M. Krystek

Die Anwendung der Wahrscheinlichkeitstheorie in der Längenmesstechnik

Die Längemesstechnik ist sowohl für eine genaue Fertigung, als auch für die Ingenieursforschung wesentlich, denn sie beeinflusst unmittelbar die Qualität gefertigter Produkte. Um nachzuweisen, ob ein Produkt innerhalb der Spezifikationsgrenzen liegt und anschließend eine Entscheidung über seine Annahme oder Ablehnung treffen zu können, muss die Messunsicherheit in die Entscheidungsfindung mit einbezogen werden. Diese Publikation behandelt das Hintergrundwissen, welches die Grundlage des GUM bildet, aber heute immer noch nicht ausreichend verstanden zu sein scheint.

1 Einleitung

Die industrielle Längenmesstechnik beschäftigt sich unter anderem mit der Messung der geometrischen Merkmale von Werkstücken. Dies schließt die Messung von Lage, Orientierung, Form, Maß (wie z.B. Abstand, Dicke und Durchmesser), Oberflächeneigenschaften (wie z.B. Welligkeit und Rauheit), Defekte (wie z.B. Kantenausbrüche) und das gegenseitige Verhältnis von Geometrieelementen eines Werkstücks (wie Parallelität, Orthogonalität, Winkligkeit, Geradheit, Ebenheit, Rundheit, usw.) ein. Auch die Lage und Orientierung von Werkstücken zueinander ist Gegenstand der industriellen Längenmesstechnik. Sie ist damit sowohl für eine genaue Fertigung, als auch für die Ingenieursforschung wichtig, denn sie beeinflusst unmittelbar das erreichbare Ergebnis.

Der Begriff Längenmessung hat eine operationelle Definition, denn er bedeutet das Ergebnis einer Messung auf die SI-Einheit der Länge, das Meter, zu beziehen, und zwar durch eine ununterbrochene Kette von Vergleichen, die alle eine angegebene Messunsicherheit haben müssen. Daraus folgt unmittelbar, dass die Messunsicherheit eine wesentliche Rolle in der Längenmessung spielt.

Um festzustellen, ob ein Längenmerkmal innerhalb vorgegebener Spezifikationsgrenzen liegt, wie sie z.B. durch einen Konstrukteur auf einer technischen Zeichnung angegeben worden sind, und anschließend eine Entscheidung über die Annahme oder Ablehnung eines Produkts zu treffen, muss die Messunsicherheit in die Aufstellung von Entscheidungskriterien mit einbezogen werden. In der Industrie ist es heute bereits weitgehend akzeptiert, das Messunsicherheitsbudgets von zunehmender Wichtigkeit für die Prozess- und Qualitätskontrolle sind. Der *Guide to the Expression of Uncertainty in Measurement* [GUM 1993], auch kurz GUM genannt, gibt den Rahmen für die Bestimmung von Unsicherheiten vor. Im Folgenden soll das wesentliche Hintergrundwissen diskutiert werden, das die Grundlage des GUM bildet, aber heute immer noch nicht ausreichend verstanden zu sein scheint.

2 Kommensurabilität von Längen

Die geometrische Definition des Begriffs Länge ist der euklidische Abstand von zwei Punkten im physikalischen Raum. Um eine Länge zu messen, muss diese mit einer bekannten Länge verglichen werden. Dieser Vergleich macht es allerdings erforderlich, dass ein Bruchteil oder ein Vielfaches der bekannten Länge genau in der zu messenden Länge aufgeht. In einem solchen Fall sagen wir, die beiden Längen seien kommensurabel, anderenfalls sie seien inkommensurabel. Demnach ist eine grundlegende Forderung der Längenmessung, dass die Längeneinheit mit jeder möglichen zu messenden Länge kommensurabel sein muss. Aber so sinnvoll diese Forderung auch ist, sie ist nicht erfüllbar.

Um diese eigenartige Aussage zu verstehen, stellen wir uns z.B. vor, dass wir einige charakteristische Längen eines geometrisch idealen Würfels zu messen haben. Die Länge der Raumdiagonalen des Würfels verhält sich zur Länge seiner Kanten wie $\sqrt{3} : 1$. Es ist aber wohlbekannt, dass $\sqrt{3}$ eine irrationale Zahl ist. Das bedeutet, das die Längeneinheit bestenfalls entweder mit der Länge der Kanten des Würfels oder mit der Länge seiner Raumdiagonale kommensurabel sein kann, aber nicht mit beiden Längen gleichzeitig. Ein weiteres Beispiel wäre die Messung des Durchmessers eines geometrisch idealen Zylinders und der Länge seines Umfangs, die ja, wie allgemein bekannt ist, sich zueinander wie $1 : \pi$ verhalten, d.h. auch hier tritt ein irrational Verhältnis auf.

