

# Und was messen wir morgen?

J. Rissler

**ZUSAMMENFASSUNG** Expositionsmessungen sind mit großen Unsicherheiten behaftet. Daher stellt dieser Artikel statistische Methoden vor, um die Titelfrage zu beantworten. Die Messungen beruhen auf ausreichend großen Stichproben für Tests, um eine Verteilung annehmen zu können. In einem zweiten Schritt werden Parameter dieser Verteilung berechnet. Das erlaubt Aussagen über Messwertbereiche und deren Wahrscheinlichkeiten. Damit kann unter anderem das Einhalten von Grenzwerten überprüft werden. Dieses Vorgehen wird mit Alternativen im deutschen Technischen Regelwerk verglichen.

## 1 Einführung

Das Messen von Expositionswerten ist nur sinnvoll, wenn man mit einiger Sicherheit Aussagen darüber machen kann, wie hoch die Exposition auch zu anderen Messzeitpunkten sein wird, zum Beispiel morgen. Bei Messungen im Alltag ist das kein Problem. Ein Tisch, der heute 1 m lang ist, wird das auch morgen noch sein.

Bei Luftexpositionen ist das anders. Das zeigt allein schon ein Blick auf die erweiterten Messunsicherheitsbereiche der validierten Verfahren, die 30 bis 50 % betragen dürfen [1]. Unser Tisch kann also bei einem solchen Verfahren morgen bei einer Wiederholungsmessung unter identischen Bedingungen 50 cm oder 1,5 m lang sein. Dass man trotzdem zu Recht stolz auf diese Verfahren ist, liegt daran, dass die eigentlich entscheidenden Einflüsse auf den Messwert gar nicht im Messverfahren begründet sind und um ein Vielfaches höher liegen können. Die Unsicherheiten der validierten Verfahren sind also klein genug, aber die Unsicherheiten der Messwerte so groß, dass es sinnlos erscheinen mag, überhaupt messen zu wollen.

Um in dieser Situation Aussagen treffen zu können, ist es nötig, dass wir uns von unserer alltäglichen Messerfahrung lösen und Expositionsmesswerte als Zufallszahlen betrachten. Damit werden sie als Ergebnisse eines statistischen Prozesses interpretiert, der durch eine Verteilungsfunktion beschrieben wird. Anders gesagt, es sind Werte eines Würfels mit sehr vielen Seiten. Daher muss man zuerst so viele Messungen machen, dass man die Werte einer einzigen Verteilungsfunktion zuordnen kann. Man muss also sicherstellen, dass man es nur mit einem Würfel und nicht mit mehreren zu tun hat (Verteilungsprüfung). Dann muss man in der Verteilungsschätzung die Parameter der Verteilungsfunktion bestimmen, also die Eigenschaften des Würfels (Anzahl der Flächen etc.). Danach kann man Bereichen von Messwerten Wahrscheinlichkeiten zuordnen.

## 2 Verteilungsprüfung

Voraussetzung dafür, dass eine Verteilung vorliegt, ist das Eingrenzen der Expositionsbedingungen. Die Messwerte sollten also unter ähnlichen Bedingungen (Materialien, Arbeitsverfahren, Ma-

## And what will we measure tomorrow?

**ABSTRACT** Exposure measurements are subject to large uncertainties. Therefore, this article presents statistical methods to answer the title question. They rest on sufficiently large samples for tests to accept a distribution. In a second step, the parameters of this distribution are calculated. This allows statements concerning ranges of measured values and their probabilities. This can be used to assess the compliance with limit values. This procedure is compared to alternatives in German technical regulations.

schinen, Erfassung etc.) erhoben worden sein. Das entspricht dem Bilden einer „similar exposure group“ in EN 689 [2]. Ob dieses Eingrenzen erfolgreich gewesen ist, kann dann die hier beschriebene Verteilungsprüfung zeigen.

