



# Erweiterte Messunsicherheit

## Lernziel:

Angabe eines Erweiterungsfaktors und ausrechnen des entsprechenden Vertrauensgrads für das Messergebnis.

## Inhalt

1. Vertrauensgrad und Erweiterungsfaktor .....	3
1.1 Einführung: Über die Interpretation der Standardunsicherheit.....	3
1.2 Vertrauensgrad eines Intervalls.....	3
1.3 Standardabweichung und Vertrauensgrad für eine Normalverteilung .....	3
1.4 Erweiterte Unsicherheit und Erweiterungsfaktor .....	4
1.5 Benützung der erweiterten Unsicherheit .....	5
2. Form der Verteilung.....	5
2.1 Wichtigkeit der Bestimmung der Verteilung.....	5
2.2 Methoden zur Bestimmung der Verteilung .....	6
2.3 Kurze Präsentation des zentralen Grenzwertsatzes .....	6
2.4 Kurze Präsentation der $t$ -Verteilung.....	6
3. Praktische Berechnung der erweiterten Unsicherheit .....	7
3.1 Fall einer Normalverteilung .....	7
3.2 Fall einer $t$ -Verteilung.....	7
3.3 Graphische Zusammenfassung.....	8
Anhang.....	9
A Begriffe aus der Statistik und Verteilungen .....	9
A.1 Unsicherheit der Standardunsicherheit .....	9
A.2 Rückblick: Erwartungswert und Varianz.....	9
A.3 Standardabweichung des Mittelwertes .....	10
A.4 Standardabweichung der Standardabweichung .....	10
A.5 Verteilungsfunktion und Wahrscheinlichkeitsdichte .....	11
A.6 Vertrauensgrad eines Intervalls.....	15
B Bestimmung der erweiterten Unsicherheit .....	15
B.1 Einführung .....	15
B.2 Grundbegriffe .....	16
B.2.1 Zentraler Grenzwertsatz.....	16
B.2.2 Die $t$ -Verteilung.....	19

<i>Eigenschaften der <math>t</math>-Verteilung</i> .....	19
B.3    Praktische Ermittlung der erweiterten Unsicherheit .....	21
B.3.1    Effektive Freiheitsgrade .....	21
B.3.2    Praktische Anwendung der $t$ -Verteilung .....	22
B.3.3    Bilaterale und unilaterale Vertrauensintervalle.....	24
B.3.4 $t$ -Verteilung und Normalverteilung in EXCEL.....	25

# 1. Vertrauensgrad und Erweiterungsfaktor

## 1.1 Einführung: Über die Interpretation der Standardunsicherheit

Es wurde bis jetzt gesehen, dass die kombinierte Standardunsicherheit  $u_c(y)$  sich universell zur Angabe eines Messergebnisses für die Messgrösse  $y$  verwenden lässt. Aus dem Ausdruck  $y \pm u_c(y)$  erfährt man, dass die beste Schätzung für  $Y = f(X_1, \dots, X_N)$  den Wert  $y$  annimmt und die Streuung der Werte, die vernünftigerweise der Messgrösse zugeordnet werden können, durch  $u_c(y)$  charakterisiert ist. Aber wie soll man  $u_c(y)$  nun interpretieren? Wie gross ist das Risiko, dass bei Wiederholung der Messung der Wert für  $Y$  ausserhalb des durch  $y \pm u_c(y)$  angegebenen Intervalls liegt?

## 1.2 Vertrauensgrad eines Intervalls

Es wurde im Rahmen von diesem Kurs schon gesehen, dass jeder Messgrösse (Eingangsgrösse  $X_i$  oder Ausgangsgrösse  $Y$ ) eine Verteilung möglicher Werte zugeordnet wird. Diese Verteilung wird üblicherweise durch ihre so genannte Wahrscheinlichkeitsdichtefunktion  $g_i(x_i)$  oder  $g(y)$  angegeben.

Eine richtig normierte Wahrscheinlichkeitsdichte hat eine Gesamtfläche gleich 1 unterhalb ihrer Kurve<sup>1</sup>. Diese Fläche entspricht der Gesamtheit der Werte, die der entsprechenden Messgrösse vernünftigerweise zugeordnet werden können.

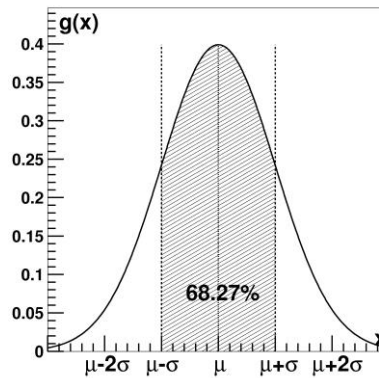
Der Bruchteil  $p$  dieser Fläche unter der Kurve, der innerhalb eines Intervalls  $[a, b]$  liegt, ist gleich dem Anteil der Werte der Messgrösse, der in diesem Intervall liegt. Dieser Bruchteil  $p$  (in % ausgedrückt) wird Vertrauensgrad oder Überdeckungswahrscheinlichkeit des Intervalls genannt.

## 1.3 Standardabweichung und Vertrauensgrad für eine Normalverteilung

Einer Standardabweichung  $\sigma$  einer Normalverteilung entspricht ein Intervall  $[\mu - \sigma, \mu + \sigma]$  um den Mittelwert  $\mu$ , das etwa 68.3% der Gesamtfläche unterhalb der Kurve im Intervall  $[-\infty, +\infty]$  umfasst. Dieser Prozentsatz  $p$  ist der Vertrauensgrad dieses Intervalls. Das bedeutet, dass wenn man einen Prozess, der Werte zufällig nach dieser Verteilung produziert, 100 mal anwendet, werden etwa 68 Werte im Intervall  $[\mu - \sigma, \mu + \sigma]$  und die übrigen 32 Werte werden ausserhalb desselben liegen.

---

<sup>1</sup> Die Wahrscheinlichkeitsinterpretation von der 1 ist die Gewissheit, d.h. der „wahre“ Wert von der Grösse, die durch diese Verteilung beschrieben wird, befindet sich notwendigerweise unter den Werten auf der Abszisse unter dem Graph.



**Figur 1:** Für eine Normalverteilung hat ein Intervall der Breite  $2\sigma$  um den Mittelwert  $\mu$  einen Vertrauensgrad von 68.3%. Ordinatenwerte sind für  $\sigma = 1$  gezeigt.

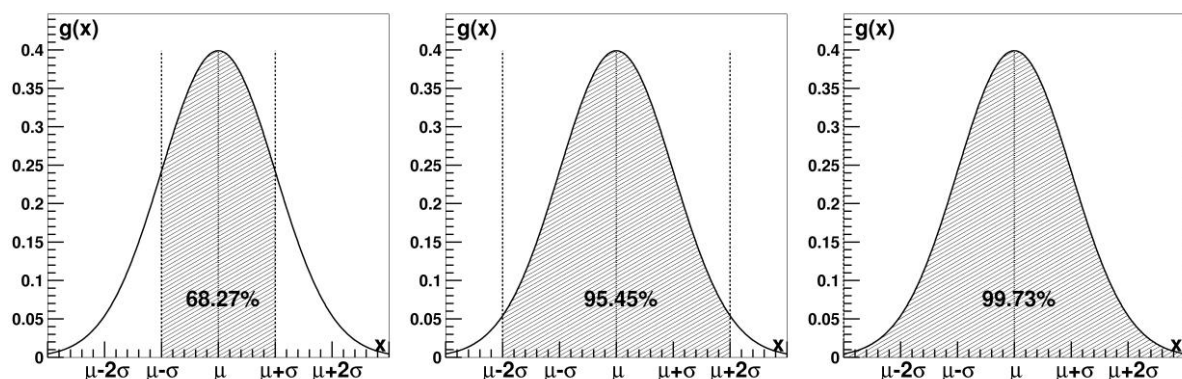
Folglich hat für eine normalverteilte Messgrösse das Intervall, das durch ihre Standardunsicherheit gegeben ist, einen Vertrauensgrad von 68.3%, was die Frage von Abschnitt 1.1 beantwortet.

## 1.4 Erweiterte Unsicherheit und Erweiterungsfaktor

Für ein Messresultat der Messgrösse  $Y$  ist es oft notwendig, die Unsicherheit als ein Intervall mit einem höheren Vertrauensgrad als demjenigen, der durch die Standardunsicherheit definiert wird, anzugeben. Sie wird dann erweiterte Unsicherheit genannt und mit  $U$  (gross  $u$ ) bezeichnet. Man erhält sie durch Multiplikation der kombinierten Standardunsicherheit  $u_c(y)$  mit einem Erweiterungsfaktor  $k$ , sodass  $U = k \cdot u_c(y)$ .

Das Ergebnis der Messung kann dann durch  $Y = y \pm U$  ausgedrückt werden, was so zu interpretieren ist, dass  $y$  der beste Schätzwert des der Messgrösse  $Y$  zugehörigen Wertes ist und dass  $y-U$  bis  $y+U$  einen Bereich darstellt, von dem erwartet werden kann, dass er einen grossen Anteil der Verteilung der Werte umfasst, die  $Y$  sinnvollerweise zugeordnet werden können.

Wenn man eine gegebene Verteilung betrachtet, dann entspricht das Intervall, das eine Halbbreite  $k \cdot \sigma$  hat, einem festen Grad des Vertrauens  $p$ . Für eine Normalverteilung sind diese Zahlen  $p = 68.3\%$ , beziehungsweise  $95.5\%$  und  $99.7\%$  für  $k = 1, 2$  und  $3$ .



**Figur 2:** Überdeckungswahrscheinlichkeiten für  $\pm\sigma$ ,  $\pm 2\sigma$  und  $\pm 3\sigma$  Intervalle um den Mittelwert  $\mu$  für eine Normalverteilung. Ordinatenwerte sind für  $\sigma = 1$  gezeigt.

Im Allgemeinen wird  $k$  zwischen 2 und 3 gewählt. Bei speziellen Anwendungen kann  $k$  auch höher sein. Die Wahl von  $k$  ist zum Teil Konvention (anders je nach Anwendungsbereich) aber auch Kompromissache:  $k$  sollte weder zu klein sein sodass, das Vertrauensgrad genügend gross ist, noch zu gross sodass die Unsicherheit  $U$  aussagekräftig bleibt. Zum Beispiel

wäre es nicht sinnvoll, eine Länge als  $2\text{ m} \pm 1\text{ m}$  mit einem Vertrauensgrad von 99.999999% anzugeben, da so eine grosse Unsicherheit (50%!) das Resultat praktisch unbrauchbar macht.

## 1.5 Benützung der erweiterten Unsicherheit

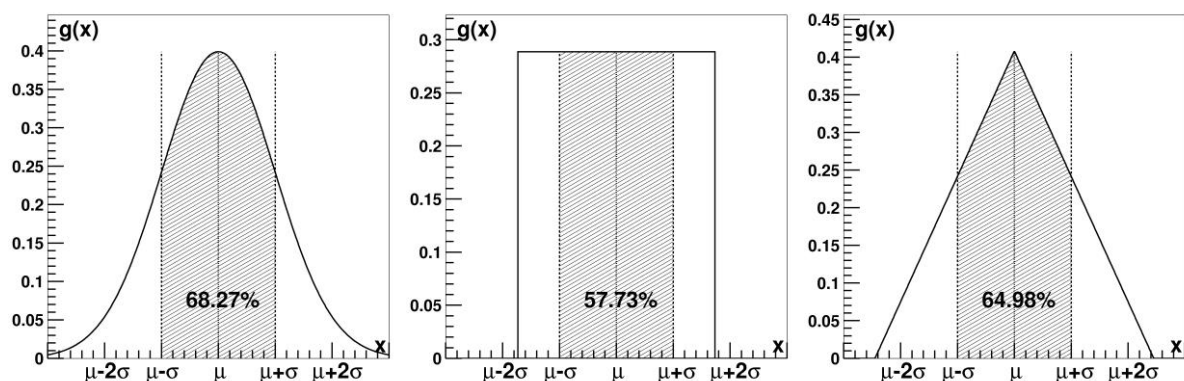
Die erweiterte Unsicherheit  $U$  dient nur dazu, das Schlussresultat von einer Messung anzugeben. Es ist wichtig, immer die Standardunsicherheit  $u$  zu verwenden, um Unsicherheitsfortpflanzungsberechnungen mit der quadratischen Additionsformel durchzuführen, sonst ist es **falsch**! Deswegen ist es bei der Angabe des Messresultats notwendig, immer den verwendeten Erweiterungsfaktor  $k$  anzugeben, sodass die Standardunsicherheit aus der erweiterten Unsicherheit jederzeit wiedererhalten werden kann.

