

Forderung nach Varianzenhomogenität – ein Ärgernis?

Gerhard Wachter und Joachim Kleiner

Varianzenhomogenität wird in allen aktuellen Regelwerken und deren Vorgängerversionen als Voraussetzung zur Kalibration und deren Auswertung gefordert. Moderne analytische Messverfahren erlauben andererseits heute Kalibrationen über mehrere Dekaden, und in nicht wenigen Fällen scheint zumindest dort die Forderung nach Varianzenhomogenität in der praktischen Analytik problematisch, um nicht zu sagen ärgerlich. Es stellt sich grundsätzlich die Frage, wie wirkt sich die Missachtung der Varianzeninhomogenität bei der linearen Regression und der Auswertung der Probenmessung aus. Wie kann man also mit dieser Forderung umgehen?

Der Reststandardabweichung s_{yx} (vergleiche Formel(2)) kommt unter den Verfahrenskenndaten eine Schlüsselrolle zu. Sie repräsentiert in der Kalibration die mittlere Wiederholbarkeit der Informationswerte innerhalb eines gegebenen Merkmalswertebereichs (Arbeitsbereich; vergleiche Abbildung 1). Ihre Unabhängigkeit vom Probengehalt wird mit Hilfe des verkürzten Varianzenhomogenitätstestes festgestellt¹.

Die Vorschrift(en)

Unabhängig davon, ob man die aus dem Jahre 1986 stammende DIN 38 402 (Teil 51) oder die zuletzt im Jahre 2001 als „2nd Edition“ überarbeitete ISO-Norm 8466-2 konsultiert, immer wird gefordert, dass Konstanz der Wiederholbarkeit innerhalb des mit der Kalibration abgesteckten Arbeitsbereichs gegeben ist. Auch die DIN 32 645 fordert zur Berechnung der Nachweis-, Erfassungs- und Bestimmungsgrenzen nach der indirekten Methode Varianzenhomogenität.² Andererseits kennt der Praktiker genügend Fälle, bei denen eben diese Forderung unpraktikabel, um nicht zu sagen „unsinnig“ zu sein scheint. Als ein Beispiel von vielen sei die Atom-Emissions-Spektrometrie genannt, die doch „über mehrere Dekaden linear“ arbeitet und entsprechende Erwartungen an den an-



Wachter



Kleiner

Die Autoren

Gerhard Wachter war nach seinem Chemiestudium an der Universität Freiburg in der Technischen Schule des Bodenseewerks Perkin-Elmer über 20 Jahre verantwortlich für Kundenschulungen zur Instrumentellen Analytik. Zusammen mit Dr. Joachim Kleiner, ebenfalls langjähriger und ehemaliger Mitarbeiter an der Technischen Schule, ist er als Fachreferent für Statistische Qualitätssicherung in der beruflichen Fortbildung tätig. Dr. Joachim Kleiner studierte Mathematik und Physik an der Universität Konstanz.

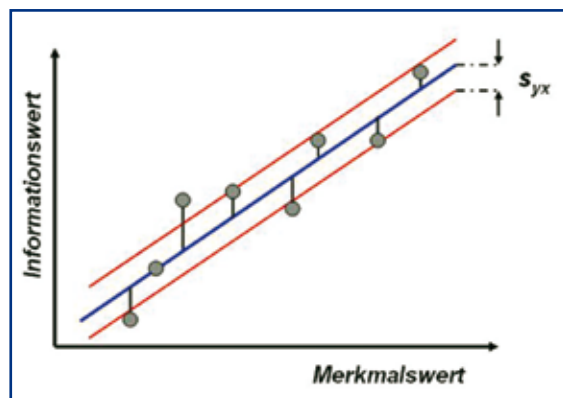


Abbildung 1: Reststandardabweichung als mittlere Streuung der Einzelwerte um die Ausgleichsfunktion.

wendbaren Arbeitsbereich fördert. Just dazu legen die oben genannten Regelwerke lapidar fest, dass bei Nichterfüllen der Varianzenhomogenität der Arbeitsbereich „dann möglichst soweit eingengt werden sollte, bis die Bedingung erfüllt ist“.³