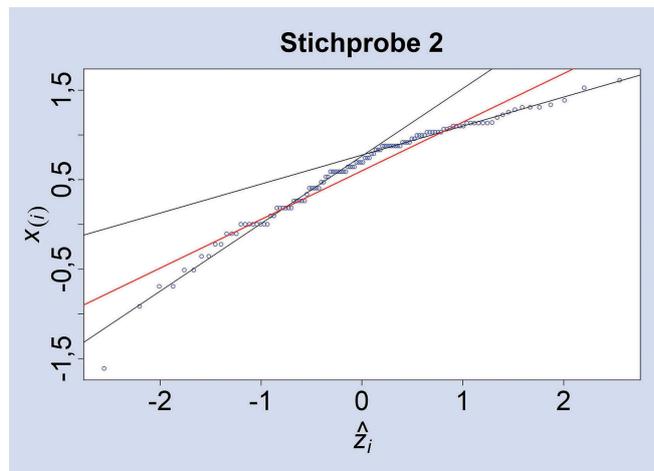
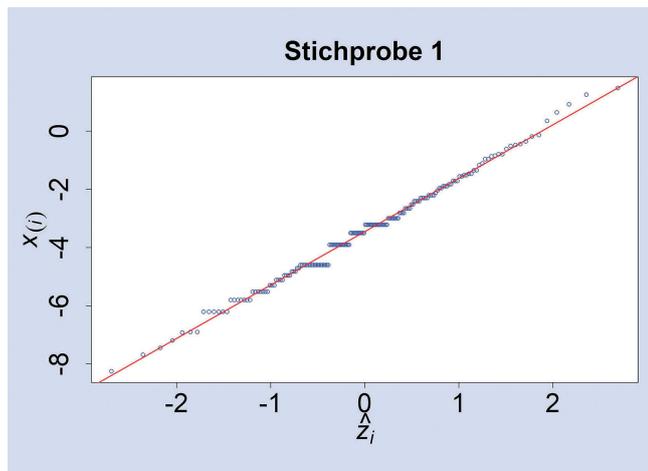
Lehrbücher der Statistik [3, 4] sehen vor, dass man zwischen  $n = 30$  und  $n = 50$  Werte braucht, um zu einer Aussage über eine Verteilung zu kommen. Diese Messwerte sollen zufällig aus allen möglichen Kombinationen von Expositionsbedingungen ausgewählt werden.

Außerdem ist die Annahme gerechtfertigt, dass die Messwerte einer Lognormalverteilung folgen (Anhang A.2). Sie ist nur für positive Werte definiert (Konzentrationen sind immer positiv), und sie hat eine höhere Wahrscheinlichkeit für hohe Werte (rechtsschief).

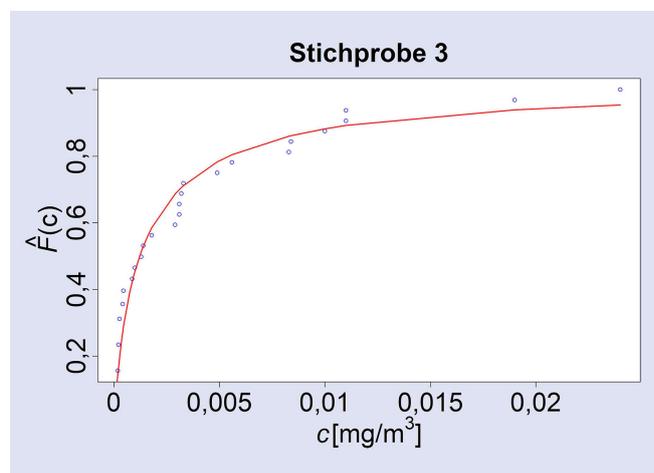
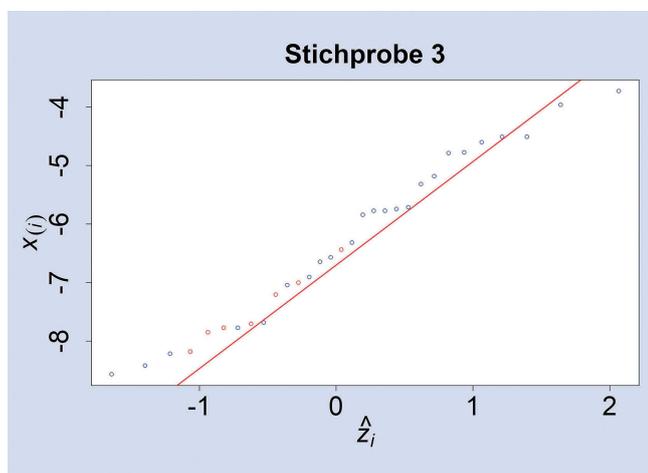
Bei einer Lognormalverteilung sind die natürlichen Logarithmen der Werte normalverteilt, und beide Verteilungsfunktionen werden durch dieselben Parameter  $\mu$  und  $\sigma$  beschrieben. Daher empfiehlt es sich zunächst, die natürlichen Logarithmen der Messwerte  $x_i = \ln(c_i)$  zu bilden, und diese dann der Größe nach zu sortieren ( $x_{(i)}$ ). Die beobachteten Quantile (Messwerte) trägt man dann in einem Quantil-Quantil-Plot gegen die theoretisch erwarteten Quantile einer Standardnormalverteilung  $\hat{z}_i$  auf (QQ-Plot, Anhang Gleichungen (A.2), (A.3)).

Falls die  $x_i$  einer Normalverteilung folgen, und damit die  $c_i$  einer Lognormalverteilung, kann man durch die Werte in dem QQ-Plot eine Regressionsgerade legen, deren Achsenabschnitt ein Schätzwert für den Lageparameter ( $\hat{\mu}$ ) ist und die Steigung einen Schätzwert für den Streuparameter ( $\hat{\sigma}$ ) darstellt. Damit ist der QQ-Plot sehr nützlich, da er nicht nur auf einen Blick zeigt, wie gut eine angenommene Verteilung (Regressionsgerade) die Messwerte beschreibt, sondern auch noch Schätzwerte für Verteilungsparameter liefern kann.

In **Bild 1** kann man links Stichprobe 1 mit  $n = 176$  sehen, die gut mit einer Normalverteilung beschrieben werden kann. Die Werte streuen gleichmäßig um die Regressionsgerade und es sind keine auffälligen Muster zu sehen. Einige Messwerte liegen aufgrund von Rundungsangaben des Analysenlabors bei demselben Wert. Im Gegensatz dazu ist rechts bei Stichprobe 2 ( $n = 118$ ) zu sehen, dass es zwei Äste gibt. Die Regressionsgerade in Rot beschreibt die Daten nicht: Diese Stichprobenwerte werden also von



**Bild 1.** Experimentelle Quantile  $x_{(i)}$  (blaue Kreise) als Funktion der theoretischen Quantile  $z_i$  mit Regressionsgeraden (rot). *Grafik: Autor*



**Bild 2.** Stichprobe mit zensierten Werten und mehreren Bestimmungsgrenzen. Lognormalverteilung mit ML-Schätzer (rote Linien); links: QQ-Plot: Experimentelle Quantile  $x_{(i)}$  (blaue Kreise) als Funktion der theoretischen Quantile  $z_i$ , Bestimmungsgrenzen (rote Kreise); rechts: nicht-parametrische Verteilungsfunktion (blaue Kreise) als Funktion der gemessenen Konzentration  $c$ . *Grafik: Autor*

mehreren Prozessen erzeugt, die auch mit mehreren Verteilungsfunktionen beschrieben werden müssen.

Falls Werte unterhalb der Bestimmungsgrenze vorliegen, können für diese zensierten Werte die Bestimmungsgrenze in den QQ-Plot eingetragen werden. Bei einer Bestimmungsgrenze ist das noch ohne weiteres machbar [2]. Bei mehreren Bestimmungsgrenzen, die etwa durch unterschiedliche Messdauern bedingt sein können, kann man die Werte unterhalb der höchsten Bestimmungsgrenze nicht mehr eindeutig sortieren.

Im QQ-Plot in **Bild 2** sind links die Bestimmungsgrenzen für Stichprobe 3 ( $n = 32$ ) als rote Kreise eingetragen und anstatt einer Regressionsgeraden eine Normalverteilung, wie sie nach der Maximum-Likelihood-Methode (Abschnitt 3) geschätzt wird. Diese Schätzung stimmt gut mit den Werten oberhalb der höchsten Bestimmungsgrenze überein.

Für die Stichproben 1 und 3 legen die QQ-Plots nahe, dass man es dort mit einer (einzigen) Verteilungsfunktion zu tun hat. Diese Vermutung wird nun mit einem Verteilungstest überprüft. Solche Tests vergleichen meist eine Schätzung einer Verteilungsfunktion mit einer Schätzung aufgrund der Stichprobe, die keine Verteilungsparameter braucht.

Falls man es mit einer Stichprobe ohne zensierte Werte zu tun hat, kann man mit dem Shapiro-Wilk-Test [5] überprüfen, ob die logarithmierten Werte einer Normalverteilung folgen. Der Test ist robust für kleine Stichproben, und die Schätzung der Verteilungsfunktion erfolgt durch die Regressionsgerade des QQ-Plots, was ein weiterer Grund ist, den QQ-Plot zu benutzen.

Bei Stichproben mit zensierten Werten bietet sich ein Test nach Efron [6] an. Hier erfolgt die Schätzung der Verteilungsfunktion durch die Maximum-Likelihood-Methode (ML-Methode, Abschnitt 3), die in Bild 2 rechts als rote Linie eingetragen ist. Die blauen Kreise sind die Schätzwerte aufgrund der Stichprobe, die keine Verteilungsparameter braucht. Wie beim QQ-Plot links sprechen auch hier die geringen Abweichungen der Kreise von der Linie dafür, dass eine Verteilung vorliegt. Außerdem kann man im Gegensatz zur linken Abbildung den Vergleich über den gesamten Bereich der Stichprobe vornehmen.

Das erwünschte Ergebnis ist, dass eine (Log-)Normalverteilung vorliegt. Dies ist die Nullhypothese sowohl beim Shapiro-Wilk-Test als auch beim Efron-Test. Daher wird das Signifikanzniveau auf  $\alpha = 0,10$  heraufgesetzt. Wenn die Testswahrscheinlichkeitswerte  $p > \alpha$  ergeben, lehnen wir die Nullhypothese nicht ab und gehen davon aus, dass eine Verteilung vorliegt.

**Tabelle.** ML-Schätzer für Parameter einer Lognormalverteilung, Schätzer für Perzentile; in Klammern sind die Standardabweichungen (sd) und die 70 %-Obergrenzen ( $\theta_{0,70}$ ) angegeben.

Stichprobe		$\hat{\mu}$ (sd)	$\hat{\sigma}$ (sd)	$\hat{q}_{0,90}(\theta_{0,70}) \left[ \frac{\text{mg}}{\text{m}^3} \right]$	$\hat{q}_{0,95}(\theta_{0,70}) \left[ \frac{\text{mg}}{\text{m}^3} \right]$
1	176	-3,46 (0,14)	1,84 (0,10)	0,334 (0,367)	0,656 (0,727)
3	32	-6,70 (0,33)	1,77 (0,27)	0,012 (0,016)	0,025 (0,033)

Die Stichproben in Bild 1 haben für den Shapiro-Wilk-Test  $p$ -Werte  $p > 0,68$  von (Stichprobe 1) sowie  $p < 4 \cdot 10^{-5}$  (Stichprobe 2). Für die Stichprobe 1 wird die Null-Hypothese nicht abgelehnt, für die Stichprobe 2 schon. Der Efron-Test für die Stichprobe 3 ergibt  $p > 0,40$ . Damit ist auch hier die Null-Hypothese nicht abgelehnt.

Die Tests bestätigen also das, was man deskriptiv in Bild 1 und 2 schon gesehen hat. Mehr noch, durch die Abbildungen kann man zusätzliche Informationen bekommen, warum etwa eine Nullhypothese abgelehnt worden ist. So kann man für die Stichprobe 2 versuchen zu ermitteln, worin sich die Messwerte der beiden Äste unterscheiden (unterschiedliche Schichten, Betriebsstätten etc.). Falls eine genauere Einteilung der Stichprobe 2 gelingt, kann es sein, dass jeder Stichprobenteil für sich sinnvoll ausgewertet werden kann.

### 3 Verteilungsschätzung

Die Analysen aus dem vorigen Abschnitt lassen uns davon ausgehen, dass eine einzige Verteilungsfunktion (ein einziger Würfel) ausreicht, um zwei von den drei untersuchten Stichproben zu beschreiben. Daher können wir erst jetzt dazu übergehen, für diese beiden Stichproben die Verteilungsfunktion zu beschreiben.

Die Verteilungsfunktion verknüpft Messwertbereiche mit Wahrscheinlichkeiten. Deren Umkehrfunktion ist die Quantilfunktion, die Wahrscheinlichkeiten mit Messwerten verknüpft. Für die Expositionsbewertung sind das sogenannte 90. und das 95. Perzentil wichtig ( $q_{0,90}$ ,  $q_{0,95}$ ). Das sind Quantile (Messwerte) mit denen die Wahrscheinlichkeiten 0,90 und 0,95 verknüpft sind, diese Messwerte oder kleinere zu beobachten. Sie beschreiben den Bereich, von dem wir ausgehen, dass er mit einer hohen Wahrscheinlichkeit (0,90 oder 0,95) die messbare Exposition enthält. Liegt dieser Bereich unterhalb eines Beurteilungswertes, kann dieser als eingehalten gelten.

Die Perzentile lassen sich also nur mit einer Verteilungsfunktion bestimmen, und dafür braucht man deren Parameter. Bei Stichproben ohne zensierte Werte oder mit nur einer Bestimmungsgrenze sieht die EN 689 [2] vor, dass die Regressionsgerade eines QQ-Plots die Verteilungsfunktion bestimmt. Diese Methode ist einfach vorzunehmen aber nicht robust, um das 90. oder 95. Perzentil zu beschreiben [7].

Daher bietet sich die Maximum-Likelihood-Methode (ML-Methode) an. Diese führt bei normalverteilten Stichproben ohne zensierte Werte (siehe Anhang A.3) zum arithmetischen Mittelwert  $\bar{x}$  und der empirischen Standardabweichung  $s$  als Schätzer für den Lage- und den Streuparameter ( $\hat{\mu} = \bar{x}$ ,  $\hat{\sigma} = s$ ):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

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

Streng genommen hat der ML-Schätzer für die Standardabweichung im Nenner  $1/\sqrt{n}$ , aber mit  $1/\sqrt{n-1}$  ist dieser Schätzer erwartungstreu.

Bei Stichproben mit zensierten Werten werden alle Werte in der Likelihood-Funktion (Gleichung (A.10)) konsistent berücksichtigt [8]. Aus dieser werden die Schätzer für  $\hat{\mu}$  und  $\hat{\sigma}$  durch Anpassung bestimmt, die im Gegensatz zum QQ-Plot aber nur noch numerisch darstellbar ist.

Die Tabelle enthält die Ergebnisse dieser Schätzung für die Stichproben 1 und 3. Die Werte für  $\hat{\mu}$  und  $\hat{\sigma}$  lassen sich in den QQ-Plots nachvollziehen (Bild 1, 2). Aus diesen lassen sich die Perzentilschätzer bestimmen:

$$\hat{q}_{0,90} = \exp\{\hat{\mu} + 1,281 \cdot \hat{\sigma}\} \frac{\text{mg}}{\text{m}^3}$$

$$\hat{q}_{0,95} = \exp\{\hat{\mu} + 1,645 \cdot \hat{\sigma}\} \frac{\text{mg}}{\text{m}^3} \quad (2)$$

Dabei ist 1,281 (1,645) das 90. (95.) Perzentil der Standardnormalverteilung. Die Einheiten sind die Einheiten der ursprünglichen Messwerte  $c_i$ .

Bei Stichprobe 3 liegt  $\hat{q}_{0,95}$  oberhalb des höchsten Messwertes der Stichprobe ( $0,024 \frac{\text{mg}}{\text{m}^3}$ ). Das ist nicht weiter verwunderlich, da etwa 19 % aller Stichproben mit  $n = 32$  das 95. