Statistik in Eignungsprüfungen verstehen

Einführung

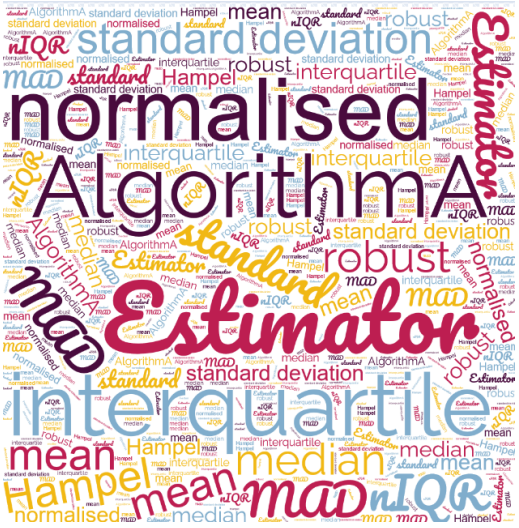
Der Eurachem-Leitfaden "Selection, Use and Interpretation of Proficiency Testing (PT) Schemes" [1] empfiehlt den Teilnehmern, bei der Auswahl eines Eignungsprüfungs-(EP-)Programms den vom Anbieter verwendeten statistischen Ansatz zu berücksichtigen. Dieses Merkblatt soll den Teilnehmern an quantitativen EP-Programmen helfen, die statistischen Parameter, die die Datenverteilung im EP-Bericht beschreiben, besser zu verstehen. Der EP- Bericht sollte auch Informationen darüber enthalten, wie der zugewiesene Wert (x_{pt}), die entsprechende Unsicherheit ($u\,(x_{pt})$) und die Standardabweichung für die Eignungsbewertung (σ_{pt}) ermittelt wurden [2].

Die in den EP angegebenen statistischen Kennwerte beschreiben in der Regel zwei Merkmale eines Datensatzes oder einer zugrundeliegenden Normalverteilung von technisch gültigen Ergebnissen. Ein Merkmal ist die Lage, die in der Regel durch eine Art von Durchschnitt beschrieben wird. Das zweite Merkmal ist die Dispersion oder Streuung der Daten. Eine weiterer wichtiger statistischer Kennwert ist die mit der Lage verbundene (Standard-)Unsicherheit, die dabei helfen kann, festzustellen, ob der gewählte Lageschätzer von (beispielsweise) einem externen Referenzwert oder von einzelnen gemeldeten Ergebnissen abweicht. Die verschiedenen statistischen Kennwerte, die für diese Zwecke verwendet werden, sind im Folgenden zusammengefasst.

Lage der von den Teilnehmern gemeldeten Ergebnisse

Eine geschätzte Lage wird in der Regel angegeben, auch wenn sie nicht als zugewiesener Wert verwendet wird. ISO 13528 [3] beschreibt verschiedene Möglichkeiten zur Schätzung der Lage unter Verwendung des arithmetischen Mittels oder des Medians, oder unter Anwendung anspruchsvoller robuster statistischer Methoden (Tabelle 1).

Die wichtigsten Punkte sind:



- Wenn Datensätze Ausreißer enthalten oder wenn die Symmetrie des Datensatzes nicht verifiziert werden kann, empfiehlt ISO 13528 dringend die Verwendung robuster Schätzer.
- Bei der Verwendung des arithmetischen Mittels müsste der EP-Anbieter normalerweise den Datensatz überprüft und Ausreißer entfernt haben.
- Alle Schätzungen der Lage können bei Datensätzen mit einer geringen Anzahl von Ergebnissen unzuverlässig sein.
- Ein signifikanter Unterschied zwischen der geschätzten Lage und einem Referenzwert kann auf einen Bias der von den Teilnehmern verwendeten Analysemethoden (vielleicht nur einer, einiger oder aller Methoden) hinweisen.
- Wenn die Teilnehmer unterschiedliche Analysenverfahren verwenden, kann der EP-Anbieter für jedes dieser Verfahren unterschiedliche geschätzte Lagen angeben.

Streuung der von den Teilnehmern gemeldeten Ergebnisse

Die Standardabweichung ist eine Schätzung der Streuung eines Datensatzes. Es gibt verschiedene Möglichkeiten, die Streuung von Daten zu schätzen, darunter MADe, nIQR, Algorithmus A, Qn oder Q-Methode (Tabelle 1). Häufig werden sie normalisiert, um eine direkte Vergleichbarkeit mit der Standardabweichung eines normalverteilten Datensatzes zu ermöglichen.

