Wiederholte Messungen zur Verbesserung der Standardunsicherheit

Einleitung

Die durch zufällige Effekte bedingte Standardunsicherheit wird häufig aus wiederholten Experimenten abgeleitet und durch die Standardabweichung *s* des gemessenen Größenwerts quantifiziert. Wenn die aus zufälligen Effekten resultierende Standardunsicherheit einer Einzelmessung der Größe benötigt wird, ist dies einfach die beobachtete Standardabweichung *s*; wenn aber das Ergebnis der Mittelwert $\bar{x}$ aus *n* Messungen ist, dann wird die Standardunsicherheit $u_{\bar{x}}$ umso kleiner je größer n wird. Damit ist:

$$u_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad \text{Gl. (1)}$$

Beispiel 1 zeigt die Anwendung von Gl. (1) zur Schätzung der Unsicherheit des Mittelwerts, sie gilt aber nicht, wenn die Unsicherheit einzelner Beobachtungen betrachtet wird.

> Beispiel 1
>
> Eine Pipette wird mit 12 Messungen kalibriert; der Mittelwert und die Standardabweichung werden berechnet. Gl. 1 beschreibt die Schätzung der Standardunsicherheit des Mittelwerts. Wenn die Pipette jedoch dazu *benutzt* wird, nur *ein* Aliquot zu liefern, ist Gl. (1) nicht anwendbar, und die Standardunsicherheit aus zufälligen Schwankungen für diese einzelne Messung ist die Standardabweichung *s*.

Damit Gl. (1) gilt, ist es erforderlich, dass alle Messungen unabhängig sind und aus einer stabilen Probe unter den **gleichen** Messbedingungen ermittelt werden. Bedingungen für alle durchgeführten Messungen können zum Beispiel sein, 1) Wiederholbedingungen, 2) Vergleichbedingungen (Reproduzierbarkeit innerhalb des Labors) oder 3) erweiterte Vergleichbedingungen (zwischen Laboratorien).

Es ist sehr wichtig, sich klar zu machen, dass die in Gl. (1) angegebene Standardunsicherheit nur die geschätzte Unsicherheit aus Zufallsschwankungen unter den Bedingungen liefert, unter denen die Beobachtungen erfolgt sind und dass sie strikt nur für voneinander unabhängige Beobachtungen gilt.

Manchmal ist es schwer zu unterscheiden, ob es sich um hinreichend unabhängige Beobachtungen handelt und Gl. (1) angewendet werden darf, denn es gibt keine einfache Regel. Die in den folgenden Abschnitten beschriebenen Beispiele sollen helfen, die Fälle zu identifizieren, in denen Gl. (1) sicher angewendet werden kann.

Beispiel für die Anwendung von Gleichung (1)

Messungen inhomogener Proben

Macht die Inhomogenität der Proben einen erheblichen Teil der Unsicherheit aus, kann der Analytiker entscheiden, eine größere Anzahl von Teilproben des Untersuchungsmaterials zu messen, um die Standardunsicherheit zu reduzieren. Wenn alle diese Messungen unter Wiederholbedingungen gemacht werden, d. h. unter denselben Messbedingungen über das gesamte Verfahren einschließlich der zufälligen Auswahl von Teilproben, sollte die Standardabweichung des Mittelwerts nach Gl. (1) berechnet werden, um die Unsicherheit zu ermitteln, die aus Zufallsschwankungen unter Wiederholbedingungen resultiert.

Beispiele, in denen Gleichung (1) nicht angewendet werden darf

Die folgenden Abschnitte beschreiben zwei Beispiele, in denen weder die Standardabweichung noch die Standardabweichung des Mittelwerts ohne weitere Datenanalyse direkt verwendet werden dürfen.

Messung gruppierter Werte

Ein Beispiel hierfür sind Daten aus der internen Qualitätskontrolle eines Messverfahrens, das eine messtägliche Kalibrierung beinhaltet. Wir möchten die Standardunsicherheit des Mittelwerts berechnen, der verwendet wird, um die Zentrallinie in einer Qualitätsregelkarte festzulegen. Bei den Daten handelt es sich um Doppelmessungen zur Qualitätskontrolle, die jeden Tag von einer stabilen Probe über einen längeren Zeitraum (z. B. p Tage) erhalten wurden, so dass wir insgesamt $2p$ Beobachtungen, d. h. p Gruppen zu je 2 Messungen haben. Jedes Paar von Doppelbestimmungen hat dieselbe Kalibrierung gemeinsam, die Doppelwerte innerhalb des Datensatzes sind daher nicht strikt unabhängig voneinander und Gl. (1) kann nicht direkt für alle $2p$ Beobachtungen angewendet werden. Die Unsicherheit des Mittelwerts kann einfach berechnet werden, indem man die Standardabweichung der Mittelwerte für jeden Tag durch $\sqrt{p}$ dividiert. Varianzanalyse kann in ähnlichen Fällen hilfreich sein. Vergleichbare Prinzipien gelten für andere Arten der Gruppierung, z. B. hinsichtlich der Analytiker, der Instrumente, u. s. w.

Messungen, bei denen die Probe oder das Messsystem zeitlich nicht stabil ist

Ein anderes gängiges Beispiel sind zeitabhängige Daten. Die Zeitabhängigkeit könnte durch eine Gerätedrift oder eine tatsächliche Konzentrationsveränderung im Zeitverlauf verursacht sein. In diesen Fällen ist die Abweichung einer bestimmten Beobachtung teilweise zufällig und teilweise ‚übertragen' aus vorangegangenen Messungen. Wiederum sind die Abweichungen, die jede Beobachtung beeinflussen, nicht voneinander unabhängig, weil ein Teil der Abweichung sukzessiven Beobachtungen gemeinsam ist; Gl. (1) kann nicht verwendet werden und komplexere statistische Techniken, die eine Korrelation erlauben, sind zur Verarbeitung solcher Daten einzusetzen.

Weitere Informationen zur Behandlung korrelierter Daten bei der Unsicherheitsabschätzung finden Sie in Eurolab Technical Report 1/2006: Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results, Appendix A.5 (www.eurolab.org).

Verfasst durch die Eurachem/CITAC
Measurement Uncertainty and Traceability Working Group, First English edition 2015
übersetzt ins Deutsche durch Eurolab-D, 2015