## 2. Form der Verteilung

Im Idealfall wird man gern einen bestimmten Wert des Erweiterungsfaktors  $k$  wählen, der einen Bereich  $Y = y \pm U = y \pm k \cdot u_c(y)$  entsprechend einem bestimmten Grad des Vertrauens liefert, z.B. 95% oder 99%. Ebenso wäre man gern in der Lage, den mit einem Bereich assoziierten Vertrauensgrad für einen gegebenen Wert von  $k$  eindeutig angeben zu können. Das lässt sich in der Praxis jedoch nicht leicht realisieren, denn dafür sind umfassende Kenntnisse der Wahrscheinlichkeitsverteilung  $g(y)$  erforderlich, die durch das Messergebnis  $y$  und die kombinierte Standardunsicherheit  $u_c(y)$  charakterisiert wird. Leider liefert die Unsicherheitsfortpflanzungsformel (siehe Modul MU-04) nur die Standardunsicherheit  $u_c(y)$ , was ein Mass für die Breite der Verteilung  $g(y)$  der Werte von  $Y$  ist, aber sie gibt keine Auskunft über die Form dieser Verteilung.

### 2.1 Wichtigkeit der Bestimmung der Verteilung

Es wurde vorher gesehen (siehe Abschnitt 1.3), dass ein Intervall  $[\mu - \sigma, \mu + \sigma]$  um den Mittelwert  $\mu$  einer Normalverteilung 68.3% der Gesamtfläche unterhalb der Kurve enthält. Wenn man jetzt ein gleich grosses Intervall  $[\mu - \sigma, \mu + \sigma]$  für eine Rechteckverteilung betrachtet, dann entspricht es nur noch 57.7% der Gesamtfläche unterhalb der Verteilungskurve. Für eine Dreieckverteilung hat dieses Intervall einen Vertrauensgrad von 65%.



**Figur 3:** Intervalle der gleichen Breite  $2\sigma$  um den Mittelwert  $\mu$  haben unterschiedliche Vertrauensgrade je nach Wahl der Verteilung. Ordinatenwerte sind für  $\sigma = 1$  gezeigt.

Anhand dieser drei Beispiele erkennt man, dass die alleinige Angabe einer Standardabweichung  $\sigma$  keinen eindeutigen Vertrauensgrad für das zugehörige Intervall definiert. Man sollte also zusätzlich entweder die zugehörige Verteilung oder den Vertrauensgrad selber angeben.

## 2.2 Methoden zur Bestimmung der Verteilung

Offenbar hängt  $g(y)$  sowohl von der Funktion  $f$ , die  $Y$  als Funktion der Eingangsgrössen  $X_i$  angibt,  $Y = f(X_1, \dots, X_N)$ , wie auch von den Verteilungen  $g_i(x_i)$ ,  $i = 1, \dots, N$  der einzelnen Eingangsgrössen  $X_i$  ab. Im Allgemeinen ist es nicht möglich,  $g(y)$  analytisch zu bestimmen, aber numerische Methoden können hier behilflich sein (siehe Modul MUS-03).

Andererseits, unter gewissen Annahmen kann man aber einen brauchbaren Ansatz für  $g(y)$  finden, ohne die Verteilung exakt bestimmen zu müssen. Diese Näherungsmethode wird vom GUM verwendet und wird jetzt hier eingeführt. Die Verwendung einer Näherung rechtfertigt sich, da die Berechnung von Intervallen mit bestimmten Vertrauensgraden im besten Fall sowieso nur eine Approximation darstellt.

Zwei neue Begriffe werden nötig sein, um die GUM Methode einzusetzen: der zentrale Grenzwertsatz und die  $t$ -Verteilung. Diese Begriffe werden auf detaillierter Art im Anhang präsentiert, und hier werden sie nur kurz eingeführt.

## 2.3 Kurze Präsentation des zentralen Grenzwertsatzes

Dieser Abschnitt gibt eine kurze Einführung zum zentralen Grenzwertsatz. Eine detailliertere Präsentation wird im Anhang B.2.1 gegeben.

Der zentrale Grenzwertsatz, wie sein Name zeigt, ist in der Statistik sowie in viele wissenschaftliche Gebiete sehr wichtig. Dieser Satz rechtfertigt die zentrale Bedeutung der Normalverteilung in der Beschreibung von vielen statistischen Modellen, von Naturphänomenen, usw. Dieser Satz zeigt, dass alle Systeme die viele vielfältige, nicht in Details bekannte Einflüsse unterworfen sind, am besten durch die Normalverteilung beschrieben werden.

Auf dem Gebiet der Messunsicherheit angewendet impliziert der zentrale Grenzwertsatz, dass wenn eine Messung viele Einflussgrössen hat, dann wird das Messresultat durch eine Normalverteilung beschrieben, unabhängig davon, wie die Eingangsgrössen verteilt sind. Dies gilt unter folgenden Voraussetzungen:

- Die Funktion  $f$ , die den Zusammenhang zwischen dem Messresultat und den Eingangsgrössen angibt, sollte nicht zu stark nichtlinear sein
- Die Eingangsgrössen sollten unabhängig sein
- Die Anzahl Eingangsgrössen sollte gross sein
- Es darf nicht sein, dass wenige Eingangsgrössen die Unsicherheit auf dem Messresultat dominieren – in anderen Worten, die Eingangsgrössen sollten etwa gleich grosse Unsicherheiten aufweisen (nach Multiplikation mit den entsprechenden Empfindlichkeitskoeffizienten).

## 2.4 Kurze Präsentation der $t$ -Verteilung

Die Normalverteilung ist adäquat zur Beschreibung von grossen Datenstichproben, aber kleine Stichproben werden eher durch eine sogenannte  $t$ -Verteilung beschrieben. Diese Verteilung sieht wie eine Glockenkurve aus, wie die Normalverteilung auch, und sie hängt von einem Parameter  $\nu$  ab, der Anzahl der Freiheitsgrade genannt wird. Wenn die Anzahl der Freiheitsgrade gross wird, wird die  $t$ -Verteilung gleich der Normalverteilung. Man kann also sagen, dass die  $t$ -Verteilung eine Verallgemeinerung der Normalverteilung ist.

In der GUM Methode wird eine  $t$ -Verteilung, deren Anzahl der Freiheitsgrade von der Anzahl der Freiheitsgrade der Eingangsgrössen abhängt, zur Beschreibung der Werte des Messresultates verwendet. Dieses Modell ist eine Approximation, die sich solange rechtfertigt, dass die Voraussetzungen des zentralen Grenzwertsatzes gültig sind. Für eine höhere Anzahl der Freiheitsgrade, z.B. mehr als fünfzehn, ist die Verwendung von einer Normalverteilung zur Beschreibung der Werte des Messresultates gerechtfertigt.

Eine detailliertere Präsentation der  $t$ -Verteilung wird im Anhang B.2.2 gegeben.

### 3. Praktische Berechnung der erweiterten Unsicherheit

Als Ausgangssituation wird angenommen, dass für alle Eingangsgrößen  $X_i$  die Schätzwerte  $x_i$  mit Standardunsicherheiten  $u(x_i)$ ,  $i=1, \dots, N$  ermittelt wurden, und die Messgrösse  $Y$  durch  $y = f(x_1, \dots, x_N)$  mit kombinierter Standardunsicherheit  $u_c(y) = \sqrt{\sum_{i=1}^N c_i^2 \cdot u^2(x_i)}$  berechnet wurde. Dabei sind die  $c_i$  die Empfindlichkeitskoeffizienten  $c_i = \frac{\partial f}{\partial x_i}$ .

#### 3.1 Fall einer Normalverteilung

Man möchte das Messresultat mit einer erweiterten Unsicherheit mit 95% Vertrauensgrad ausdrücken. Für eine Normalverteilung wurde vorhergesehen, dass ein Erweiterungsfaktor  $k = 2$  den gewünschten Vertrauensgrad entspricht, und man kann also  $U = k \cdot u_c(y) = 2 \cdot u_c(y)$  setzen.

Man kann das Messresultat auch in einen kleinen Abschnitt folgendermassen formulieren (z.B. für eine Längenmessung  $L$ ):

*$L = (20.000670 \pm 0.000042) \text{ mm}$ , wobei die angegebene Messunsicherheit das Produkt der kombinierten Standardunsicherheit  $u = 21 \text{ nm}$  mit einem Erweiterungsfaktor  $k = 2$  ist, der auf der Normalverteilung beruht. Der Messwert ( $L$ ) und die dazugehörige erweiterte Messunsicherheit ( $U = k \cdot u$ ) geben den Bereich ( $L \pm U$ ) an, der den Wert der gemessenen Grösse mit einer Wahrscheinlichkeit von ca. 95% enthält.*

#### 3.2 Fall einer $t$ -Verteilung

Im allgemeinen Fall der GUM-Methode wird eine  $t$ -Verteilung angenommen, aber das Prinzip bleibt unverändert, nur den Erweiterungsfaktor wird anders sein beim gleichen Vertrauensgrad.

Man nimmt also an, dass die Werte, die die Messgrösse  $Y$  annimmt durch eine  $t$ -Verteilung beschrieben werden, die von einem Parameter  $\nu$  (die Anzahl der Freiheitsgrade) abhängt.

Zur Ermittlung der Anzahl der Freiheitsgrade bzw. des Erweiterungsfaktors geht man folgendermassen vor:

- 1) Ordne jeder Standardunsicherheit  $u(x_i)$  der Eingangsgrößen eine Anzahl Freiheitsgrade  $\nu_i$  zu. Im Abschnitt B.3.1 werden Beispiele für typische Eingangsgrößen, deren Standardabweichung nach Typ A und B ermittelt wurde, gegeben.
- 2) Rechne die effektive Anzahl der Freiheitsgrade  $\nu_{\text{eff}}$  von  $u_c(y)$  unter Verwendung der Welch-Satterthwaite-Formel (B.1) aus:

$$\nu_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^n \frac{u_i^4(y)}{\nu_i}}, \quad \text{wobei} \quad u_i(y) \equiv c_i \cdot u(x_i)$$

Runde  $\nu_{\text{eff}}$  auf die nächstniedrigere ganze Zahl.

- 3) Benutze Tabelle 2 auf Seite 23, die den Zusammenhang zwischen dem Vertrauensgrad und der Breite des Unsicherheitsintervalls für die  $t$ -Verteilung angibt:
  - a. Entweder legt man einen erwünschten Grad des Vertrauens  $p$  fest, sucht diesen Wert in der Tabelle auf der ersten Zeile ganz oben, und liest in der gleichen Spalte den Wert des zugehörigen Erweiterungsfaktor  $k_p$  auf der Zeile, die  $\nu = \nu_{\text{eff}}$  entspricht, ab. Wenn nötig interpoliert man zwischen zwei tabellierten Werten.

- b. Oder man wählt einen Erweiterungsfaktor  $k_p=2$  oder  $k_p=3$ , sucht diesen Wert in der Tabelle auf der Zeile  $\nu = \nu_{eff}$ , und liest der Spalte zuoberst den zugehörigen Vertrauensgrad  $p$  ab. Wenn nötig interpoliert man zwischen zwei tabellierten Werten.

Im Grenzfall eines grossen  $\nu_{eff}$  geht die  $t$ -Verteilung in die Normalverteilung über (letzte Zeile in der Tabelle mit  $\nu=\infty$ ).