Die Vorgehensweise

Nach DEV A51 wird der gewünschte Arbeitsbereich abgesteckt und es werden von den Merkmalsniveaus sowohl am unteren wie oberen Arbeitsbereichsende je zehn unabhängige Wiederholmessungen durchgeführt. Jeder dieser Datensätze wird mit Hilfe des Grubbs'schen Ausreißertestes auf Ausreißer geprüft und darf dem eigentlichen Test auf Varianzenhomogenität (F-Test) nur dann zu geführt werden, wenn er nicht mehr als einen zu eliminierenden Ausreißer enthält. Der Varinazen-F-Test wird mit 99%-iger statistischer Sicherheit durchgeführt.

Statistische Auswerteprogramme wie beispielsweise SQS⁴ bieten auf normenkonformen Arbeitsblättern eine bequeme Ausführung und Auswertung des geforderten Varianzenhomogenitäts-Testes. Die manuell eingegebenen oder importierten Rohdaten der beiden Datensätze werden automatisch auf Ausreißer getes-

¹ aus DIN 38 402, Teil 51: Begriffe (3.9)

² siehe DIN 32 645, (11) Mathematische Voraussetzungen

³ siehe DIN 38 402, Teil 51, 5.1.2 Überprüfung der Varianzenhomogenität

⁴ SQS ist ein Excel-Add-on-Programm zur statistischen Qualitätssicherung im analytischen Labor und wird vertrieben von Dr. Joachim Kleiner, individuelle Software & Schulung, D-78345 Moos

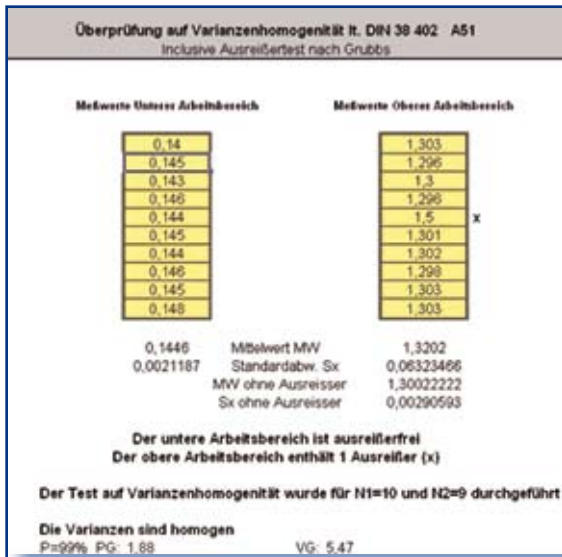


Abbildung 2: Überprüfen auf Varianzhomogenität nach DIN 38 402 (A51) mit Hilfe von SQS.

tet und die ausreißerfreien Datensätze über die F-Statistik miteinander verglichen. Das in Worte gefasste Ergebnis wird mit Zusatzangaben wie Datenumfänge (N_1 und N_2), Prüf- (PG) und Vergleichsgröße (VG) ergänzt. Außerdem können eigene Bemerkungen des Anwenders zur Transparenz der Prüfung dienen, die zusammen mit den Rohdaten und deren Auswertung als Datei oder als Ausdruck gesichert werden können.

Wird Varianzhomogenität bestätigt, ist der Weg frei für eine Kalibration über den abgesteckten Bereich. Bei Varianzeninhomogenität sehen die genannten Regelwerke, wie bereits erwähnt, eine Einschränkung des Arbeitsbereichs vor. Eine gewichtete Regression wie sie als Alternative diskutiert wird, ist in keinem der einschlägigen Regelwerke derzeit beschrieben oder gefordert. Außerdem ist eine solche Kalibration mit einem nicht zu vernachlässigenden Mehraufwand verbunden. Der Praktiker ist daher geneigt, auf die Erfüllung der Varianzhomogenität zu verzichten, nicht selten „mit einem gewissen Unwohlsein“. Die Frage stellt sich dann: Worin liegt eigentlich die Begründung am Festhalten der Forderung nach Varianzhomogenität?

