Fehlerfortpflanzung

Sehr häufig werden mehrere Mittelwerte von Stichproben miteinander verrechnet. Besonders häufig geschieht das in Quotienten, wie dem Phosphorgehalt des Sestons / der Trockenmasse oder dem zellspezifischen Chlorophyllgehalt und alle prozentualen Anteile von chemischen oder Größenfraktionen. Wenn Zähler und Nenner an unabhängigen Stichproben ermittelt werden, z. B. aus 3 Unterproben der Chlorophyllgehalt und aus 3 weiteren werden die Zellen gezählt, gehen die beiden Standardabweichungen in eine Fehlerfortpflanzung ein. Manchmal werden jedoch beide Parameter tatsächlich an separaten Unterproben einander zugeordnet gemessen, z. B. der C- und N-Gehalt in 3 separaten Filterrückständen. Dann werden die Quotienten zuerst gebildet und dann Mittelwert sowie Standardabweichung des C:N-Verhältnisses berechnet.

Fehlerfortpflanzung in Summen entsteht oft durch die Addition mehrerer Standardabweichungen (auch Unsicherheiten genannt). So werden verschiedene Fehlerquellen im Qualitätsmanagement summiert, wie die Reproduzierbarkeit einer Probe, von Proben in einer Messserie oder zwischen Messserien.

Differenzen und Produkte sind in der Umweltanalytik und aquatischen Ökologie etwas seltener. So werden Differenzen gebildet, wenn verschiedene Filtrate (Vollprobe, 200, 20 und 2 µm Filtrat) gemessen und voneinander abgezogen werden, um die Fraktionen >200, 20-200, 2-20 und <2 µm zu erhalten. Produkte werden für gewichtete Mittel oder zur Berechnung von Stoffmengen aus Konzentration und Volumen gebildet. Sehr häufig hat dann aber ein Faktor keine Standardabweichung, hier die mittlere Größenklasse und das Volumen. Dann muss und kann keine Fehlerfortpflanzung gerechnet werden.

Normalverteilung

- Mittelwert und Standardabweichung charakterisieren Lage und Streuung von Messwerten einer normal verteilten Stichprobe.
- Standardabweichung eines arithmetrischen Mittelwertes $\bar{x}_i$ als Streuungsmaß metrischer Messwerte berechnet aus der Summe der Abweichungsquadrate normiert über die Freiheitsgrade
- Diese Parameter dürfen nicht für andere, insbesondere schiefe, Verteilungen berechnet werden.
- Die Kombination von Median und Standardabweichung ist nicht richtig.
- Üblicherweise wird angenommen, dass Mehrfachmessungen z. B. einer Konzentration normal verteilt sind.
- Die Standardabweichung kann über den Mittelwert normiert (Variationskoeffizient in %) werden, so dass Streuungen großer und kleiner Werte miteinander verglichen werden können.

Standardabweichung $$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

Variationskoeffizient $$Vk = \frac{s}{\bar{x}} \quad \text{oder} \quad Vk' = \frac{s}{\bar{x}} \cdot 100\%$$

Fehlerfortpflanzung nach Doerffel

- zur Berechnung der Standardabweichung von Summen, Differenzen, Produkten und Quotienten zweier unabhängiger Einzelergebnisse x_1 und x_2
- unter der Voraussetzung der stochastischen Unabhängigkeit der Einzelergebnisse

Summe $$\bar{x}_1 + \bar{x}_2 \pm \sqrt{{s_{\bar{x}_1}}^2 + {s_{\bar{x}_2}}^2}$$

Differenz $$\bar{x}_1 - \bar{x}_2 \pm \sqrt{{s_{\bar{x}_1}}^2 + {s_{\bar{x}_2}}^2}$$

Produkt $$\bar{x}_1 \cdot \bar{x}_2 \pm \sqrt{{\bar{x}_1}^2 \cdot {s_{\bar{x}_2}}^2 + {\bar{x}_2}^2 \cdot {s_{\bar{x}_1}}^2}$$

Quotient $$\frac{\bar{x}_1}{\bar{x}_2} \pm \frac{\sqrt{{\bar{x}_1}^2 \cdot {s_{\bar{x}_2}}^2 + {\bar{x}_2}^2 \cdot {s_{\bar{x}_1}}^2}}{{\bar{x}_2}^2}$$

Weitere statistische Kenngrößen zur Fehlereinschätzung

- ohne Kenntnis der Verteilungsfunktion ist Standardfehler des Mittelwertes schätzbar

Standardfehler $$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- für Zellzählungen (Poisson-Verteilung) gilt

$$\mu_u = \left(\frac{c}{2} - \sqrt{n}\right)^2$$ μ_u = untere Vertrauensgrenze

$$\mu_o = \left(\frac{c}{2} + \sqrt{n+1}\right)^2$$ μ_o = obere Vertrauensgrenze

n = gezählte Zellen
c = 1,96 für eine Irrtumswahrscheinlichkeit von 95%

- und daraus näherungsweise der prozentuale Zählfehler Vk'

$$Vk' = \pm \frac{1.96}{\sqrt{n}} \cdot 100\%$$

Unbekannte bzw. parameterfreie Verteilungen

Für andere Verteilungen gibt es spezifische Streuungskenngrößen. Ganz allgemeingültig sind die Percentilen einer Verteilung, wie sie in Box-Whisker-Plots.

- Der Median ist der Lageparameter der Verteilung. Für eine Normalverteilung stimmen arithmetrisches Mittel und Median überein.

- Die Box umfasst die Hälfte aller Werte um den Median herum. Das ist der Interquartilabstand ($Q_{25\%}$ bis $Q_{75\%}$).
- Die Whisker (Strecken) reichen (je nach Software, z.T. wählbar) von der 5 oder 10%-Percentile bis zur 90 bzw. 95%-Percentile.
- Ausreißer können als Symbole angezeigt werden. Nur Werte, die einen bestimmten Abstand zum Median überschreiten, sind Ausreißer.