Man könnte nun argumentieren, dass es in beiden Fällen ausreichend ist, nur eine der beiden beteiligten Längen zu messen, z.B. im ersten Fall die Kantenlänge des idealen Würfels und im zweiten Fall den Durchmesser des idealen Zylinders, und anschließend die anderen interessierenden Längen durch Anwendung der gut bekannten Regeln der Geometrie zu berechnen. Tatsächlich ist dies aber kein Ausweg, denn welche der Längen auch immer zur Messung ausgewählt wird, sie könnte inkommensurabel mit der Längeneinheit sein; und nicht nur dass wir diese Tat-

sache niemals wissen können, es gibt auch keinerlei Möglichkeit zu prüfen, ob die geforderte Kommensurabilität vorhanden ist oder nicht.
Die Schlussfolgerung ist demnach, dass wir uns mit der Tatsache abzufinden haben, dass wir selbst unter idealen Umständen nicht in der Lage sind, eine Länge genau zu messen, oder besser gesagt, falls es zutreffen sollte, dass wir sie zufällig genau messen, dann werden wir es nicht wissen. Das bedeutet, dass wir eine unvermeidliche immanente Unsicherheit in der Längenmessung haben.
Die Aussage des letzten Absatzes kann als von rein philosophischem Interesse angesehen werden, denn für jeden praktischen Anwendungszweck können wir uns auf die mathematische Tatsache stützen, dass jede irrationale Zahl sich mit jeder beliebigen Genauigkeit durch eine rationale Zahl approximieren lässt, d.h. wenn wir nur die Längeneinheit mit einer gewissen Genauigkeit kennen, dann können wir im Prinzip jede beliebige Länge mit einer vergleichbaren Genauigkeit messen, was bedeutet, dass wir unsere Längenmessung so genau machen können, wie wir nur wünschen. In der Praxis ist also die Grenze der Genauigkeit ausschließlich durch die Genauigkeit gegeben, mit der die Längeneinheit bekannt ist.

3 Information und Wahrscheinlichkeit

Es gibt natürlich zahlreiche andere Einflüsse, wie z.B. die Umgebungstemperatur, die wesentlich ernster zu nehmende Beiträge zur Messunsicherheit in der Längenmesstechnik liefern, als die immanente Unsicherheit, die im vorhergehenden Abschnitt diskutiert worden ist. Allerdings kann diese besondere Unsicherheit als ein gutes Beispiel für eine systematische Abweichung dienen, die im Prinzip unbekannt ist, weil, wie wir bereits gesehen haben, diese Abweichung auftreten kann, aber nicht notwendigerweise auftreten muss.
Wenn wir nach der Wahrscheinlichkeit fragen, dass eine derartige systematische Abweichung von Null verschieden ist, dann können wir uns sicherlich nicht auf die Wahrscheinlichkeitsdefinition stützen, die auf der Vorstellung wiederholter Experimente beruht, wie dies üblicherweise in der orthodoxen Statistik getan wird, denn zum Einen macht es keinen Sinn sich unendlich oft wiederholte Experimente vorzustellen, welche zwar denkbar sind, aber niemals durchgeführt werden, und zum Anderen, selbst wenn sie durchführbar wären, würden wir nur die Häufigkeitsverteilung bezüglich der Wiederholbarkeit unseres Experiments erhalten, die uns keinerlei Informationen über den Wert unserer speziellen systematischen Abweichung liefern würde, vorausgesetzt sie ist überhaupt vorhanden.
Im Gegensatz zu der orthodoxen Statistik beruht die Wahrscheinlichkeitsdefinition in der Bayes-Statistik auf der wie auch immer vorhandenen Kenntnis, welche wir über die betrachtete Situation besitzen, selbst wenn kein zufälliger Prozess im Spiel ist. Dies allein ermöglicht uns die Behandlung systematischer Abweichungen.
Wir können versuchen, diese Aussage ein wenig klarer zu machen, indem wir unser Beispiel einer idealen Längenmessung verwenden. Wir wollen zunächst annehmen, dass die zu messende Länge kein rationales Vielfaches der Längeneinheit ist. Wir entscheiden nun, wie wir diese irrationale Zahl durch eine rationale Zahl approximieren, wobei wir höchstens eine Längenabweichung ε zulassen wollen. Folglich führen wir wissentlich eine systematische Abweichung ein, welche allerdings den absoluten Wert $|\varepsilon|$ nicht überschreitet, d.h. die systematische Abweichung nimmt einen gewissen Wert aus dem reellen Intervall $(-\varepsilon,\ +\varepsilon)$ an, kann aber nicht Null sein. Wenn wir nun die Annahme fallenlassen, dass wir wissen, ob die zu messende Länge ein irrationales Vielfaches der Längeneinheit ist oder nicht, dann nimmt die systematische Abweichung immer noch einen Wert aus dem reellen Intervall $(-\varepsilon,\ +\varepsilon)$ an, aber in diesem Fall ist der Wert Null genauso wahrscheinlich, wie jeder andere Wert innerhalb dieses Intervalls. Auf der Grundlage unserer Kenntnis ordnen wir folglich jedem Wert aus dem reellen Intervall $(-\varepsilon,\ +\varepsilon)$ die gleiche Wahrscheinlichkeit zu, als unsere systematische Abweichung auftreten zu können, d.h. wir legen eine rechteckförmige Wahrscheinlichkeitsdichtefunktion (Gleichverteilung) fest, die mit unserer Kenntnis im Einklang ist, aber ansonsten so uninformativ ist wie nur irgend möglich.
Dieses Beispiel demonstriert, auf welche Weise es die Bayes-Statistik im Prinzip erlaubt, einer Kenntnis über systematische Effekte eine Wahrscheinlichkeitsverteilung zuzuordnen. Tatsächlich war die Möglichkeit, dazu in der Lage zu sein, systematische Messabweichungen in konsistenter Weise zu behandeln, der Grund für die Einführung des GUM, der hauptsächlich auf der Bayes'schen Theorie der Messunsicherheit [Weise & Wöger 1992] beruht. Die früher verwendete Fehlerrechnung, welche auf der orthodoxen Statistik beruht, war dagegen nicht in der Lage, derartige Probleme zu behandeln.
In der Bayes-Statistik sind Wahrscheinlichkeiten stets bedingte Wahrscheinlichkeiten, denn sie beruhen auf den Informationen, welche wir unter den jeweils vorhandenen Umständen in einer bestimmten Situation besitzen. In der Längenmesstechnik, wie auch auf jedem anderen Gebiet der Messtechnik, gibt es im Prinzip zwei verschiedene Arten von Informationen. Die eine ist üblicherweise bereits vorhanden, bevor irgendeine Messung ausgeführt worden ist, während die andere aus den gemessenen Daten abgeleitet wird. Der einzige Grund, warum wir überhaupt Messungen vornehmen ist, weil wir unsere Kenntnis vergrößern wollen.
Die Information, welche wir besitzen, noch bevor irgendeine Messung durchgeführt worden ist, d.h. was wir *a priori* wissen, ist manchmal sehr umfassend und sollte folglich sicherlich nicht ignoriert werden. Dies ist, nebenbei gesagt, was gemeint ist, wenn im GUM gefordert wird *alle* zur Verfügung stehenden Informationen in das Messunsicherheitsbudget mit einzubeziehen. Die Information, nachdem wir einige Messdaten erhalten haben, d.h. was wir *a posteriori* wissen, ist gewöhnlich, im Vergleich zur Information die wir vorher hatten, durch die Messung vergrößert worden, es sei denn, die neue Information ist inkonsistent mit unseren Vorkenntnissen. Wenn der letztere Fall eintritt, müssen wir in Betracht ziehen, dass unser Modell möglicherweise falsch ist oder noch verbessert werden sollte.