Die Begründung

Kern einer statistischen Behandlung von Verfahren ist die Gewinnung der Verfahrenskennwerte, an Hand derer nicht nur das Verfahren in seiner Leistungsfähigkeit beschrieben und dokumentiert, sondern auch verglichen und beurteilt werden kann. Schließlich dienen Verfahrenskennwerten auch zur Auswertung analytischer Messungen. Zu diesen Verfahrenskennwerten gehören unter anderem auch Vertrauens- und Prognoseband einer Regression (Kalibration). Über die Breite des Prognosebandes wird beispielsweise

die Aussagenschärfe, das heisst das Intervall für den „wahren“ Wert des gefundenen Merkmalswert (meist Konzentrations- oder Massenwert) bei vorgegebener statistischer Sicherheit gewonnen (vergleiche Abbildung 3; Auswertung von Analyseergebnissen nach DIN 38 402, Teil 51 beziehungsweise ISO 8466-1 und ISO 8466-2). In der heutigen Sichtweise spricht man von der Ergebnisunsicherheit der Konzentration bezüglich der Kalibrationsumrechnung des Informationswertes.

Die Berechnung des Prognosebereichs für den Merkmalswert x in einer linearen Regression ($y = b \cdot x + a$) basiert auf Formel (1).

Formel (1):

$$x_{\text{min,max}} = \frac{y - a}{b} \pm \frac{s_{yx}}{b} \cdot t(f, P) \cdot \sqrt{\frac{1}{N} + \frac{1}{M} + \frac{(y - \bar{y})^2}{b^2 \cdot \sum_{i=1}^N (x_i - \bar{x})^2}}$$

mit:

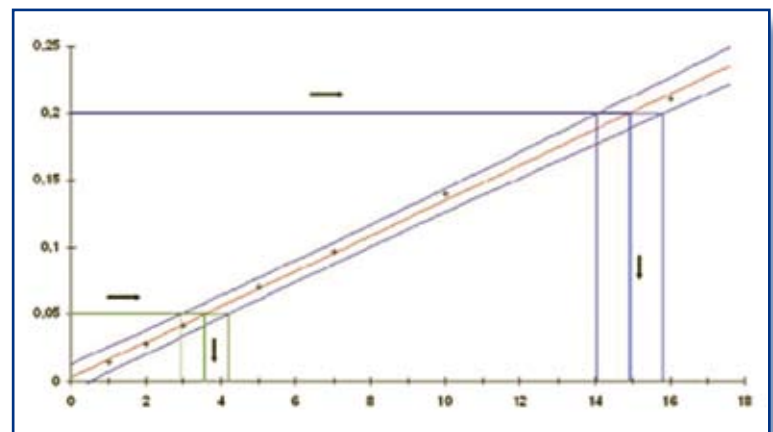
- y Informationswert (z.B. Extinktion in Peakhöhe)
- a Achsenabschnitt aus der Regressionsrechnung
- b Anstieg (Empfindlichkeit) aus der Regressionsrechnung
- s_{yx} Reststandardabweichung (siehe Formel (2))
- t(f;P) t-Wert (mit Freiheitsgrad f und stat. Sicherheit P)
- N Zahl der Standards
- M Zahl der Wiederholmessungen pro Standard

Formel (2):

$$s_{yx} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y})^2}{N - 2}}$$

Damit Formel (1) zur Berechnung des Prognosebereichs eines gefundenen Merkmalswertes eingesetzt werden kann, muss aber für den Arbeitsbereich x_1 bis x_n die Reststandardabweichung s_{yx} als konstant vorausgesetzt werden können, und nichts anderes wird über den Varianzhomogenitätstest geprüft. Dabei ist zu

Abbildung 3: Prinzip der Auswertung von Analyseergebnissen nach DIN 38 402, Teil 51 bzw. ISO 8466-1.



beachten, dass nicht die strikte numerische Gleichheit der Varianzen aus beiden Datensätzen gefordert ist. Gefordert wird, dass sich die beiden Varianzen mit 99%-iger statistischer Sicherheit nur zufällig und nicht signifikant unterscheiden.

Auswirkung einer Varianzeninhomogenität

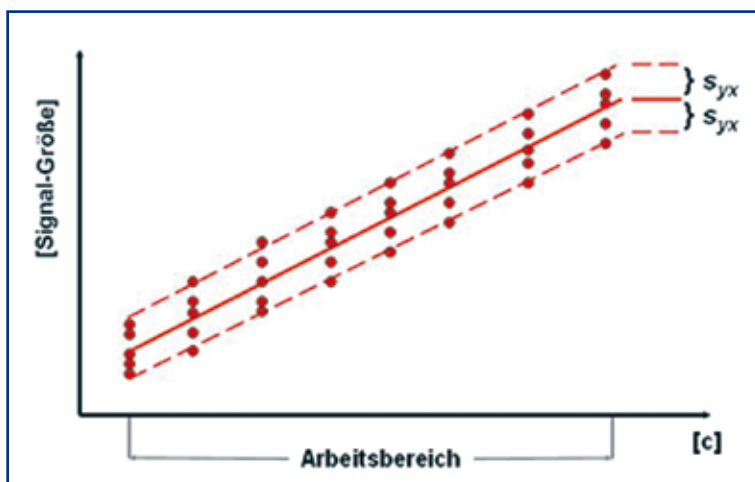
Wenn auf Varianzenhomogenität innerhalb des Arbeitsbereichs eines kalibrierten Verfahrens aus praktischen Gründen verzichtet wird, sind die daraus abzuleitenden Konsequenzen mit Blick auf die Anforderungen an die entsprechende Methode zu berücksichtigen. Diese werden überall dort kritisch, wo die Reststandardabweichung in die Beurteilung beziehungsweise Auswertung direkt oder indirekt eingeht.

Linearität

Als objektives Maß für lineares Verhalten des Informationswertes in Abhängigkeit vom Merkmalswert gilt der Mandel'sche Linearitätstest, wie er in DIN 38 402, Teil 51 (DEV A51) bzw. ISO 8466-1 beschrieben ist. Dabei wird geprüft, ob die Verbesserung der Reststandardabweichung beim Übergang vom linearen zum quadratischen Regressionsmodell signifikant ist oder nur zufällig.

Da Linearität über einen vorgegebenen Arbeitsbereich geprüft wird und dafür die Reststandardabweichung als kritische Größe dient, ist bei vorliegender Varianzeninhomogenität von einer größeren Testunschärfe auszugehen, insbesondere wenn im oberen Konzentrationsbereich eine signifikante Verschlechterung der Wiederholbarkeit vorliegt (Abbildung 5). Die Schärfe des Mandel'schen Linearitätstestes hängt ansonsten hauptsächlich von den Freiheitsgraden und damit von der Anzahl der eingesetzten Standards ab (vergleiche Entwicklung der F-Werte in Abhängigkeit vom Freiheitsgrad).

Abbildung 4: Reststandardabweichung bei Varianzenhomogenität.



Regressionskoeffizienten und Verfahrenskenndaten

In die Berechnung der Regressionskoeffizienten Achsenabschnitt a und Steigung (Empfindlichkeit) b geht die Reststandardabweichung direkt nicht ein. Varianzeninhomogenität ist daher für die Berechnung von a und b unkritisch, aber kritisch bei der Bewertung der gesamten Kalibration, weil dazu die Verfahrenskenndaten Reststandardabweichung s_{yx} und die Verfahrenstandardabweichung s_{x0} eingehen.

Vertrauensband

Über die Regressionskoeffizienten der Kalibration werden die Probeninformationswerte in deren Merkmalswerte umgewandelt. Zur Abschätzung der Unsicherheit aus der Kalibration dient das Vertrauensband der Kalibration (vergleiche Formel (1); Abbildung 3).

Da bei der Berechnung des Vertrauensbereiches von der „Konstanz der Reststandardabweichung s_{yx} “ ausgegangen wird, ist bei tatsächlich vorliegender Varianzeninhomogenität zumindest an einem der beiden Enden des Arbeitsbereichs von einer zu optimistischen Abschätzung der Unsicherheit aus der Kalibration auszugehen. Ob diese zu optimistische Einschätzung im unteren oder oberen Konzentrationsbereich liegt, hängt davon ab, ob die signifikant größere Varianz im unteren oder oberen Arbeitsbereich vorliegt.

Eine zu optimistische Einschätzung im oberen Bereich ist meist weniger kritisch als im unteren Bereich, es sei denn im oberen Konzentrationsbereich sind Grenzwerte zu beurteilen, die nicht überschritten werden dürfen.