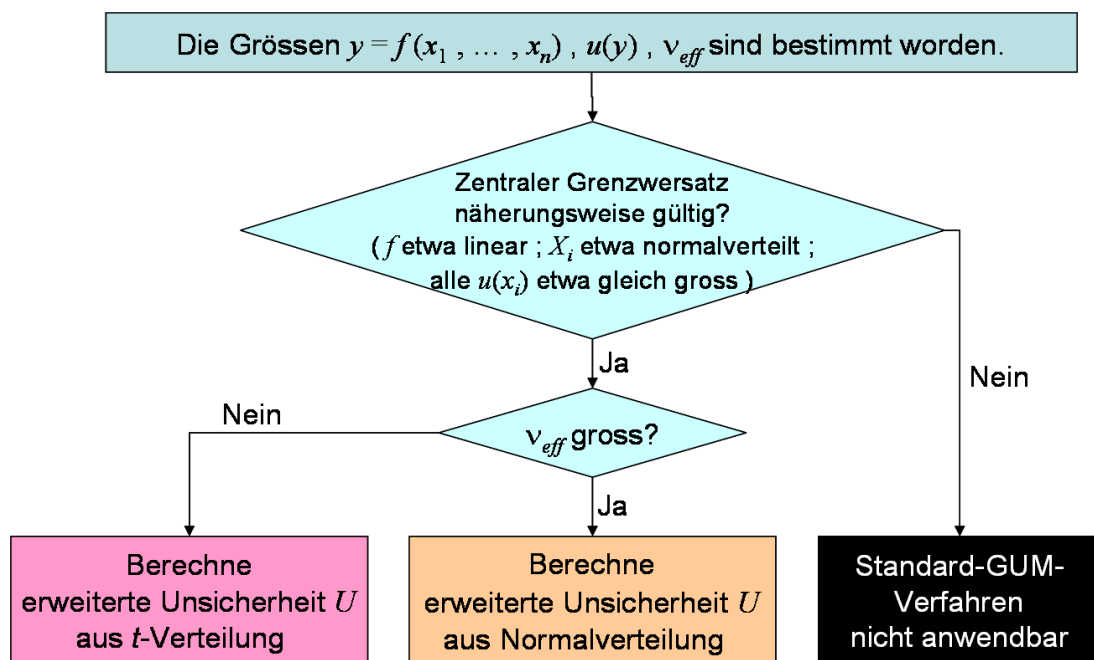
- 4) Man protokolliert das Messergebnis mit seiner erweiterten Unsicherheit  $U$  unter Verwendung einer Formulierung der Art (z.B. für eine Länge  $L$ ):

$L = (20.000670 \pm 0.000049) \text{ mm}$ , wobei die Zahl nach dem Formelzeichen  $\pm$  den Zahlenwert einer erweiterten Unsicherheit  $U = k \cdot u$  angibt;  $U$  errechnet sich aus einer kombinierten Standardunsicherheit  $u = 21 \text{ nm}$  und einem Erweiterungsfaktor  $k = 2.31$ , der auf der  $t$ -Verteilung für  $\nu = 8$  Freiheitsgrade beruht und ein Intervall angibt, das einem geschätzten Vertrauensgrad von 95% entspricht.

Man merke, dass die erweiterte Unsicherheit, die mit einer  $t$ -Verteilung erhalten wird, höher als mit einer Normalverteilung bei identischen Standardunsicherheit und Vertrauensgrade ist. Die Verwendung der Normalverteilung kann also zu einer Unterschätzung der Unsicherheit führen, wenn die Anzahl der Freiheitsgrade klein ist.

### 3.3 Graphische Zusammenfassung

Dieses GUM-Standardverfahren ist streng genommen nur dann gültig, wenn gewisse Voraussetzungen erfüllt sind (es handelt sich im Wesentlichen um die Gültigkeitsbedingungen des so genannten „zentralen Grenzwertsatzes“, siehe Abschnitt 2.3). Wenn diese Voraussetzungen nicht alle ganz erfüllt sind, liefert dieses Verfahren meistens noch vertretbare Resultate. Aber in Fällen, wo die Voraussetzungen markant verletzt sind, könnte eine falsche erweiterte Unsicherheit resultieren. Zum Beispiel, wenn  $f$  starke Nichtlinearitäten aufweist, oder wenn  $u_c(y)$  von einer einzelnen Unsicherheitskomponente  $u_i(y)$  dominiert wird, die aus einer Rechteckverteilung ermittelt wurde, wäre die so erhaltene erweiterte Unsicherheit unzuverlässig.





## Anhang

Hier wird nun für den interessierten Leser eine strengere Herleitung der Begriffe, die in diesem Modul eingeführt wurden gegeben. Selbstverständlich handelt es sich um eine einfache Einführung in dieser Thematik und ist kein Ersatz für ein spezialisiertes Buch.

### A Begriffe aus der Statistik und Verteilungen

#### A.1 Unsicherheit der Standardunsicherheit

Dieser Abschnitt führt Grundbegriffe über Statistik und Stichproben ein. Er dient u.a. dazu, die Formel (A.8) zu begründen, die später benötigt wird, um eine Anzahl Freiheitsgrade den Eingangsgrößen mit Typ B Standardunsicherheit zuzuordnen (siehe Abschnitt B.3.1). Wichtige Begriffe werden im Text definiert und sind bei ihrem ersten Auftritt unterstrichen.

#### A.2 Rückblick: Erwartungswert und Varianz

Sei  $\{x_i\}$  die Menge aller möglichen Messwerte einer physikalischen Grösse  $X$  (die Grundgesamtheit). Diese Grundgesamtheit enthalte  $M$  Elemente (meistens ist  $M = \infty$ ), die auch Zufallsvariablen genannt werden, da bei jeder Messung eine Vielzahl nicht kontrollierbarer Einflussfaktoren dafür sorgen, dass nacheinander ermittelte Werte zufälligen Schwankungen unterliegen.

Der Mittelwert  $\mu = \mu(x)$  (oder Erwartungswert  $E(x)$ ) von  $X$  ist dann

$$E(x) = \mu = \frac{1}{M} \sum_{i=1}^M x_i \quad . \quad (\text{A.1})$$

Die Varianz  $\sigma^2 = \sigma^2(x)$  (auch  $V(x)$  geschrieben) der Zufallsvariablen  $X$  ist das Mittel der quadratischen Abweichungen der Werte vom Mittelwert  $\mu$ :

$$V(x) = \sigma^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2 \quad . \quad (\text{A.2})$$

$\sigma^2$  ist ein Mass für die Streuung der Werte um den Mittelwert. Die Quadratwurzel der Varianz  $\sigma$  heisst Standardabweichung.

In der Praxis hat man keinen Zugang zu allen Elementen der Grundgesamtheit und man muss die Untersuchung auf eine reduzierte Auswahl  $n$  von Elementen (eine Stichprobe) beschränken, um eine Aussage über die Grundgesamtheit zu machen. Diese  $n$  Einzelmessungen werden zufällig aus der Grundgesamtheit aller möglichen Messwerte der Grösse  $X$  „ausgelost“.

Als Schätzung für den unbekannten Mittelwert  $\mu$  wird der Stichproben-Mittelwert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{A.3})$$

verwendet. Die Varianz der Stichprobe ist gegeben durch

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (\text{A.4})$$

sie ist eine Schätzung für die Varianz  $\sigma^2$  der Grundgesamtheit.

### A.3 Standardabweichung des Mittelwertes

Der Mittelwert  $\bar{x}$  von den  $n$  Stichprobenwerten ändert sich von Stichprobe zu Stichprobe, und er kann deshalb selber als Zufallsvariable aufgefasst werden. Es stellt sich somit die Frage nach der Streuung dieser Zufallsgrösse, wenn die Stichproben wiederholt werden. In der Statistik wird gezeigt, dass der Erwartungswert des Mittelwertes  $\mu(\bar{x})$  einer Stichprobe durch den Erwartungswert der Einzelwerte  $\bar{x}$  geschätzt wird und die Varianz des Mittelwerts einer Stichprobe  $\sigma^2(\bar{x})$  wird durch die Varianz der Einzelwerte dividiert durch den Stichprobenumfang  $n$  geschätzt

$$s^2(\bar{x}) = \frac{s^2(x)}{n} \quad . \quad (\text{A.5})$$

Entsprechend ist eine Schätzung der Standardabweichung des Mittelwerts einer Messreihe gegeben durch

$$s(\bar{x}) = \frac{s(x)}{\sqrt{n}} \quad . \quad (\text{A.6})$$

### A.4 Standardabweichung der Standardabweichung

Nun haben wir wiederum die Situation, dass  $s(\bar{x})$  sich von Stichprobe zu Stichprobe ändert. Somit ist diese Grösse selber auch eine Zufallsvariable. Wenn man annimmt, dass die Werte  $x$  in der Grundgesamtheit normalverteilt sind, kann man zeigen, dass der Erwartungswert  $E[s(\bar{x})] \approx \sigma(\bar{x})$  und die Standardabweichung  $\sigma[s(\bar{x})]$  zu  $s(\bar{x})$  näherungsweise gegeben ist durch [2]

$$\sigma[s(\bar{x})] \approx \frac{\sigma(\bar{x})}{\sqrt{2(n-1)}} \quad . \quad (\text{A.7})$$

Die relative Unsicherheit von  $s(\bar{x})$  kann also durch

$$\frac{\sigma[s(\bar{x})]}{\sigma(\bar{x})} \approx \frac{1}{\sqrt{2(n-1)}} \quad (\text{A.8})$$

abgeschätzt werden. Das ist eine „Unsicherheit der Unsicherheit“ von  $\bar{x}$ , die sich aus dem rein statistischen Grund der begrenzten Probenahme ergibt. Folgende Tabelle gibt  $\sigma[s(\bar{x})]/\sigma(\bar{x})$  für einige Werte von  $n$  an. Diese Werte wurden aber anhand des genauen Ausdrucks für  $\sigma[s(\bar{x})]/\sigma(\bar{x})$  und nicht der Näherung (A.7) gerechnet.

**Tabelle 1:** Standardabweichung der empirischen Standardabweichung des Mittelwertes  $\bar{x}$  aus  $n$  unabhängigen Beobachtungen einer normalverteilten Zufallsgrösse  $x$  dividiert durch die Standardabweichung des Mittelwertes (aus dem GUM [1], Tabelle E.1).

Anzahl der Beobachtungen $n$	$\sigma[s(\bar{x})]/\sigma(\bar{x})$ (%)
2	76
3	52
4	42
5	36
10	24
20	16
30	13
50	10

Man sieht also, dass es eine ziemlich grosse Anzahl Einzelmessungen braucht, um die „Unsicherheit der Unsicherheit“ klein zu machen.

## A.5 Verteilungsfunktion und Wahrscheinlichkeitsdichte

Wenn man nun die  $n$  Elemente der oben erwähnten Stichprobe in nicht-abnehmender Reihenfolge ordnet und dann in  $r$  diskrete Werteklassen aufteilt, dann bekommt man eine Häufigkeitsverteilung. Jede Klasse  $j$  enthält die Elemente  $x_{i_j}$ , deren Wert im entsprechenden Bereich  $[x_j^{\min}, x_j^{\max})$  liegt, wobei die obere Grenze der Klasse  $j$  mit der unteren Grenze der Klasse  $j+1$  übereinstimmt:  $x_j^{\max} = x_{j+1}^{\min}$ . Jede Klasse wird durch ihren Medianwert  $\langle x_j \rangle$  bezeichnet (z.B. wenn eine Klasse die Werte von  $x_j^{\min} = 2.1$  bis  $x_j^{\max} = 2.2$  enthält, dann ist  $\langle x_j \rangle = 2.15$ ).  $\langle x_j \rangle$  stimmt nicht unbedingt mit dem Mittelwert der  $x_{i_j}$  in dieser Klasse überein.

Alle einzelnen Werte  $x_i$ ,  $i = 1, \dots, n$  haben hier das gleiche Gewicht 1, sodass eine Klasse  $j$ , die  $n_j$  Elemente enthält das Gewicht  $n_j$  hat. Man kann dann für jede Klasse ihr relatives Gewicht  $m_j/m = p_j$  ausrechnen. Offenbar ist  $0 \leq p_j \leq 1$  und

$$\sum_{j=1}^r p_j = 1 \quad . \quad (\text{A.9})$$

Im Grenzfall einer sehr grossen Anzahl Elemente  $n$  nennt man den Limes

$$\lim_{n \rightarrow \infty} n_j/n = p_j \quad (\text{A.10})$$

die Wahrscheinlichkeit einen Wert in der Klasse  $j$  zu finden:  $p_j = \Pr[x_j^{\min} \leq X < x_j^{\max}]$ . Die  $r$  Klassen, mit den  $p_j$  versehen, bilden dann eine Wahrscheinlichkeitsverteilung.