In der Bayes-Statistik sind den *a priori* bzw. den *a posteriori* Kenntnissen Wahrscheinlichkeitsdichtefunktionen zugeordnet, die als Prior-Verteilung bzw. Posterior-Verteilung bezeichnet werden. Diese beiden Funktionen sind durch das Bayes'sche Theorem miteinander verbunden, das für eine einzelne Größe durch die Beziehung

$$p(X|\mathbf{x}) = Cp(\mathbf{x}|X)p(X) \tag{1}$$

gegeben ist, wobei X die betrachtete Größe, $\mathbf{x} = (x_1, \dots, x_n)^T$ ein Vektor, der die n gemessenen Daten repräsentiert, $p(X)$ die Prior-Verteilung, $p(X|\mathbf{x})$ die Posterior-Verteilung, $p(\mathbf{x}|X)$ die so genannte Likelihood-Funktion und C eine Normierungskonstante sind. Die Likelihood-Funktion ist ein Maß für die Wahrscheinlichkeit die Daten $\mathbf{x}$ zu erhalten, wenn die Größe X bekannt ist. Das Bayes'sche Theorem zeigt sehr gut, wie die für eine gewisse Größe vorhandene Prior-Wahrscheinlichkeitsdichtefunktion durch die Information, welche in den durch die Messung erhaltenen Daten vorhanden ist, über die Likelihood-Funktion modifiziert wird, um so schließlich die Posterior-Wahrscheinlichkeitsdichtefunktion dieser Größe zu erhalten.

Üblicherweise, insbesondere in der Forschung, hat die Likelihood-Funktion ein sehr scharfes Maximum, wodurch die hohe Relevanz der Daten angezeigt wird, während die Prior-Verteilung sehr oft vergleichsweise flach verläuft, weil unser Vorwissen noch sehr gering ist. Die Posterior-Verteilung wird also durch die Likelihood-Funktion dominiert, d.h. durch die Messdaten. In einigen Fällen allerdings, zum Beispiel für einen Prozess in der Massenproduktion, der gut unter Kontrolle ist und demnach als ausreichend stabil bekannt ist, hat die Prior-Verteilung ein sehr scharfes Maximum und die durch die Prüfung des Werkstücks erhaltenen Daten erhöhen die Information, welche man bereits besitzt, nicht sehr viel. Unter derartigen Umständen ist die Posterior-Verteilung durch die Prior-Verteilung, anstatt durch die Likelihood-Funktion, dominiert und es ist möglich (und vielleicht aus ökonomischen Gründen empfehlenswert), die Unsicherheit eines bestimmten Geometrieelements des Produkts abzuschätzen, ohne überhaupt eine Messung durchzuführen. Dies scheint auf den ersten Blick befremdlich zu sein, ist aber im Sinne der Bayes-Statistik vollkommen in Ordnung und entspricht damit auch dem GUM, der auf ihr beruht. Die Idee hinter dieser Vorgehensweise ist, dass es keine Rolle spielt, auf welche Weise wir unsere Information erhalten haben, es zählt ausschließlich, dass die Information gültig, problemgerecht und ausreichend ist, um in unser Messunsicherheitsbudget mit einbezogen zu werden. Dies erlaubt es uns, Messergebnisse zu verwenden, welche vorher durch Andere oder an anderen Orten erhalten worden sind, welche in Kalibrierscheinen, Tabellen über Materialeigenschaften, wissenschaftlichen Publikationen, Protokollen von Produktionsprozessen, oder ähnlichen Dokumentationen angegeben worden sind. Wir könnten sogar Expertenwissen nutzen, vorausgesetzt, wir sind vorsichtig genug keine bedenkliche Information zu verwenden, was der Regel der Bayes-Statistik widersprechen würde, dass man nicht raten soll, aber alle vorhandenen Tatsachen verwenden soll.

Nachdem nun versucht worden ist, einige der Vorteile der Bayes-Theorie und insbesondere die Verwendung des Bayes'schen Theorems für statistische Schlussfolgerungen und die Datenauswertung aufzuzeigen, müssen wir etwas darüber sagen, wie die Prior-Verteilung und die Likelihood-Funktion erhalten werden können.