Die Konsequenz aus einer zu optimistischen Fehleinschätzung der Unsicherheit im unteren Arbeitsbereich lässt sich am besten am Beispiel der indirekten Methode zur Bestimmung der Nachweis-, Erfassungs- und Bestimmungsgrenzen nach DIN 32 645 beurteilen (vergleiche Abbildung 6).

Wie aus Abbildung 6 leicht zu erkennen ist, verschlechtern sich mit zunehmender Vertrauensbandbreite die Grenzwerte am unteren Arbeitsbereichsende. Liegt Varianzeninhomogenität durch signifikant höhere Varianzwerte im unteren Arbeitsbereich verglichen mit dem oberen vor, würden die erhaltenen Werte für die Nachweis-, Erfassungs- und Bestimmungsgrenze zu niedrig geschätzt werden, weil die in die Vertrauensbandbreite eingehende Reststandardabweichung s_{yx} am unteren Arbeitsbereichsende auf Grund ihres „Mittelwertcharakters“ zu gering eingeschätzt wird. Umgekehrt würde Varianzeninhomogenität durch signifikant niedrigere Varianzen im unteren Arbeitsbereich gegenüber dem oberen zu einer pessimistischen Schätzung der Grenzwerte und damit zu mehr Sicherheit in der Grenzwertbetrachtung führen.

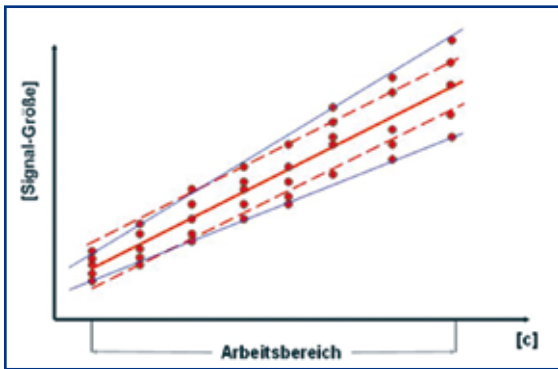


Abbildung 5: Auswirkung einer Varianzeninhomogenität auf die Reststandardabweichung (gestrichelte Linien Reststandardabweichung unter der Annahme einer Varianzenhomogenität).

Folgerungen

Wenn also Varianzeninhomogenität im Arbeitsbereich, das heißt Abhängigkeit der Wiederholstandardabweichung vom jeweiligen Gehalt der analytischen Probe vorliegt, verlieren Verfahrenskenndaten wie die Reststandardabweichung s_{yx} , die Verfahrensstandardabweichung s_{x_0} oder der Vertrauensbereich VB der Kalibration ihre allgemeine Gültigkeit. Neben den Empfehlungen, sich auf geeignete Arbeitsbereiche zu beschränken, bleiben dem Praktiker folgende Alternativen:

a) Er missachtet die Forderung nach Varianzenhomogenität wohl wissend um die oben genannten Risiken eventueller Fehleinschätzungen. Eine Bewertung und Dokumentation dieser Risiken ist insbesondere gegenüber dem Auftraggeber dann nicht nur wünschenswert.

b) Als Ausweg wird auch die „Gewichtete Regression“ diskutiert⁵, deren Umsetzung im Gegensatz zur Kalibration nach den genannten Regelwerken⁶ unabhängige Mehrfachmessungen auf allen Kalibrierlevels zwingend voraussetzt und damit einen beträchtlichen Mehraufwand in der Kalibrationsvorbereitung und bei der Messung der Standards bedeutet. Außerdem kann zur Zeit auf kein Regelwerk verwiesen werden, das die dafür notwendigen Rahmenbedingungen beschreibt.

c) Ein pragmatisches Vorgehen orientiert sich daran, wie kritisch die geforderte Gehaltsbestimmung ist. Insbesondere bei Verfahren mit einer Kalibration über einen erweiterten Arbeitsbereich wird sichergestellt, dass wenigstens im kritischen Bereich Varianzenhomogenität nachgewiesen werden kann, wenn diese schon nicht über den gesamten Arbeitsbereich mit ein und derselben Kalibration erhalten wird. Kritische Bereiche liegen beispielsweise dort vor, wo Grenzwertüberwachungen zuverlässige Analyseergebnisse mit