Erwartungswert und Varianz der  $\{x_i\}$  sind

$$E(x) = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \left[ \sum_{j=1}^r \left( \sum_{i_j=1}^{n_j} x_{i_j} \right) \right] \cong \frac{1}{n} \left[ \sum_{j=1}^r \langle x_j \rangle \cdot n_j \right] \underset{n \text{ gross}}{\cong} \sum_{j=1}^r \langle x_j \rangle \cdot p_j \quad (\text{A.11})$$

$$\begin{aligned} V(x) &= \frac{1}{n-1} \sum_{i=1}^n (x_i - E(x))^2 = \frac{1}{n-1} \left[ \sum_{j=1}^r \left( \sum_{i_j=1}^{n_j} (x_{i_j} - E(x))^2 \right) \right] \\ &\cong \frac{1}{n-1} \left[ \sum_{j=1}^r (\langle x_j \rangle - E(x))^2 \cdot n_j \right] \underset{n \text{ gross}}{\cong} \sum_{j=1}^r (\langle x_j \rangle - E(x))^2 \cdot p_j \end{aligned} \quad (\text{A.12})$$

Beispiel: Die Stichprobe bestehend aus den Werten

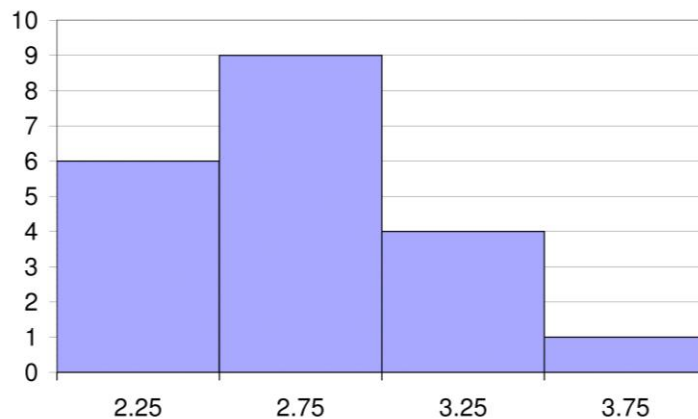
{2.10, 2.31, 3.36, 2.88, 2.90, 2.08, 2.54, 2.79, 3.43, 2.58, 2.29, 2.33, 2.70, 3.09, 3.81, 3.12, 2.85, 2.99, 2.67, 2.35}

wird in nicht-abnehmender Reihenfolge geordnet:

{2.08, 2.10, 2.29, 2.31, 2.33, 2.35, 2.54, 2.58, 2.67, 2.70, 2.79, 2.85, 2.88, 2.90, 2.99, 3.09, 3.12, 3.36, 3.43, 3.81}.

Wählt man Klassen mit 0.50 Breite, dann kann man zwischen 2.00 und 4.00 vier Klassen definieren und die Anzahl Elemente in jeder dieser Klassen aufzählen:

Klassen-nummer $j$	Wertebereich	$\langle x_j \rangle$	Liste der Elemente in der Klasse $x_{ij}$	Anzahl Elemente $n_j$	Gewicht $p_j$
1	[2.00 – 2.50)	2.25	2.08, 2.10, 2.29, 2.31, 2.33, 2.35	6	0.30
2	[2.50 – 3.00)	2.75	2.54, 2.58, 2.67, 2.70, 2.79, 2.85, 2.88, 2.90, 2.99	9	0.45
3	[3.00 – 3.50)	3.25	3.09, 3.12, 3.36, 3.43	4	0.20
4	[3.50 – 4.00)	3.75	3.81	1	0.05



**Figur 4:** Häufigkeitsverteilung.

Die totale Anzahl Elemente in der Stichprobe ist  $6+9+4+1 = 20$ , und die Wahrscheinlichkeiten in jeder Klasse kann man schätzen durch  $p_1 = 6/20 \approx 0.30$ ,  $p_2 = 9/20 \approx 0.45$ ,  $p_3 = 4/20 \approx 0.20$ ,  $p_4 = 1/20 \approx 0.05$  (um die  $p_j$  genau zu bestimmen würde man unendlich viele Elemente benötigen).

Mittelwert der Stichprobe (A.3):  $\bar{x} = (2.10 + 2.31 + \dots + 2.35) / 20 = 2.76$

Varianz der Stichprobe (A.4):

$$s^2 = [ (2.10-2.76)^2 + (2.31-2.76)^2 + \dots + (2.35-2.76)^2 ] / 19 = 0.21$$

Erwartungswert aus der Wahrscheinlichkeitsverteilung (A.11):

$$E(x) = 2.25 \cdot 0.30 + 2.75 \cdot 0.45 + 3.25 \cdot 0.20 + 3.75 \cdot 0.05 = 2.75$$

Varianz der Wahrscheinlichkeitsverteilung (A.12):

$$V(x) = [ (2.25-2.75)^2 \cdot 0.30 + (2.75-2.75)^2 \cdot 0.45 + (3.25-2.75)^2 \cdot 0.20 + (3.75-2.75)^2 \cdot 0.05 ] = 0.175$$

Offenbar ist die Anzahl der Werte  $n = 20$  und die Anzahl der Klassen  $r = 4$  zu klein, um eine gute Wahrscheinlichkeitsverteilung zu konstruieren.

Die bereits definierten Klassen werden durch ihr Gewicht  $p_j$  charakterisiert, was die Wahrscheinlichkeit angibt, einen Wert  $x_i$  von  $X$  zwischen  $[x_j^{\min}, x_j^{\max})$  zu finden:  $p_j = \Pr[x_j^{\min} \leq X < x_j^{\max}]$ . Man kann sich stattdessen auch für die Wahrscheinlichkeit interessieren, einen Wert  $x_i$  von  $X$  kleiner als einen vorgegebenen Wert  $x$  zu bekommen:  $\Pr[X \leq x]$ . Wenn man diesen Ausdruck als Funktion  $G(x)$  von  $x$  auffasst, dann nennt man sie die Verteilungsfunktion der Variablen  $X$  oder auch kumulative Wahrscheinlichkeitsverteilung:

$$G(x) = \Pr[X \leq x] \tag{A.13}$$

Der Wert dieser Funktion ist die Summe der Gewichte der Klassen  $j = 1, \dots, r^{\max}$  mit  $\langle x_j \rangle \leq x$ :

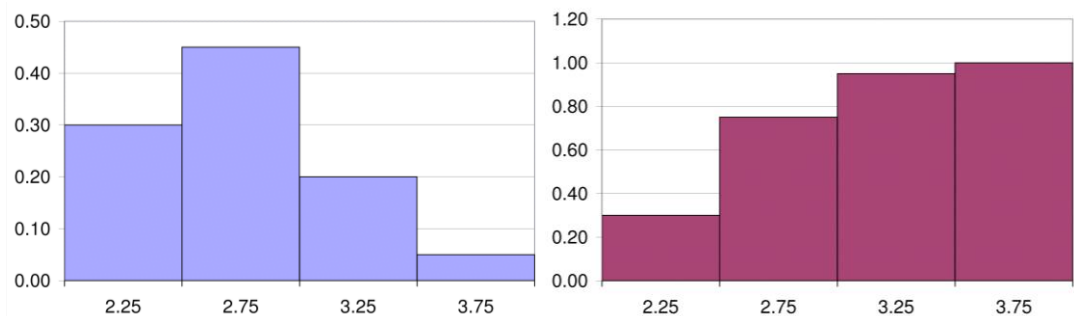
$$G(x) = \Pr[X \leq x] = \sum_{j=1}^{r^{\max}} \Pr[x_j^{\min} \leq X < x_j^{\max}] = \sum_{j=1}^{r^{\max}} p_j \quad . \quad (\text{A.14})$$

Daraus folgt, dass  $0 \leq G(x) \leq 1$  für alle  $x$  ist. Man kann dann die Wahrscheinlichkeit, dass die Werte der Zufallsvariablen  $X$  zwischen  $x_1$  und  $x_2$  liegen, mit

$$\Pr[x_1 \leq X \leq x_2] = \Pr[X \leq x_2] - \Pr[X \leq x_1] = G(x_2) - G(x_1)$$

ausdrücken.

Beispiel: Im obigen Beispiel ist  $G(3.0) = \sum_{j=1}^2 p_j = 0.30 + 0.45 = 0.75$ . In der Stichprobe hat man tatsächlich 15 Werte  $\leq 3.0$ , sodass  $15 / 20 = 0.75$ . Die Histogramme für die Wahrscheinlichkeitsverteilung und die Verteilungsfunktion sehen so aus:



**Figur 5:** Histogramme der Wahrscheinlichkeitsverteilung (links) und der Verteilungsfunktion.

Die Wahrscheinlichkeitsverteilung (Figur 5 links) ist durch Einsetzen des Wertes  $p_j$  in der Säule der Klasse  $j$  konstruiert worden. Um aus dieser Wahrscheinlichkeitsverteilung eine Wahrscheinlichkeitsdichte zu konstruieren muss die Höhe der Säule  $p_j$  durch ihre Breite  $\Delta x_j$  dividiert werden (in diesem Beispiel ist  $\Delta x_j = 0.5$ ). Dann ist  $p_j = (\text{Höhe} \times \text{Breite})$  der Säule ( $\equiv$  Fläche der Säule), und die Verteilungsfunktion (A.14) ist durch

$$G(x) = \sum_{j=1}^{r^{\max}} \left( \frac{p_j}{\Delta x_j} \right) \cdot \Delta x_j$$

gegeben.

### Kontinuierliche Wahrscheinlichkeitsverteilung

Um von einer diskreten Verteilung zu einer kontinuierlichen Verteilung überzugehen, lässt man die Anzahl der Klassen, in denen die Elemente einer diskreten Verteilung eingeteilt sind, sehr gross (gegen Unendlich) und die Breite der einzelnen Klassen gleichzeitig sehr klein werden (gegen Null). Dabei werden die Summen durch Integrale ersetzt und die diskreten Wahrscheinlichkeiten  $p_j$  durch den Ausdruck  $g(x) \cdot dx$ , wobei  $g(x)$  Wahrscheinlichkeitsdichtefunktion heisst (sie entspricht der Höhe der Säule  $p_j / \Delta x_j$  und  $dx$  ihrer Breite  $\Delta x_j$ ). Analog zur Gleichung (A.9) gilt

$$\int_{-\infty}^{+\infty} g(x) \cdot dx = 1 \quad (\text{A.15})$$

(der Integrationsbereich wurde von  $-\infty$  bis  $+\infty$  genommen, aber wenn  $g(x)$  Werte grösser Null nur in einem beschränkten Intervall annimmt, kann man den Integrationsbereich entsprechend verkleinern). Die Zahl  $g(x_0) \cdot dx$  gibt die Wahrscheinlichkeit an, dass der Wert von  $x$

zwischen  $x_0 - dx/2$  und  $x_0 + dx/2$  liegt. Die Ausdrücke für Erwartungswert (A.11) und Varianz (A.12) lauten nun

$$E(x) = \int_{-\infty}^{+\infty} x \cdot g(x) \cdot dx \quad (\text{A.16})$$

$$V(x) = \int_{-\infty}^{+\infty} (x - E(x))^2 \cdot g(x) \cdot dx \quad (\text{A.17})$$

Die Verteilungsfunktion (A.14) wird

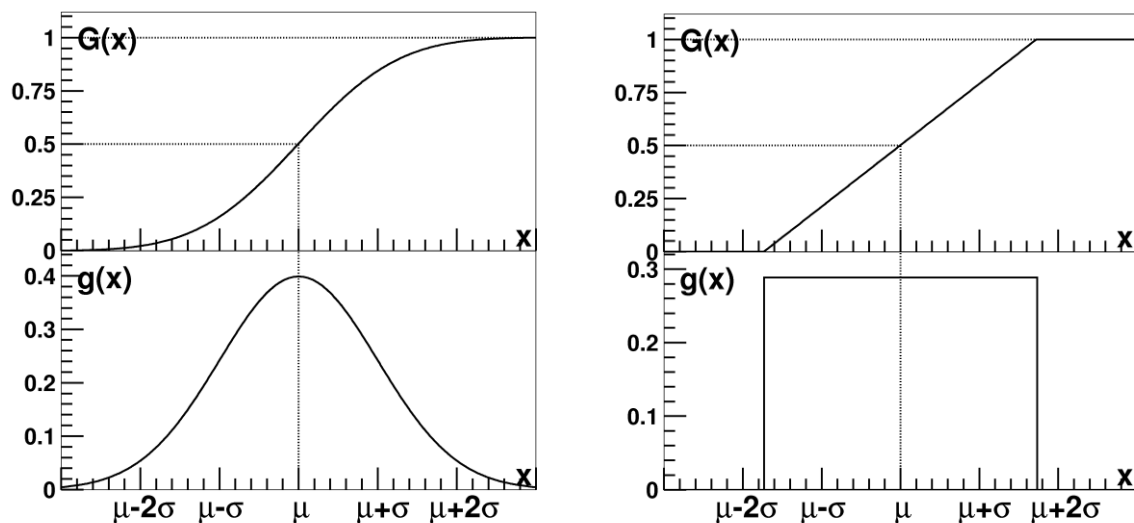
$$G(x) = \int_{-\infty}^x g(z) \cdot dz \quad (\text{A.18})$$

woraus folgt, dass die Wahrscheinlichkeitsdichte  $g(x)$  die Ableitung der Verteilungsfunktion  $G(x)$  ist:

$$\frac{dG(x)}{dx} = g(x) \quad (\text{A.19})$$

Man kann auch die Wahrscheinlichkeit, Werte zwischen  $x_1$  und  $x_2$  zu finden, folgendermaßen ausdrücken:

$$\begin{aligned} Pr[x_1 \leq X \leq x_2] &= Pr[X \leq x_2] - Pr[X \leq x_1] = G(x_2) - G(x_1) \\ &= \int_{-\infty}^{x_2} g(x') \cdot dx' - \int_{-\infty}^{x_1} g(x') \cdot dx' = \int_{x_1}^{x_2} g(x') \cdot dx' \end{aligned} \quad (\text{A.20})$$



**Figur 6:** Wahrscheinlichkeitsdichtefunktion  $g(x)$  (unten) und Verteilungsfunktion  $G(x)$  (oben) für eine Normalverteilung (links) und eine Rechteckverteilung (rechts). Ordinatenwerte von  $g(x)$  sind für  $\sigma = 1$ .