Wir haben oben bei dem Beispiel der idealen Längenmessung gesehen, wie Information einer Wahrscheinlichkeitsdichtefunktion zugeordnet werden kann. Natürlich würden die meisten Leute es nicht für falsch halten, jedem Wert aus einem gegebenen Intervall die gleiche Wahrscheinlichkeit zuzuordnen, wenn ein Wert genausogut möglich ist, wie der andere, und keine weitere Information vorhanden ist. Dies folgt aus dem wohlbekannten *Prinzip des unzureichenden Grundes* oder dem so genannten Bayes-Laplace-Postulat [Laplace 1814]. Was wir oben intuitiv getan haben, lässt sich mathematisch durch die Anwendung des *Prinzips der maximalen Entropie* [Jaynes 1957] verallgemeinern. Um die Likelihood-Funktion $p(\mathbf{x}|X)$ zu erhalten, vorausgesetzt die Prior-Verteilung ist auf irgend eine Weise bereits bekannt, verwendet dieses

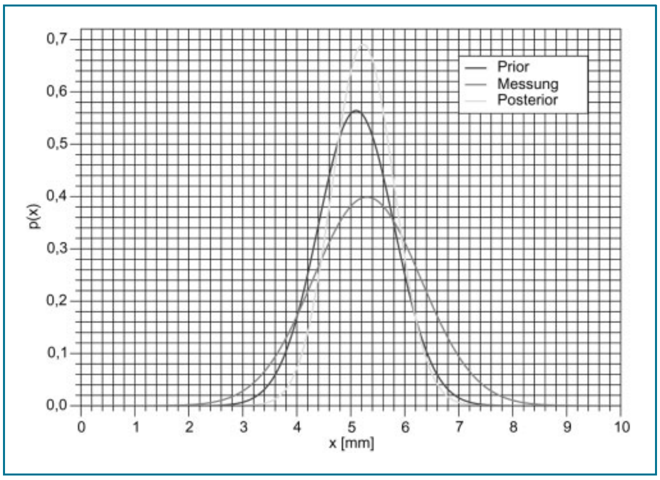

Abb. 1: Zum Bayes'schen Theorem

Prinzip die Variationsrechnung nach Lagrange, um die Informationsentropie [Shannon 1948]

$$S = -\int p(X)p(\mathbf{x}|X)\log p(\mathbf{x}|X)\mathrm{d}X, \tag{2}$$

zu maximieren, unter der üblichen Normalisierungsbedingung, dass das Integral über die gesuchte Wahrscheinlichkeitsdichtefunktion gleich Eins sein muss, wie dies natürlich für jede Wahrscheinlichkeitsdichte gefordert werden muss, und, so weit anwendbar, weiteren Nebenbedingungen, welche die Information repräsentieren, die für die spezielle Messaufgabe vorhanden ist.

Ein wichtiger Punkt, der hier erwähnt werden muss, ist das Modell des physikalischen Messverfahrens, das immer als einer der wichtigsten Anteile unserer vorhandenen Kenntnisse mit einbezogen werden muss. Das Modell kann entweder durch eine oder mehrere mathematische Gleichungen und Relationen gegeben sein, oder genauso gut durch eine algorithmische Spezifikation oder ein geeignetes Computerprogramm. Die letztere Möglichkeit wird üblicherweise gewählt, wenn die Monte-Carlo-Methode verwendet wird, um die Messunsicherheit zu bestimmen. Diese Methode wird seit Kurzem durch das Erscheinen eines neuen Supplements zum GUM [ISO 2006] besser unterstützt.

Das Prinzip der maximalen Entropie wurde in der Vergangenheit bereits erfolgreich auf viele Probleme angewendet, es bleibt aber immer noch die Schwierigkeit, die notwendige Prior-Verteilung festzulegen. In den meisten Fällen wird lediglich eine konstante Prior-Verteilung gewählt [Jeffreys 1939], aber andere Möglichkeiten sind ebenfalls in der Literatur diskutiert worden [Jaynes 1968].

4 Erwartungswert und Unsicherheit

Nach der Bayes'schen Theorie der Messunsicherheit ist die Posterior-Verteilung die mathematische Repräsentation des Zustands der Kenntnis bezüglich des Wertes der interessierenden Messgrößen, nachdem die Messung durchgeführt worden ist. Alle diese Größen betreffenden interessierenden Werte können unmittelbar aus ihrer Wahrscheinlichkeitsdichtefunktion erhalten werden.

Wenn wir die Messgröße mit X, ihre Posterior-Verteilung mit $p(X|\mathbf{x})$ und ihre Daten durch x bezeichnen, lassen sich zwei Werte spezifizieren, die von unmittelbarem Interesse sind. Der erste ist der so genannte Erwartungswert von X

$$\mathrm{E}X = \int Xp(X|\mathbf{x})\mathrm{d}X, \tag{3}$$

und der zweite ist die Varianz der Wahrscheinlichkeitsdichtefunktion, gegeben durch

$$\mathrm{Var}X = \mathrm{E}(X - \mathrm{E}X)^2. \tag{4}$$

Dieser Wert ist ein Maß für die Dispersion der Posterior-Verteilung um den Erwartungswert der betrachteten Größe.