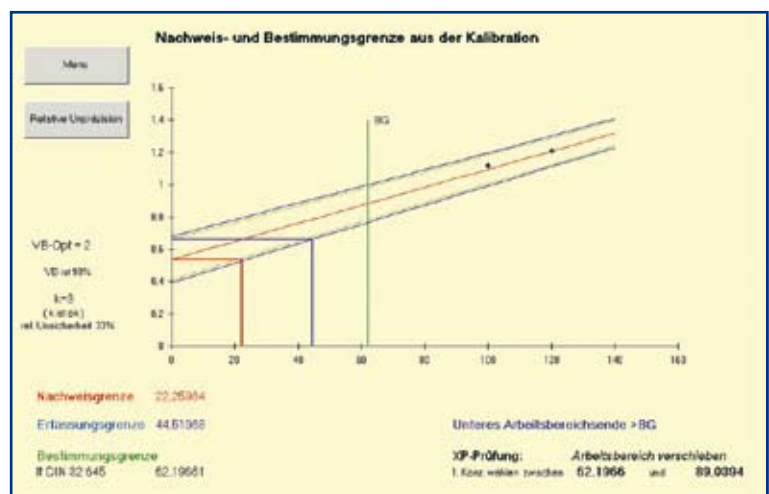
Angaben zu deren Sicherheit unabdingbar machen. Ein kritischer Bereich ist auch die untere Abgrenzung des Arbeitsbereichs (Nachweis-, Erfassungs- und Bestimmungsgrenze). Für Analytbestimmungen außerhalb des Bereichs mit vergleichbarer Wiederholstandardabweichung sind die mit dem Ergebnis verbundenen Risiken (siehe oben) abzuschätzen und zu dokumentieren.

Zusammenfassung

Varianzenhomogenität gewährleistet eine richtige Einschätzung der Ergebnisunsicherheit bezüglich der Kalibrationsumrechnung über den gesamten Arbeitsbereich. Bei Varianzenhomogenität ist das Vertrauensband symmetrisch zur Arbeitsbereichsmittle angeordnet. Varianzeninhomogenität hat keinen Einfluss auf die Regressionskoeffizienten und damit auf die Kalibrationsfunktion, sowie die daraus resultierende Umrechnung des Informationswertes in den zugehörigen Konzentrationswert. Wird die Kalibration nur hinsichtlich dieser Umrechnung benutzt, gibt es für den Praktiker keinen zwingenden Grund, die Forderung nach Varianzenhomogenität zu befolgen. Varianzeninhomogenität wirkt sich „nur“ in der Beurteilung der Ergebnisunsicherheit bezüglich der Kalibrationsumrechnung aus und diesbezüglich besonders an den Rändern des Arbeitsbereichs. Kritisch sind die analytischen Situationen, bei denen die Varianz am unteren Arbeitsbereichsende signifikant größer als am oberen Arbeitsbereichsende ist. Damit werden nämlich analytischen Kenngrößen wie Nachweis-, Erfassungs- und Bestimmungsgrenzen mit der Annahme einer konstanten Reststandardabweichung über den Arbeitsbereich zu klein bestimmt. Unkritisch ist die entgegengesetzte Situation, dabei werden die Kenngrößen als zu groß abgeschätzt verglichen mit der Realität. Bei Probenmessungen im Zentrum des Arbeitsbereichs gleichen sich die Ungenauigkeiten aus und liefern bei Missachtung der Varianzenhomogenität die gleichen Abschätzungen für das Ergebnisunsicherheitsintervall wie beim Einhalten der Varianzenhomogenität.



Abbildung 6: Bestimmung der Nachweis-, Erfassungs- und Bestimmungsgrenze aus dem Vertrauensband der Kalibration nach DIN 32 645 mit Hilfe von SQS.



⁵ Dr. Stefan Schömer, GIT Fachz. Lab. 9/96, 904-907 (1996)

⁶ DIN 38 402, Teil 51 (1986) sowie ISO 8466-1 (1990) und ISO 8466-2 (2001)