**Beispiel:** Berechnung der Standardabweichung  $\sigma$  einer Rechteckverteilung mit (A.17). Betrachten wir eine Rechteckverteilung mit Mittelwert  $x_0$  und totale Breite  $2a$ . Die Wahrscheinlichkeitsdichte lautet dann

$$g(x) = \begin{cases} 1/(2a) & \text{für } x_0 - a \leq x < x_0 + a \\ 0 & \text{sonst} \end{cases}$$

(die Höhe der Verteilung ist  $1/(2a)$  sodass die Gesamtfläche gleich 1 ist). Dann gilt nach (A.17) (unter Verwendung der Tatsache, dass die so genannte Stammfunktion von  $x^2$  gleich  $x^3/3$  ist):

$$\sigma^2(x) = \int_{-\infty}^{+\infty} (x - x_0)^2 \cdot g(x) \cdot dx$$

$$\begin{aligned}
&= \int_{x_0-a}^{x_0+a} (x-x_0)^2 \cdot \frac{1}{2a} \cdot dx = \frac{1}{2a} \cdot \frac{1}{3} (x-x_0)^3 \Big|_{x_0-a}^{x_0+a} \\
&= \frac{1}{2a} \cdot \frac{1}{3} [(x_0+a-x_0)^3 - (x_0-a-x_0)^3] \\
&= \frac{1}{2a} \cdot \frac{1}{3} [(a)^3 - (-a)^3] = \frac{1}{2a} \cdot \frac{1}{3} \cdot 2a^3 = \frac{a^2}{3}
\end{aligned}$$

Dann ist  $\sigma = \sqrt{a^2/3} = a/\sqrt{3}$ .

## A.6 Vertrauensgrad eines Intervalls

Wenn man eine gegebene Verteilung betrachtet, dann entspricht das Intervall, das eine Halbbreite  $k \cdot \sigma$  hat, einem festen Grad des Vertrauens  $p$ . Für eine Normalverteilung sind diese Zahlen  $p = 68.3\%$ , beziehungsweise  $95.5\%$  und  $99.7\%$  für  $k = 1, 2$  und  $3$ . Sie werden mit Formel (A.20) berechnet:

$$p = \int_{y-k \cdot u(y)}^{y+k \cdot u(y)} g(x) \cdot dx, \text{ wobei } g(x) = \frac{e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi} \cdot \sigma} \text{ mit } \mu \equiv y \text{ und } \sigma \equiv u(y) \quad (\text{A.21})$$

oder äquivalent:

$$p = \int_{-k}^k \tilde{g}(x) \cdot dx \text{ mit } \tilde{g}(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \text{ (skalierte und zentrierte Normalverteilung).}$$

Figur 2 auf Seite 4 zeigt diese verschiedenen Intervalle. Eigentlich wird diese Rechnung für die Normalverteilung numerisch durchgeführt (es gibt keinen analytischen Ausdruck für die Stammfunktion dieser Verteilung), und man verwendet tabellierte Werte wie z.B. die letzte Zeile unten in Tabelle 2 auf Seite 23.

Figur 3 auf Seite 5 zeigt wie der Vertrauensgrad für ein Intervall fester Breite von der betrachteten Verteilung abhängt. In diesem Beispiel sieht man, dass ein Intervall von einem Standardabweichung  $[\mu-\sigma, \mu+\sigma]$  für eine Normalverteilung einen Vertrauensgrad von  $68.3\%$ , für eine Rechteckverteilung  $57.7\%$  und für eine Dreieckverteilung  $65\%$  hat.

Dieses Resultat für die Rechteckverteilung kann sehr leicht ausgehend vom Beispiel im Abschnitt A.5 berechnet werden: die Gesamtfläche unter der Rechteckverteilung der Breite  $2a$  und der Höhe  $1/2a$  ist gleich dem Produkt dieser beiden Werte, also  $2a/(2a) = 1$  und die Fläche innerhalb des 1-Standardabweichung Intervalls um den Mittelwert mit  $\sigma = a/\sqrt{3}$  ist wiederum gleich Breite mal Höhe des Intervalls  $= 2\sigma/(2a) = 2a/(\sqrt{3} \cdot 2a) = 1/\sqrt{3} = 0.577$ .

## B Bestimmung der erweiterten Unsicherheit

### B.1 Einführung

In diesem Kapitel wird die allgemeine Frage untersucht, wie man aus dem Schätzwert  $y$  der Messgrösse  $Y$  und der zugehörigen kombinierten Standardunsicherheit  $u_c(y)$  eine Unsicherheit  $U_p = k_p \cdot u_c(y)$  erhält, die ein Intervall  $y - U_p \leq Y \leq y + U_p$  definiert, das eine feste Überdeckungswahrscheinlichkeit  $p$  hat. Es geht also um die Festlegung des Erweiterungsfaktors  $k_p$ , der ein Intervall um das Messergebnis  $y$  erzeugt, von dem erwartet werden kann, dass er einen grossen, festgelegten Anteil  $p$  der Verteilung von Werten umfasst, die der Messgrösse  $Y$  vernünftigerweise zugeordnet werden können.

Wie oben gesehen, kann dies nur dann genau gemacht werden, wenn die Wahrscheinlichkeitsverteilung  $g(y)$  der Werte von  $Y$  explizit vorliegt. Offenbar hängt  $g(y)$  sowohl von der Funktion  $f$ , die  $Y$  als Funktion der Eingangsgrössen  $X_i$  angibt,  $Y = f(X_1, \dots, X_N)$ , wie auch von den Verteilungen  $g_i(x_i)$ ,  $i = 1, \dots, N$  der einzelnen Eingangsgrössen  $X_i$  ab. Im Allgemeinen ist es nicht möglich,  $g(y)$  analytisch zu bestimmen, aber numerische Methoden können hier behilflich sein (siehe Modul MUS-03). Unter gewissen Annahmen kann man aber einen brauchbaren Ansatz für  $g(y)$  finden, ohne die Verteilung exakt bestimmen zu müssen.

In den meisten praktischen Messsituationen ist die Berechnung von Intervallen mit bestimmten Vertrauensgraden im besten Fall nur eine Näherung. Selbst die empirische Standardabweichung des Mittelwertes aus mehrmaligen Beobachtungen einer Grösse, die durch eine Normalverteilung beschrieben wird, kann für kleinere Stichproben eine grössere Unsicherheit aufweisen (siehe Abschnitt A.4).

## B.2 Grundbegriffe

In diesem Abschnitt werden Begriffe erklärt, die für die Bestimmung der erweiterten Unsicherheit nach der GUM Standardmethode nützlich sein werden.

### B.2.1 Zentraler Grenzwertsatz

Die Unsicherheitsfortpflanzungsformel (siehe Modul MU-04) gibt an, wie gross die kombinierte Standardunsicherheit von  $Y$  ist, ausgehend von den Standardunsicherheiten der Eingangsgrössen und eine Linearisierung der Funktion  $f(X_1, \dots, X_N)$ . Sie gibt aber keine Auskunft über die Verteilung  $g(y)$  der Werte von  $Y$ . Der zentrale Grenzwertsatz erlaubt es, wenn gewisse Voraussetzungen erfüllt sind, eine Aussage über die Form von  $g(y)$  zu machen.

Allgemein gilt: Wenn eine Zufallsvariable  $Y$  sich als eine Linearkombination einer beliebigen Anzahl unabhängiger Zufallsvariablen  $X_i$  ausrechnen lässt,  $Y = \sum_{i=1}^N c_i X_i$ , dann sind Erwartungswert und Varianz von  $Y$  gleich der Summe derjenigen von  $X_i$ :

$$E(Y) = \sum_{i=1}^N c_i \cdot E(X_i) \quad \text{und} \quad \sigma^2(Y) = \sum_{i=1}^N c_i^2 \cdot \sigma^2(X_i) \quad .$$

Diese Ausdrücke wurden im Rahmen von diesem Kurs schon gesehen: es handelt sich einfach um die Formeln für die Funktion  $f$  und die Fortpflanzung der Unsicherheit im Spezialfall wo das Messresultat  $Y$  eine strikt lineare Funktion der Eingangsgrössen  $X_i$  ist.

Diese Aussage gilt für beliebige Verteilungen  $g_i(X_i)$ . Falls zusätzlich alle  $g_i(X_i)$  Normalverteilungen sind, dann ist auch  $Y$  normalverteilt. Falls aber nicht alle  $X_i$  normalverteilt sind, dann kann man keine allgemeine Aussage über die Form von  $g(Y)$  machen. Der zentrale Grenzwertsatz macht eine Aussage für den Fall, wo nicht alle  $X_i$  normalverteilt sind und die Anzahl  $N$  der  $X_i$  gross ist.

**Zentraler Grenzwertsatz:** Eine Linearkombination unabhängiger Zufallsvariablen  $X_i$  ist asymptotisch normalverteilt:

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N c_i \cdot X_i \rightarrow N \left( \sum_{i=1}^N c_i \cdot E(X_i), \sum_{i=1}^N c_i^2 \cdot \sigma^2(X_i) \right)$$

oder auch:

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N c_i \cdot X_i - \sum_{i=1}^N c_i \cdot E(X_i)}{\sqrt{\sum_{i=1}^N c_i^2 \cdot \sigma^2(X_i)}} \equiv \lim_{n \rightarrow \infty} \frac{Y - E(Y)}{\sigma(Y)} \rightarrow N(0,1)$$



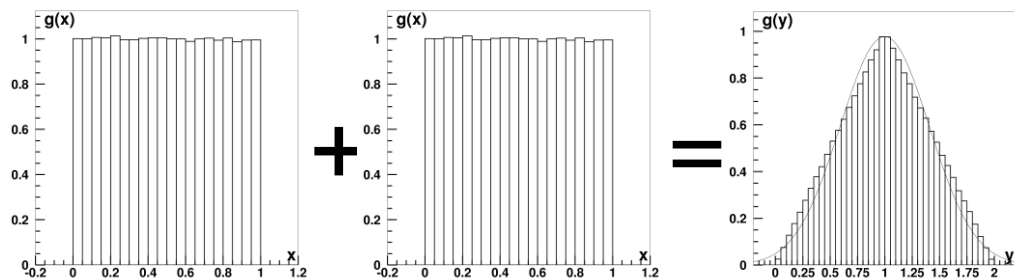
wobei  $N(\mu, \sigma^2)$  eine Normalverteilung mit Mittelwert  $\mu$  und Varianz  $\sigma^2$  bezeichnet (nicht zu verwechseln mit der Anzahl  $N$  der Zufallsvariablen  $N$ !).

Die zwei obigen Ausdrücke sind äquivalent, da, wenn  $Y$  wie  $N(\mu, \sigma^2)$  verteilt ist, ist  $(Y - \mu)$  wie  $N(0, \sigma^2)$  verteilt und  $(Y - \mu)/\sigma$  wie  $N(0, 1)$ .