Eurachem
A FOCUS FOR ANALYTICAL CHEMISTRY IN EUROPE

Die wichtigsten Punkte sind:

- Die Streuung wird in denselben Einheiten wie die Messgröße ausgedrückt oder als Prozentsatz angegeben.
- Bei der Schätzung der klassischen Standardabweichung müsste der EP-Anbieter normalerweise den Datensatz überprüft und Ausreißer entfernt haben.
- Alle Schätzungen der Streuung können bei Datensätzen mit einer geringen Anzahl von Ergebnissen unzuverlässig sein.

Tabelle 1 - Vorteile (☺) und Nachteile (☹) der verschiedenen statistischen Ansätze

Lage	**Arithmetischer Mittelwert**	**Median** (robuster Mittelwert)	--	--	**Algorithmus A** (robuster Mittelwert)	**Hampelschätzer** (robuster Mittelwert)
Streuung	**Klassische Standard-abweichung (s)**	--	**Skalierter Median der absoluten Abweichungen (MADe)**	**Normierter Interquartil-abstand (nIQR)**	**Robuste Standard-abweichung (s*)**	**Robuste Standard-abweichung Q_n oder Q (s*)**
Einfach	☺ Ja	☺ Ja	☺ Ja	☺ Ja	☹ Nein	☹ Nein
Anwendbar für	☹ Normalverteilte Daten	☺ Alle Daten	☹ Normalverteilte Daten (1)	☹ Normalverteilte Daten (1)	☺ Alle Daten	☺ Alle Daten
Robust gegen Ausreißer	☹ Nein	☺ Ja	☺ Ja	☺ Ja	☺ Ja	☺ Ja
Resistenz gegen kleinere Modi , RMM (2)	☹ Schlecht	☺ Gut	😐/☺ Moderat/Gut	😐 Moderat	😐 Moderat	😐/☺ Moderat/Gut
Effizienz, Eff (3)	☺ Hoch	😐 Moderat	☹ Gering	☹ Gering	☺ Hoch	☺ Hoch
Bruchpunkt, BP(4)	☹ Sehr niedrig (5)	☺ Hoch	☺ Hoch	😐 Moderat	😐 Moderat	☺ Hoch

(1) Die Faktoren, die zur Umrechnung von MAD in MADe und IQR in nIQR verwendet werden, gehen von einem normalverteilten Datensatz aus.
(2) RMM - Fähigkeit eines Schätzers, dem Bias zu widerstehen, der durch eine kleine Gruppe von abweichenden Ergebnissen verursacht wird.
(3) Für die Normalverteilung gilt: Eff = Varianz des Schätzers / Varianz des Mittelwerts;
Anmerkung: Eine höhere Effizienz bedeutet in der Regel eine geringere Unsicherheit der Schätzung.
(4) BP - Anteil der Werte, die Ausreißer sein können, ohne daß die Schätzung beeinträchtigt wird.
(5) Ausreißertest erforderlich.

Standardunsicherheit der Lage

Die Standardunsicherheit der Lage ist ein Maß für die Zuverlässigkeit des Mittelwerts, das für die Teilnehmer wichtig sein kann.

Die wichtigsten Punkte sind:

- Die Standardunsicherheit des arithmetischen Mittels wird berechnet als $u(\bar{x}) = s/\sqrt{p}$, wobei s die klassische Standardabweichung und p die Anzahl der berichteten Ergebnisse ist.
- Für andere Schätzer des Mittelwerts schlägt ISO 13528 die Verwendung eines zusätzlichen (konservativen) Multiplikationsfaktors von 1,25 vor, $u(\bar{x}) = 1.25\, s^*/\sqrt{p}$, wobei s^* die robuste Standardabweichung berechnet mit MADe, nIQR, Algorithmus A, Q_n oder Q-Methode und p die Anzahl der berichteten Ergebnisse ist.

Mehr Informationen / weiterführende Literatur

[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. International Standardization Organization, Geneva, CH.

Informationen über Eignungsprüfungsveranstalter und –systeme erhalten Sie von Ihrer nationalen Akkreditierungsstelle, von der EPTIS-Webseite (www.eptis.org) oder von anderen nationalen oder internationalen Organisationen.

Produced by the Eurachem Proficiency Testing Working Group - www.eurachem.org
First German edition, March 2025
Translation and adaptation into German by EUROLAB-DE

Subscribe to Eurachem e-News at https://www.eurachem.org/subscribe