Die Posterior-Verteilung enthält auch Informationen über denjenigen Wert, welcher als der beste Schätzwert für den wahren Wert der Messgröße angesehen werden kann. Nach den heutigen internationalen Konventionen, ist als bester anzugebender Schätzwert der Wert anzusehen, um den die Streuung der Wahrscheinlichkeitsdichtefunktion am kleinsten ist, d.h. um den die Wahrscheinlichkeitsdichtefunktion am stärksten konzentriert ist [Gauss 1809]. Wenn wir einen beliebigen Wert, sagen wir ξ, auswählen, dann ist die Dispersion der Wahrscheinlichkeitsdichtefunktion um diesen Wert gegeben durch

$$\mathrm{E}(X - \xi)^2 = \mathrm{Var}X + (\mathrm{E}X - \xi)^2. \tag{5}$$

Offenbar hat die Dispersion ein Minimum für $\xi = \mathrm{E}X$. Demnach ist die Dispersion bei diesem speziellen Wert gleich der Varianz der Wahrscheinlichkeitsdichtefunktion. Die einem beliebigen Wert ξ zugeordnete Messunsicherheit ist als Quadratwurzel der Dispersion definiert, d.h. durch

$$u(\xi) = \sqrt{\mathrm{E}(X - \xi)^2}. \tag{6}$$

Deshalb ist die kleinstmögliche Messunsicherheit $u(x)$ dem besten Schätzwert $x = \mathrm{E}X$ zugeordnet, wie dies auch im GUM angegeben ist. Die kleinstmögliche Messunsicherheit wird als Standardmessunsicherheit bezeichnet.

In der Längenmesstechnik wird, hauptsächlich verursacht durch Umgebungseinflüsse, wie z.B. Temperaturänderungen oder andere so genannte Einflussgrößen, häufig die Größe

$$Y = X + X_{\mathrm{syst}} \tag{7}$$

anstelle der interessierenden Messgröße X gemessen, wobei x_{syst} als systematische Messabweichung bezeichnet wird. Um also den besten Schätzwert der interessierenden Größe X zu erhalten, müssen wir unseren Messwert um die systematische Abweichung korrigieren, vorausgesetzt, dass dieser Wert bekannt ist. Dies wird auch im GUM empfohlen. Die dem Erwartungswert EX zugeordnete Standardmessunsicherheit ist

$$u(x) = \sqrt{u^2(y) + u^2(x_{\mathrm{syst}})}. \tag{8}$$

Diese Gleichung ist allerdings nur dann gültig, wenn die empfohlene Korrektur bezüglich einer bekannten systematischen Abweichung durchgeführt worden ist. Wenn dies absichtlich, aus welchem Grund auch immer, unter-

blieben ist, muss die Unsicherheit erhöht werden, indem man den quadrierten Wert der systematischen Abweichung zu den Termen unter der Quadratwurzel in der Gleichung (8) addiert [Lira & Wöger 1998]. Es ist zwingend notwendig sich an diese Vorschrift zu halten, um die Konformität der Unsicherheitsfortpflanzung sicherzustellen, denn nur diese garantiert die Rückführbarkeit auf die SI-Einheit der Länge.
Es soll hier noch angemerkt werden, dass, vorausgesetzt eine bekannte systematische Messabweichung ist tatsächlich vorhanden und diese hat zufälligerweise den Erwartungswert Null, wir immer noch die zugeordnete Unsicherheit berücksichtigen müssen, denn diese Unsicherheit verschwindet nicht, außer wenn die systematische Messabweichung vollständig bekannt wäre, eine Annahme, die niemals gerechtfertigt werden kann.

5 Korrelationen

Wenn zwei Größen X und Y (vollständig oder teilweise) korreliert sind, dann können wir ihre Kovarianz

$$\mathrm{Cov}XY = \mathrm{E}(X-x)(Y-y), \tag{9}$$

einführen, welche eine Verallgemeinerung der durch die Gleichung (3) gegebenen Varianz darstellt. Kovarianzen sind ein Maß für Paarkorrelationen der entsprechenden Größen X und Y. Üblicherweise werden Varianzen und Kovarianzen in einer Matrix zusammengefasst, welche als Unsicherheitsmatrix oder Kovarianzmatrix bezeichnet wird.
In den folgenden Fällen sind die Werte einer Messgröße voneinander abhängig:

- Die Werte werden mit dem gleichen Messgerät gemessen.
- Die Werte werden mit unterschiedlichen Messgeräten gemessen, aber es wird das gleiche Normal zur Kalibrierung verwendet.
- Eine Reihe von Messwerten wird gefiltert.
- Die Werte werden unter Verwendung derselben Information berechnet.

Wir wollen als Beispiel für die Auswirkung einer Korrelation von Messwerten den Fall einer systematischen Messabweichung x_{syst} betrachten. Es seien n Werte mit dem gleichen Messinstrument gemessen, das eine systematische Abweichung besitzen soll, die bekannt ist. Das Modell lautet in diesem Fall

$$Y_i = X_i + X_{\mathrm{syst}}, \quad i = 1, \ldots, n. \tag{10}$$

Für unabhängige, normalverteilte Größen $X_i (i = 1, \ldots, n)$ erhalten wir dann für die den Größen Y_i zugeordneten Standardunsicherheiten

$$u(y_i) = \sqrt{u^2(x_i) + u^2(x_{\mathrm{syst}})}, \quad i = 1, \ldots, n, \tag{11}$$

und für die Kovarianzen der Größen Y_i

$$\mathrm{Cov}Y_iY_j = u^2(x_{\mathrm{syst}}), \quad i \neq j, \quad i,j = 1, \ldots, n. \tag{12}$$