Die Wichtigkeit dieses Satzes in vielen Anwendungsgebieten wurde im Abschnitt 2.3 unterstrichen.

Diese Aussage gilt näherungsweise für  $N < \infty$  falls die Varianzen  $\sigma^2(x_i)$  der einzelnen  $X_i$  etwa gleich gross sind, d.h. falls es kein  $X_i$  gibt, dass eine viel grössere Varianz als die andere hat. Je näher die Verteilung der  $X_i$  einer Normalverteilung ist, desto weniger  $X_i$  werden benötigt, um eine Normalverteilung für  $Y$  zu erreichen.

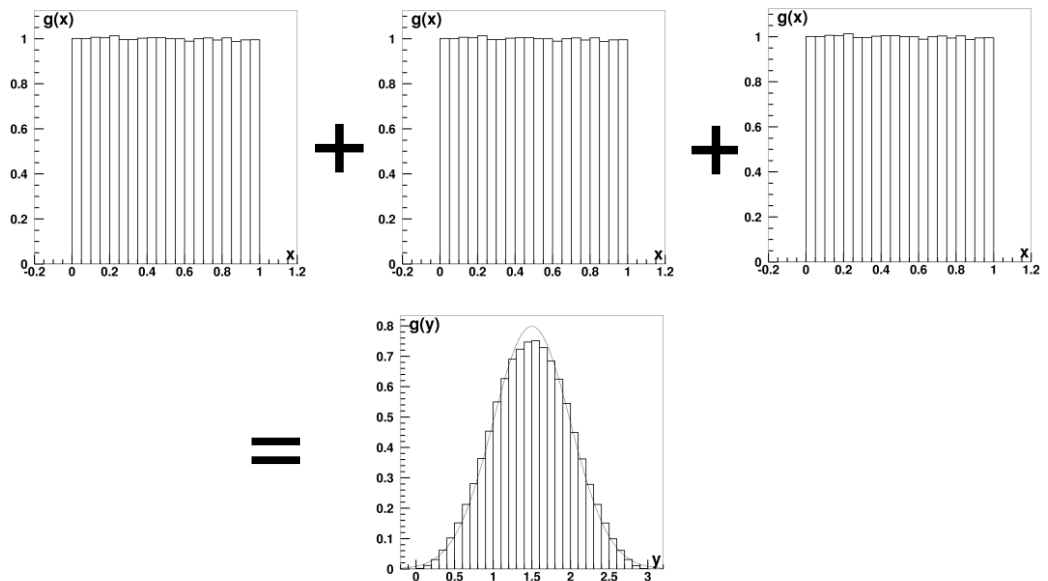
Beispiel: Betrachten wir  $N$  unabhängige Zufallsvariablen  $X_1, \dots, X_N$ , die nach einer Rechteckverteilung mit Erwartungswert 0.5 und totaler Breite 1.0 verteilt sind (die Standardabweichung ist dann  $1/(2\sqrt{3}) = 0.288675 \dots$  für alle). Wenn man die Variable  $Y = X_1 + \dots + X_N$  ausrechnet, dann bekommt man folgende Verteilung für  $N=2$ :



Bemerkung: das „+“ Zeichen steht hier für die Addition von 2 Zufallsvariablen, die nach diesen Verteilungen erhalten werden, und nicht für Addition von 2 Funktionen.

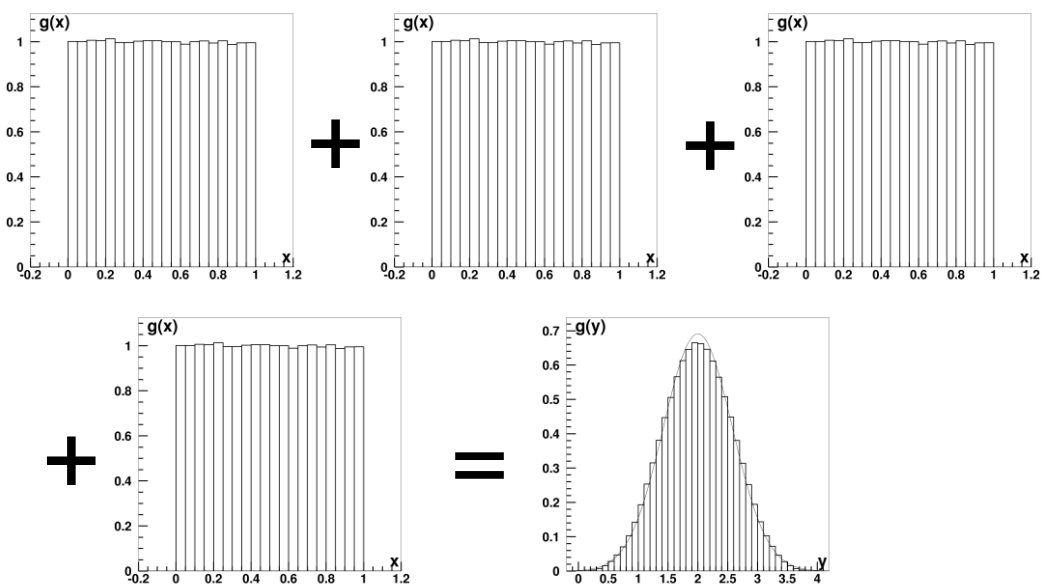
Die resultierende Verteilung ist eine Dreieckverteilung. Eine Normalverteilung gleicher Varianz wurde als dünne Linie auf der Dreieckverteilung überlagert, um zu zeigen, wie gut (bzw. wie schlecht) der zentrale Grenzwertsatz sich in diesem Fall anwenden lässt.

Für  $N=3$  bekommt man folgende Verteilung:

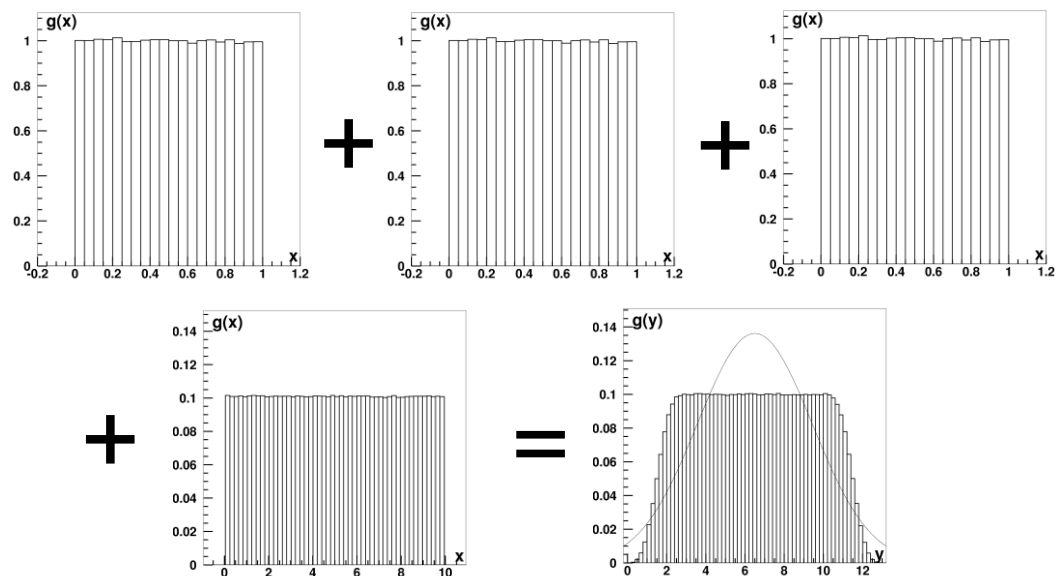


Offenbar ist die Übereinstimmung mit einer Normalverteilung schon bedeutend besser als mit  $N=2$ .

Für  $N=4$  ist sie noch besser:



Die erhaltene Verteilung strebt also relativ schnell gegen eine Normalverteilung mit zunehmendem  $N$ . Dies gilt aber nur deswegen, weil alle Rechteckverteilungen gleiche Breite haben. Wenn man nun eine einzige rechteckverteilte Grösse mit Breite 10.0 und 3 rechteckverteilte Grössen mit Breite 1.0 addiert, dann ist das Resultat viel weniger überzeugend:



Wenn festgelegt werden kann, dass die Bedingungen des zentralen Grenzwertsatzes annähernd erfüllt sind, dann ist die Verwendung eines Wertes für  $k_p$  entsprechend Normalverteilung eine vernünftige erste Näherung zur Berechnung der erweiterten Unsicherheit  $U_p = k_p \cdot u_c(y)$ , die einen Vertrauensgrad  $p$  liefert (vgl. Gleichung (A.21)).

## B.2.2 Die $t$ -Verteilung

In der Praxis wird der Wert  $y$  einer Messgröße  $Y$  aus  $E(Y) \approx y = \sum_{i=1}^N c_i \cdot x_i$  und die kombinierte Standardunsicherheit  $u_c(y)$  aus  $\sigma^2(Y) \approx u_c^2(y) = \sum_{i=1}^N c_i^2 \cdot u^2(x_i)$  geschätzt, wobei  $x_i$  der Schätzwert von  $X_i$  und  $u(x_i)$  die geschätzte Standardabweichung zum Schätzwert  $x_i$  sind. Ausserdem wird die Schätzung des Erwartungswertes  $E(Y)$  untersucht, und nicht die Zufallsvariable  $Y$  selber. Deswegen sollte für die Berechnung eines Intervalls mit einem festgelegten Vertrauensgrad nicht die Verteilung der Variable  $(Y - E(Y))/\sigma(Y)$ , sondern die Verteilung der Variable  $(y - E(Y))/u(y)$  verwendet werden.

Allgemein gilt: Sei  $X$  eine normalverteilte Zufallsgröße mit dem Erwartungswert  $\mu$  und der Standardabweichung  $\sigma$ . Sei  $\bar{x}$  der arithmetische Mittelwert aus  $n$  Beobachtungen  $x_i$  von  $X$  und  $s(\bar{x}) \equiv s(x)/\sqrt{n}$  die empirische Standardabweichung von  $\bar{x}$ . Dann ist die Verteilung der Variablen  $t = (\bar{x} - \mu)/s(\bar{x})$  die  $t$ -Verteilung mit  $v = n-1$  Freiheitsgraden.

Die  $t$ -Verteilung ist also die richtige Verteilung, um Messgrößen, die aus dem Mittelwert wiederholter Beobachtungen erhalten werden (Typ A Unsicherheit), zu beschreiben.

### *Eigenschaften der $t$ -Verteilung*

Die  $t$ -Verteilung wird auch Student-Verteilung genannt. Dieser Name rührt vom englischen Chemiker und Statistiker William Sealy Gosset (1876-1937) her, der seine Arbeiten unter dem Pseudonym Student publizierte. Er musste so vorgehen, weil sein Arbeitsgeber, die Arthur Guinness & Son Brauerei in Dublin, seinen Angestellten verboten hatte, irgendwelche Publikationen zu veröffentlichen, um sicher zu stellen, dass keine Geschäftsgeheimnisse verraten werden. Seine Arbeiten hatte er unter anderem für die Selektion der besten Gersensorte angewendet, ausgehend von der Studie von kleineren Stichproben.

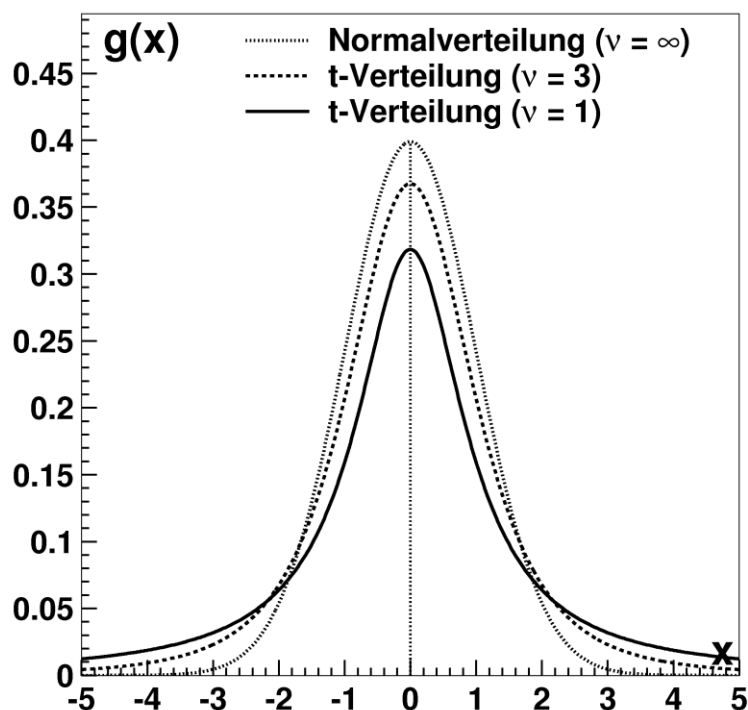
Die  $t$ -Verteilung ist eine Glockenkurve, ähnlich wie die Gaussverteilung (Figur 7). Sie lautet

$$g_v(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{v\pi}} \frac{1}{\left(1 + \frac{t^2}{v}\right)^{\frac{v+1}{2}}}, \text{ wobei } \Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx, \quad z > 0 \quad \text{die so genannte Gamma Funktion ist.}$$

Bemerkung: Für die Gamma Funktion gilt allgemein  $\Gamma(z+1) = z \cdot \Gamma(z)$ . Unter Verwendung von  $\Gamma(1) = 1$  und  $\Gamma(1/2) = \sqrt{\pi}$  kann man dann die Werte der Gamma Funktion für alle ganzen und halbganzen Zahlen (wie es in der  $t$ -Verteilung immer der Fall ist) sehr einfach rekursiv ausrechnen.

In der Praxis werden wir diese Formeln nicht verwenden, sondern wir werden tabellierte Werte dieser Verteilung benutzen.

Es gibt eigentlich nicht nur eine  $t$ -Verteilung, sondern unendlich viele, da sie von einem Parameter  $v = 1, \dots, \infty$ , der Anzahl Freiheitsgrade genannt wird, abhängt. Die Anzahl der Freiheitsgrade ist ein Mass für die Unsicherheit der Standardabweichung  $s(\bar{x})$ , die im Ausdruck  $t = (\bar{x} - E(X))/s(\bar{x})$  vorkommt. Im Grenzfall  $v \rightarrow \infty$  geht sie in die Normalverteilung über:  $t_v \xrightarrow{v \rightarrow \infty} N(0,1)$ .



**Figur 7:** Vergleich von  $t$ -Verteilungen unterschiedlicher Anzahl Freiheitsgrade  $v$  mit einer Normalverteilung  $N(0,1)$ .

Zwei wichtige Feststellungen können aus diesem Graph gewonnen werden:

- Die  $t$ -Verteilung strebt sehr schnell gegen die Normalverteilung: der Gipfel dieser Kurven befindet sich mit  $v = 3$  Freiheitsgraden schon halbwegs zwischen der Verteilung mit der kleinsten Anzahl der Freiheitsgrade ( $v = 1$ ) und der Normalverteilung. Konkret wird eine  $t$ -Verteilung mit etwa zwanzig Freiheitsgraden von einer Normalver-

teilung graphisch ununterscheidbar. Das gerechtfertigt, dass man eine Normalverteilung verwendet, wenn die Anzahl der Freiheitsgrade genügend hoch ist.

- Je höher die Anzahl der Freiheitsgrade, desto höher ist der Gipfel der Glockenkurve. Das impliziert, dass ein Intervall fester Breite (z.B.  $\pm 1$ ) um den Mittelwert (0) eine kleinere Fläche unterhalb der Kurve enthalten wird für eine  $t$ -Verteilung als für eine Normalverteilung. Umgekehrt, wenn man ein Intervall mit vorgegebener Überdeckungswahrscheinlichkeit möchte (z.B. 95%), dann wird er breiter sein müssen für eine  $t$ -Verteilung als für eine Normalverteilung. Das wird für uns die einzige praktische Konsequenz der Verwendung der  $t$ -Verteilung anstelle einer Normalverteilung sein: die Unsicherheit wird bei vorgegebenem Vertrauensgrad grösser.

## B.3 Praktische Ermittlung der erweiterten Unsicherheit

### B.3.1 Effektive Freiheitsgrade

Um die  $t$ -Verteilung anzuwenden muss man zuerst den Parameter  $\nu$  bestimmen. Für die Messgrösse  $Y = f(X_1, \dots, X_N)$  gilt eigentlich die Annahme einer  $t$ -Verteilung nur näherungsweise, denn ihre Standardunsicherheit ist die Summe aus den unterschiedlichen Varianzkomponenten der  $X_i$ , und nicht aus Variablen mit gleichen Varianzen. Jedoch lässt sich die Verteilung von  $Y$  mit Hilfe einer  $t$ -Verteilung mit  $\nu_{eff}$  effektiven Freiheitsgraden ermitteln, die man aus der Welch-Satterthwaite-Formel erhält [3], [4]:

$$\frac{u_c^4(y)}{\nu_{eff}} = \sum_{i=1}^N \frac{u_i^4(y)}{\nu_i} \quad . \quad (B.1)$$

Dabei ist  $u_c(y)$  die kombinierte Standardunsicherheit von  $Y$ , und  $u_i(y) \equiv c_i \cdot u(x_i)$  der Beitrag der Eingangsgrösse  $X_i$  zur Standardunsicherheit von  $Y$ , mit dem Empfindlichkeitskoeffizienten  $c_i$  und der Standardunsicherheit  $u(x_i)$ . Schliesslich ist  $\nu_i$  die zu  $u(x_i)$  gehörige Anzahl der Freiheitsgrade. Wenn der erhaltene Wert von  $\nu_{eff}$  keine ganze Zahl ist, dann wird er auf die nächstniedrigere ganze Zahl abgerundet, bevor man die  $t$ -Verteilung anwendet (so ist man auf der „sicheren Seite“, da einer niedrigeren Anzahl Freiheitsgrade ein breiteres Intervall für eine vorgegebene Überdeckungswahrscheinlichkeit entspricht). Formel (B.1) ist nur für unabhängige Eingangsgrössen  $X_i$  gültig.

Bemerkung: Die Welch-Satterthwaite Formel (B.1) für die „Kombination von Freiheitsgraden“ weist gewisse Ähnlichkeiten mit der Formel für die Kombination von parallel geschalteten elektrischen Widerständen

$$\frac{1}{R} = \sum_{i=1}^N \frac{1}{R_i}$$

auf. Der Unterschied liegt darin, dass bei der letzteren Formel alle Terme das Gewicht „1“ im Zähler haben, wobei in (B.1) die Freiheitsgrade mit der vierten Potenz der zugehörigen Unsicherheitskomponente gewichtet werden. Das impliziert, dass wenn ein  $u_i(y)$  viel grösser als alle anderen ist, dann wird  $\nu_{eff}$  im Wesentlichen gleich dem zugehörigen  $\nu_i$  sein.

Es ist also notwendig, bevor diese Formel zur Berechnung von  $\nu_{eff}$  angewendet wird, die den Standardunsicherheiten  $u(x_i)$  der Eingangsvariablen  $X_i$ ,  $i = 1, \dots, N$  zugehörigen Freiheitsgrade  $\nu_i$  zu bestimmen. Hier sind ein paar Beispiele für typische Eingangsgrössen, deren Standardunsicherheit nach Typ A und B ermittelt wurde, gegeben:

- Für eine Grösse vom Typ A, die als Mittelwert aus  $n$  unabhängigen Beobachtungen  $x_{ij}$ ,  $j = 1, \dots, n$  gewonnen wurde, ist  $v_i = n-1$ .
- Wenn  $n$  unabhängige Beobachtungen zur Ermittlung von Steigung und Achsenabschnitt einer Geraden mit der Methode der kleinsten Quadrate verwendet werden, ist  $v_i = n-2$ .
- Allgemeiner für eine Ausgleichung von  $r$  Parametern und  $n$  Messpunkten nach der Methode der kleinsten Quadrate ist  $v_i = n-r$  die Anzahl der Freiheitsgrade für jeden Parameter.
- Für eine Standardunsicherheit einer Grösse  $X_i$ , die aus einem Zertifikat entnommen wurde, sollte die Anzahl der Freiheitsgrade auch im Zertifikat angegeben werden (wenn das nicht der Fall ist, dann wird  $v_i = \infty$  angenommen).
- Für eine Standardunsicherheit, die nach Ermittlungsmethode B gewonnen wurde und die als genau bekannt angesehen werden kann gilt  $v_i = \infty$ , sonst kann man  $v_i = n-1$  unter Verwendung der Gleichung (A.8) abschätzen:

$$v_i = n - 1 \approx \frac{1}{2} \cdot \frac{\sigma^2(\bar{x})}{\sigma^2[s(\bar{x})]} \approx \frac{1}{2} \left( \frac{\Delta u(x_i)}{u(x_i)} \right)^{-2} \quad (\text{B.2})$$

Die Anzahl der Freiheitsgrade ist ein Mass für die Unsicherheit der Standardabweichung  $s(\bar{x})$  und so sollte man die relative Unsicherheit von  $u(x_i)$  abschätzen, um in (B.2)  $v_i$  zu bekommen. Dies ist eine subjektive Grösse, deren Wert durch eine wissenschaftliche Aussage auf der Grundlage der vorhandenen Informationen gewonnen wird.

Beispiel: Wird die Unsicherheit eines Längenmessgerätes aufgrund von systematischen Einflüssen mit  $u(x_1) = 0.1 \mu\text{m}$  abgeschätzt, und wird diese Abschätzung als bis zu 20% zuverlässig angesehen, dann ist aus (B.2)  $v_1 \approx 1/2 \cdot (\Delta u(x_1)/u(x_1))^{-2} \approx 1/2 \cdot (0.20)^{-2} = 12.5$ . Wenn dann dieses Gerät zur Messung einer Grösse  $X_2$  eingesetzt wird, die als Mittelwert  $x_2$  von  $n=8$  Beobachtungen  $x_{2i}$  gewonnen wird, dann ist  $u(x_2) = 1/\sqrt{n} \cdot \sqrt{\sum_{i=1}^n (x_{2i} - x_2)^2 / (n-1)} \underset{(z.B.)}{=} 0.26 \mu\text{m}$  mit  $v_2 = n-1=7$  Freiheitsgraden. Aus (B.1) kann man die der kombinierten Standardunsicherheit  $u(y) = \sqrt{c_1^2 \cdot u^2(x_1) + c_2^2 \cdot u^2(x_2)}$  zugehörige effektive Anzahl der Freiheitsgrade ausrechnen (Annahme:  $c_1=c_2=1$ ):

$$v_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^2 \frac{u_i^4(y)}{v_i}} = \frac{[c_1^2 \cdot u^2(x_1) + c_2^2 \cdot u^2(x_2)]^2}{\frac{c_1^4 \cdot u^4(x_1)}{v_1} + \frac{c_2^4 \cdot u^4(x_2)}{v_2}} = \frac{(0.1^2 + 0.26^2)^2}{\frac{0.1^4}{12.5} + \frac{0.26^4}{7}} \cong 9.1 \approx 9$$

### B.3.2 Praktische Anwendung der t-Verteilung

Wenn man die Überdeckungswahrscheinlichkeit  $p$  eines vorgegebenen Unsicherheitsintervalls  $U_p = k_p \cdot u_c(y)$  einer Messgrösse  $Y$  um den besten Schätzwert  $y$  bestimmen will, dann muss man die Fläche unterhalb der Kurve im Intervall  $y - U_p \leq Y \leq y + U_p$  bestimmen. Diese Beziehung kann man folgendermassen umschreiben (man sollte eigentlich  $E(Y)$  statt  $Y$  schreiben, da man ein Intervall um den Erwartungswert von  $Y$  sucht):

$$\begin{aligned} & Pr[y - U_p \leq E(Y) \leq y + U_p] \\ &= Pr[y - k_p u_c(y) \leq E(Y) \leq y + k_p u_c(y)] \quad (\text{setze } U_p = k_p \cdot u_c(y) \text{ ein}) \\ &= Pr[-k_p u_c(y) \leq E(Y) - y \leq k_p u_c(y)] \quad (\text{subtrahiere } y) \\ &= Pr[-k_p \leq (E(Y) - y)/u_c(y) \leq k_p] \quad (\text{dividiere durch } u_c(y)) \\ &= Pr[k_p \geq (y - E(Y))/u_c(y) \geq -k_p] \quad (\text{multipliziere mit } (-1)) \end{aligned}$$

$$= \Pr[-k_p \leq (y - E(Y))/u_c(y) \leq k_p] \quad (\text{Beziehung von rechts nach links schreiben}).$$

Man kann hier die oben erwähnte Variable  $t = (y - E(Y))/u_c(y)$  erkennen, die einer  $t$ -Verteilung mit  $\nu$  Freiheitsgraden folgt.

In Worten heisst dies folgendes: Um die Überdeckungswahrscheinlichkeit  $p$  eines Unsicherheitsintervalls  $y \pm k_p \cdot u_c(y)$  zu bestimmen, muss man den Erweiterungsfaktor  $k_p$  mit einer Variablen  $t_p(\nu)$  identifizieren und dann die Überdeckungswahrscheinlichkeit einer  $t$ -Verteilung mit  $\nu = \nu_{\text{eff}}$  Freiheitsgraden bestimmen, d.h. den Wert des Integrals (A.20)

$$p = \int_{-t_p}^{t_p} g_\nu(t) \cdot dt \quad . \quad (\text{B.3})$$

Für eine Normalverteilung entspricht dies Gleichung (A.21). Folgende Tabelle gibt die Werte von  $t_p(\nu)$  für verschiedene Überdeckungswahrscheinlichkeiten  $p$  und verschiedene  $\nu$ -Werte.

**Tabelle 2:** Wert von  $t_p(\nu)$  aus der  $t$ -Verteilung für die Anzahl  $\nu$  der Freiheitsgrade, die ein Intervall  $-t_p(\nu)$  bis  $+t_p(\nu)$  definieren, das den Anteil  $p$  der Verteilung umfasst (aus dem GUM [1], Tabelle G.2).

Anzahl $\nu$ der Freiheitsgrade	Anteil $p$ in %					
	68,27	90	95	95,45	99	99,73
1	1,84	6,31	12,71	13,97	63,66	235,8
2	1,32	2,92	4,30	4,53	9,92	19,21
3	1,20	2,35	3,18	3,31	5,84	9,22
4	1,14	2,13	2,78	2,87	4,60	6,62
5	1,11	2,02	2,57	2,65	4,03	5,51
6	1,09	1,94	2,45	2,52	3,71	4,9
7	1,08	1,89	2,36	2,43	3,50	4,53
8	1,07	1,86	2,31	2,37	3,36	4,28
9	1,06	1,83	2,26	2,32	3,25	4,09
10	1,05	1,81	2,23	2,28	3,17	3,96
11	1,05	1,80	2,20	2,25	3,11	3,85
12	1,04	1,78	2,18	2,23	3,05	3,76
13	1,04	1,77	2,16	2,21	3,01	3,69
14	1,04	1,76	2,14	2,20	2,98	3,64
15	1,03	1,75	2,13	2,18	2,95	3,59
16	1,03	1,75	2,12	2,17	2,92	3,54
17	1,03	1,74	2,11	2,16	2,90	3,51
18	1,03	1,73	2,10	2,15	2,88	3,48
19	1,03	1,73	2,09	2,14	2,86	3,45
20	1,03	1,72	2,09	2,13	2,85	3,42
25	1,02	1,71	2,06	2,11	2,79	3,33
30	1,02	1,70	2,04	2,09	2,75	3,27
35	1,01	1,70	2,03	2,07	2,72	3,23
40	1,01	1,68	2,02	2,06	2,70	3,2
45	1,01	1,68	2,01	2,06	2,69	3,18
50	1,01	1,68	2,01	2,05	2,68	3,16
100	1,005	1,660	1,984	2,025	2,626	3,077
$\infty$	1,000	1,645	1,960	2,000	2,576	3,000

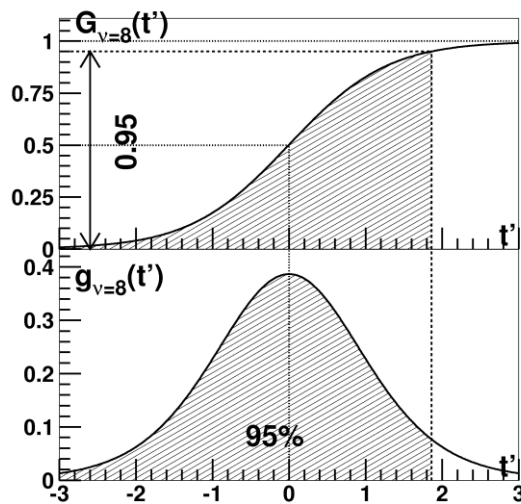
Aus dieser Tabelle kann man z.B. entnehmen, dass für eine Messgrösse mit  $\nu=8$  Freiheitsgraden, ein Intervall mit einem Vertrauensgrad von 95% einen Erweiterungsfaktor  $k_{95}=2.31$  verlangt. Für eine normalverteilte Grösse (d.h.  $\nu=\infty$  Freiheitsgrade, siehe letzte Zeile in der Tabelle) wäre der Erweiterungsfaktor  $k_{95}=1.96$ . Umgekehrt kann man sehen, dass ein Erwei-

terungsfaktor  $k=3$  einem Intervall mit einem Vertrauensgrad von etwa 97.7% entspricht für  $\nu=8$  (lineare Interpolation zwischen den Werten für  $k=2.37$  (95.45%) und  $k=3.36$  (99%)). Für eine Normalverteilung ( $\nu=\infty$ ) ist  $p=99.73\%$  für  $k=3$ .

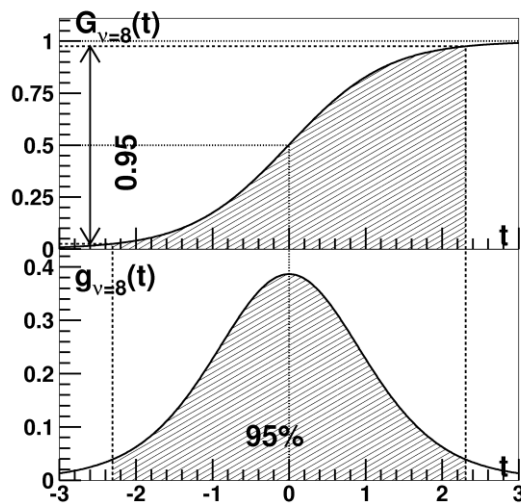
Beispiel: Für das Längenmessgerät im obigen Beispiel wurde  $\nu_{eff} = 9$  gefunden. Dann ist der Erweiterungsfaktor, der einem Intervall mit einer Überdeckungswahrscheinlichkeit von etwa 95% entspricht,  $k_{95}=2.26$  (siehe obige Tabelle). Mit  $u_c(y)=0.28$  ist die Erweiterte Unsicherheit  $U_{95} = 2.26 \cdot 0.28 = 0.63$ .

### B.3.3 Bilaterale und unilaterale Vertrauensintervalle

Manchmal werden die Werte des Integrals  $\int_{-\infty}^{t'} g_{\nu}(t) \cdot dt$  (das ist die Verteilungsfunktion  $G_{\nu}(t'_p)$  (A.18)) mit Integrationsbereich von  $-\infty$  bis  $t'$  tabelliert (unilaterale Situation, siehe Figur 8). In der obigen Tabelle 2 wurde hingegen Ausdruck (B.3) mit Integrationsbereich von  $-t$  bis  $t$  verwendet (bilaterale Situation, siehe Figur 9).



**Figur 8:** 95% Vertrauensintervall für die  $t$ -Verteilung mit  $\nu=8$  Freiheitsgraden. Diese Figur zeigt eine unilaterale Situation: die 5% Restfläche ist ganz rechts, und links reicht das 95% Intervall bis  $-\infty$ .



**Figur 9:** 95% Vertrauensintervall für die  $t$ -Verteilung mit  $\nu=8$  Freiheitsgrade. Diese Figur zeigt eine bilaterale Situation: die 5% Restfläche ist symmetrisch, je 2.5% links und rechts des Vertrauensintervalls verteilt.



Folgendes Beispiel zeigt, wie man ausgehend von einer unilateralen Situation, ein bilaterales Vertrauensintervall angeben kann.

Beispiel: Für  $v=8$  und  $p=95\%$  hat man ein symmetrisches Intervall  $[-t_p, t_p]$  mit  $t_p=2.31$  (siehe Tabelle 2 und Figur 9) aber man findet auch den gleichen Anteil der Gesamtfläche im Intervall  $[-\infty, t'_p]$  mit  $t'_p=1.86$  (Figur 8). Im ersten Fall ist die Restfläche (5%) auf beiden Seiten der Verteilung symmetrisch verteilt (je 2.5%), und im zweiten Fall sind die 5% nur auf der rechten Seite. Um die Grenze  $t_p$  des bilateralen Intervalls ausgehend von tabellierten Werten  $t'_p$  für die unilaterale Situation zu bekommen, muss man die obere Grenze eines Intervalls mit derselben Restfläche auf der rechten Seite nehmen. So ist  $t'_{97.5} = 2.31$  (unilateral) gleich  $t_{95}$  (bilateral), da 2.5% der Gesamtfläche unterhalb der Kurve sich rechts von diesem Wert befindet.

Umgekehrt, wenn man ausgehend von dem Wert  $t_p=2.31$  eines symmetrischen Intervalls  $[-t_p, t_p]$  die zugehörige Überdeckungswahrscheinlichkeit  $p$  aus einer Tabelle von  $t'_p$ -Werte (unilateral) ausrechnen will, dann muss man gemäss (A.20)  $p = G(t_p) - G(-t_p)$  nehmen (d.h. 2 Werte aus der Tabelle subtrahieren). Meistens sind die negativen Werte  $-t_p$  nicht tabelliert, aber wegen der Symmetrie der  $t$ -Verteilung um den Nullpunkt gilt  $G(-t_p) = 1 - G(t_p)$ . Für  $t_p=2.31$  aus einer unilateralen Tabelle ist dann  $G(2.31) = 0.975$  und  $G(-2.31) = 1 - 0.975 = 0.025$ , sodass  $p = G(2.31) - G(-2.31) = 0.975 - 0.025 = 0.95$  für ein symmetrisches Intervall  $[-2.31, 2.31]$ .

### B.3.4 $t$ -Verteilung und Normalverteilung in EXCEL

In EXCEL wird die Funktion  $\text{TVERT}(t;v;2)$  zu Verfügung gestellt, um den Vertrauensgrad eines Intervalls  $[-t, t]$  mit  $v$  Freiheitsgraden auszurechnen. Eigentlich gibt diese Funktion den Anteil der Fläche unterhalb der  $t$ -Verteilung an, der ausserhalb dieses Intervalls liegt, und man muss dann  $1-\text{TVERT}(t;v;2)$  verwenden, um  $p$  zu finden. Umgekehrt rechnet die Funktion  $\text{TINV}(1-p;v)$  die obere Grenze  $t$  des Vertrauensintervalls  $[-t, t]$  mit Überdeckungswahrscheinlichkeit  $p$ .

Für eine Normalverteilung ist die entsprechende Funktion für ein unilaterales Vertrauensintervall  $[-\infty, x]$   $\text{STANDNORMVERT}(x)$ . Um den Vertrauensgrad für einen Erweiterungsfaktor  $k = x$  (bilateral) zu bekommen muss man (wie bereits im obigen Beispiel erklärt)  $\text{STANDNORMVERT}(x) - \text{STANDNORMVERT}(-x)$  berechnen. Umgekehrt rechnet die Funktion  $\text{STANDNORMINV}(p)$  die (unilaterale) obere Grenze  $x'$  des Intervalls  $[-\infty, x']$  mit Überdeckungswahrscheinlichkeit  $p$ , sodass die obere Grenze des symmetrischen Intervalls  $[-x, x]$  mit  $\text{STANDNORMINV}((p+1)/2)$  erhalten wird.

## Literaturverzeichnis

- [1] GUM: BIPM, IEC, IFCC, ISO, IUPAC, OIML, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*, JCGM 100:2008, First edition 2008 (verfügbar über die Suchfunktion auf <http://www.bipm.org>)
- [2] J. F. Kenney and E. S. Keeping, *The distribution of the Standard Deviation*, §7.8 in *Mathematics of Statistics*, Pt. 2, 2<sup>nd</sup> ed., Princeton, NJ: Van Nostrand, pp. 170-173, 1951
- [3] B. L. Welch, *The generalization of 'Student's' problem when several different population variances are involved*, *Biometrika* **34**, 28-35, 1947
- [4] F. E. Satterthwaite, *An approximate distribution of estimates of variance components*, *Biometrics Bull.* **2** (6), 110-114, 1946