3rd ed. 2021). Available from www.eurachem.org.

[2] Eurachem leaflet "Understanding PT performance assessment". Available from www.eurachem.org.

[3] ISO 13528:2022 - Statistical methods for use in proficiency testing by interlaboratory comparison.

Information about PT providers and schemes can be obtained from your national accreditation body, from the EPTIS website (www.eptis.org) or from other national or international organizations.