Die systematische Messabweichung führt also auch dann zu einer Korrelation der Größen Y_i $(i = 1, \ldots, n)$, wenn sie korrigiert worden ist, bzw. wenn $\mathrm{E}X_{\mathrm{syst}} = 0$ ist!
Für die Unsicherheit des Mittelwerts

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{13}$$

erhalten wir

$$u(\bar{y}) = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} u^2(x_i) + u^2(x_{\mathrm{syst}})}. \tag{14}$$

Man sieht dann unmittelbar, dass

$$\lim_{n\to\infty} u(\bar{y}) = u(x_{\mathrm{syst}}) \tag{15}$$

gilt, d.h. für eine große Anzahl von Messungen geht die Standardunsicherheit des Mittelwertes asymptotisch gegen die Unsicherheit der systematischen Abweichung. Die Unsicherheit des Mittelwerts wird mit zunehmender Anzahl von Messwerten immer stärker durch die Unsicherheit der systematischen Abweichung dominiert!

6 Konfidenzbereiche

Nachdem wir den Erwartungswert und die ihm zugeordnete Standardunsicherheit berechnet haben, sind wir daran interessiert auch zu wissen, wieviel Vertrauen wir in unsere Ergebnisse haben können. Tatsächlich möchten wir gerne wissen, wie wahrscheinlich es ist, den Wert der Messgröße innerhalb eines gewissen Intervalls um den Erwartungswert herum zu finden.
Es ist natürlich aus der Definition der Wahrscheinlichkeitsdichtefunktion klar, dass der Wert der Messgröße mit Sicherheit innerhalb des Geltungsbereichs der Posterior-Verteilung liegt, denn das Integral über dieses spezielle Intervall erstreckt ergibt stets Eins. Wenn wir das Integrationsintervall symmetrisch um den Erwartungswert der jeweiligen Größe herum einschränken, werden wir natürlich eine geringere Wahrscheinlichkeit dafür erhalten, den Wert der Messgröße innerhalb dieses kleineren Intervalls zu finden, d.h.

$$\alpha = \int_{\mathrm{E}X-U}^{\mathrm{E}X+U} p(X|\mathbf{x})\mathrm{d}X \tag{16}$$

lässt sich als die Wahrscheinlichkeit dafür ansehen, den Wert der Messgröße innerhalb des Intervalls $(\mathrm{E}X - U,\ \mathrm{E}X + U)$ zu finden, wobei $p(X|\mathbf{x})$ die Posterior-Verteilung der Größe X ist, wenn die Daten x gegeben sind, und U legt die symmetrischen Grenzen des betrachteten Intervalls um den Erwartungswert fest. Wenn wir nun eine gewisse Wahrscheinlichkeit α festlegen, welche Überdeckungswahrscheinlichkeit (Grad des Vertrauens) genannt wird, dann werden wir dazu ein gewisses Intervall erhalten, das dieser Wahrscheinlichkeit eindeutig zugeordnet ist und Konfidenzbereich genannt wird, indem wir den Wert U nach der Gleichung (16) berechnen. Aus diesen Erklärungen wird somit klar, das der Konfidenzbereich ein Intervall ist, welches mit einer festgelegten Wahrscheinlichkeit den Wert der Messgröße enthalten kann, aber ihn nicht zwangsläufig enthalten muss. Diese Definition ist offensichtlich unterschiedlich von der durch die orthodoxe Statistik gegebenen Definition, denn sie be-

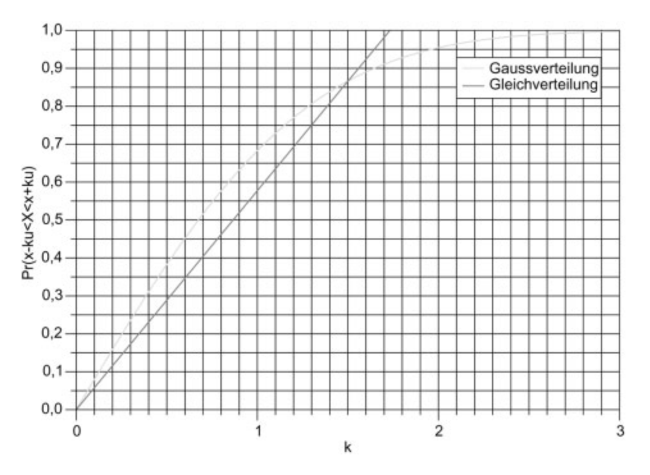

Abb. 2: Die Abhängigkeit des Erweiterungsfaktors k *von der Überdeckungswahrscheinlichkeit für zwei unterschiedliche Wahrscheinlichkeitsverteilungen*

ruht nicht auf einer Häufigkeitsverteilung der Messgröße, sondern statt dessen auf der Posterior-Verteilung, also neben der aus den Messdaten erhaltenen Information auch auf dem gesamten Vorwissen.
Unglücklicherweise ist der Wert U durch Konvention über die Gleichung $U = ku$ zur Standardunsicherheit u in Relation gesetzt worden und wird im GUM erweiterte Unsicherheit genannt. Diese Bezeichnung ist allerdings irreführend, denn aus den vorhergehenden Ausführungen ist wohl hinreichend klar geworden, das U selbst keinesfalls die Bedeutung einer Unsicherheit hat. Das Multiplizieren der Standardunsicherheit mit einer beliebigen Konstante liefert nämlich keinerlei neue Information, sondern nur dieselbe Information in einer anderen Form, d.h. k ist nicht frei wählbar, sondern muss berechnet werden, um zu einer sinnvollen Interpretation von U zu gelangen.
Der oben verwendete Faktor k wird im GUM Erweiterungsfaktor genannt. In nationalen und internationalen Normen und Richtlinien wird häufig $k = 2$ verwendet, wobei eine Wahrscheinlichkeit von 95 % angenommen wird. Dies ist allerdings nur dann richtig, wenn die Posterior-Verteilung gerade eine Normalverteilung (Gauß-Verteilung) ist. Für eine Rechteckverteilung z.B. muss man $k = 1{,}65$ verwenden, wenn eine Überdeckungswahrscheinlichkeit von 95 % gefordert wird (siehe Abb. 2). Ganz allgemein muss k als Verhältnis $k = U/u$ berechnet werden, wobei U durch die Gleichung (16) festgelegt ist. Demnach hängt k, genauso wie U, nicht nur von der Überdeckungswahrscheinlichkeit ab, sondern auch von der Posterior-Verteilung, welche im allgemeinen keine Normalverteilung ist und manchmal eine Form haben kann, die davon sehr verschieden ist.
In der Praxis kann die erweiterte Unsicherheit U im allgemeinen nur numerisch berechnet werden. Anstatt die üblichen numerischen Integrationsalgorithmen anzuwenden, ist es aus Gründen der Effizienz manchmal wesentlich vorteilhafter die Monte-Carlo-Methode einzusetzen [Cox & Siebert 2006]. Dies ist insbesondere dann sinnvoll, wenn multivariate Wahrscheinlichkeitsdichtefunktionen vorkommen.

7 Verifikation und Entscheidungsregeln

In der industriellen Längenmesstechnik liegt der Sinn der Inspektion eines Werkstücks als Teil des Qualitätsmanagementsystems in der Überprüfung, ob die durch den Konstrukteur beabsichtigten Spezifikationen eingehalten worden sind. Üblicherweise wird die Spezifikation eines Merkmals durch einen Zielwert (Nennwert) und ein Toleranzintervall um diesen Wert herum angegeben.
Die Inspektion erlaubt letztlich die Entscheidung über die Annahme oder Zurückweisung des Werkstücks. Heute stimmen alle Experten darin überein, dass Entscheidungsregeln nicht allein auf den Messwerten des Merkmals beruhen können, sondern dass auch der Erwartungswert und die ihm zugeordnete Messunsicherheit mit einbezogen werden müssen.
Um die Gedanken zu verstehen, welche hinter den Entscheidungsregeln stehen, wollen wir das folgende Szenario annehmen. Ein Werkstück wurde bezüglich eines Merkmals X geprüft. Von dem Prüflabor haben wir lediglich den Erwartungswert EX und die ihm zugeordnete Standardunsicherheit $u(x)$ erhalten. Wir kennen auch die untere und obere Spezifikationsgrenze X_U bzw. X_L. Die Frage ist nun: Was ist auf der Grundlage dieser Informationen die Wahrscheinlichkeit dafür, dass das Merkmal X innerhalb der Spezifikationsgrenzen liegt? Offenbar benötigen wir die Wahrscheinlichkeitsdichtefunktion um diese Frage zu beantworten. Wir verwenden zu diesem Zweck das Prinzip der maximalen Entropie, aus dem sich in unserem speziellen Fall eine Normalverteilung für das gemessene Merkmal ergibt. Nach einigen länglichen Berechnungen erhalten wir die Wahrscheinlichkeit (siehe dazu auch die Abb. 3)

$$\Pr(X_L \le X \le X_U) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{X_U - \mathrm{E}X}{u(x)\sqrt{2}}\right) + \operatorname{erf}\left(\frac{\mathrm{E}X - X_L}{u(x)\sqrt{2}}\right)\right], \tag{17}$$

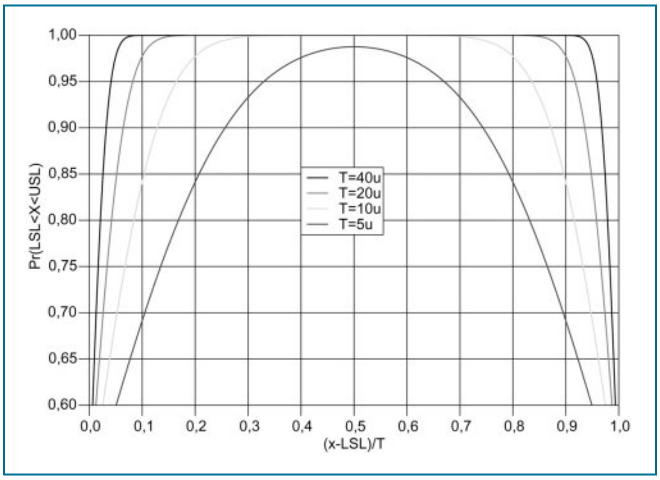

Abb. 3: Die Wahrscheinlichkeit dafür, dass ein Wert x *innerhalb der Toleranzgrenzen* LSL *(untere Spezifikationsgrenze) und* USL *(obere Spezifikationsgrenze) liegt, wenn die zugeordnete Standardunsicherheit* u *und die Breite der Toleranzzone* $T = USL - LSL$ *gegeben ist*

X innerhalb der Spezifikationsgrenzen zu finden, wobei die Abkürzung $\mathrm{erf}(x)$ die Fehlerfunktion [Erdelyi 1953] bezeichnet. Unsere Entscheidungsregel könnte nun sein, das Werkstück nur dann anzunehmen, wenn diese Wahrscheinlichkeit einen bestimmten Wert, sagen wir 95 %, überschreitet, die wir als kritische Wahrscheinlichkeit bezeichnen wollen, und anderenfalls das Werkstück zurückzuweisen.

Wir wollen noch die Konsequenzen aus der Gleichung (17) im Lichte einer internationalen Norm [ISO 1998] untersuchen, welche häufig in der industriellen Messtechnik herangezogen wird. Diese Norm legt fest, dass wir unser Werkstück dann anzunehmen haben, wenn $X_L + U < \mathrm{E}X < X_U - U$ gültig ist. Wenn wir den üblichen Wert $k = 2$ zugrunde legen, dann erhalten wir eine kritische Wahrscheinlichkeit, die etwas größer als 95 % ist, vorausgesetzt, dass die erweiterte Messunsicherheit kleiner als die Hälfte der Breite des Toleranzintervalls ist, was üblicherweise garantiert werden kann. Demnach legt die Norm eine gültige Entscheidungsregel für die Annahme fest.

Es sei noch bemerkt, dass man selbstverständlich auch andere Entscheidungsregeln festlegen kann, als die in unserem Beispiel gewählte. Die Entscheidungsregeln richten sich nach dem beabsichtigten Zweck der Entscheidung. Man sollte aber bei jeder Art von Entscheidungsregel immer von der Posterior-Verteilung des betreffenden Merkmals ausgehen, denn nur diese enthält die gesamte bekannte Information.

8 Zusammenfassung

Am Beispiel der Längenmesstechnik ist versucht worden, einen kurzen Überblick über die grundlegenden Prinzipien zu geben, die hinter dem GUM stehen, der hauptsächlich auf den Grundsätzen der Bayes-Statistik beruht, obwohl dies in diesem Dokument selbst nicht sehr klar zum Ausdruck kommt. Es ist der Zusammenhang zwischen der vorhandenen Information und der Wahrscheinlichkeitsdichtefunktion aufgezeigt worden und wie man diese durch die Anwendung des Prinzips der maximalen Entropie erhalten kann. Es sollte klar geworden sein, warum die aus dem Bayes'schen Theorem erhaltene Posterior-Verteilung alle Kenntnisse enthält, die wir über unsere Messung besitzen, und wie wichtige Werte, wie Erwartungswert, Unsicherheit und auch Konfidenzbereiche aus dieser Wahrscheinlichkeitsdichtefunktion erhalten werden können. Schließlich sind Prüf- und Entscheidungsregeln kurz diskutiert worden, und wie diese zu den Toleranzintervallen und den Unsicherheiten in Beziehung stehen.

Viele wichtige Probleme sind allerdings offengelassen worden, wie z.B. eine eingehende Diskussion der systematischen Messabweichungen, die durch Umgebungsbedingungen oder durch Unzulänglichkeiten des verwendeten Messinstruments hervorgerufen werden, aber es erscheint notwendig sicherzustellen, dass zunächst die grundlegenden Dinge vollständig verstanden sind, bevor über kompliziertere Probleme nachgedacht werden kann.

Literatur

[1] [Bayes 1763] Bayes T.: „An Essay Towards Solving a Problem in the Doctrine of Chances"; Phil. Trans. Roy. Soc. **53**; pp. 376–398, 1763

[2] [Cox & Siebert 2006] Cox, M. G.; Siebert, B. R. L.: „The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty"; Metrologia **43**; pp. 178–188, 2006

[3] [Erdelyi 1953] Erdelyi, A.: „Higher Transcendental Functions, Vol. 2"; MacGraw-Hill; New York 1953

[4] [Gauss 1809] Gauss, C. F.: „Theoria motus corporum coelestium in sectionibus conicis solem ambientium"; Perthes; Hamburg 1809

[5] [GUM 1993] ISO: „Guide to the Expression of Uncertainty in Measurement"; ISO; Geneva 1993

[6] [ISO 1998] ISO 14253-1: „Geometrical Product Specifications (GPS) – Inspection by measurement of workpieces and measuring equipment – Part 1: Deci-

sion rules for proving conformance or non-conformance with specifications“; ISO; Geneva 1998

[7] [ISO 2006] ISO Guide: „Guide to the Expression of Uncertainty in Measurement (GUM) – Supplement 1: Propagation of distributions using a Monte Carlo method“; ISO; Geneva 2006

[8] [Jaynes 1957] Jaynes, E. T.: „Information Theory and Statistical Mechanics“; Phys. Rev. **106**; pp. 620–630, 1957

[9] [Jaynes 1968] Jaynes, E. T.: „Prior Probabilities“; IEEE Trans. Syst. Sci. Cybern. **SSC-4**; pp. 227–241, 1968

[10] [Jeffreys 1939] Jeffreys, H.: „Theory of Probability“; Clarendon Press; Oxford 1939

[11] [Laplace 1814] Laplace, P. S.: „Essai Philosophique sur les Probilités“; Courcier; Paris 1814

[12] [Lira & Wöger 1998] Lira, I.; Wöger, W.: „Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects“; Meas. Sci. Technol. **9**; pp. 1010–1011, 1998

[13] [Shannon 1948] Shannon, C. E.: „A Mathematical Theory of Communication“; Bell Sytem Techn. Journ. **27**; pp. 379–623, 1948

[14] [Weise & Wöger 1992] Weise, K.; Wöger, W.: „A Bayesian theory of measurement uncertainty“; Meas. Sci. Technol. **4**; pp. 1–11, 1992

Anschrift der Verfasser:
Dr. rer. nat. Michael Krystek, Physikalisch-Technische Bundesanstalt (PTB), 38116 Braunschweig, Bundesallee 100, E-Mail: michael.krystek@ptb.de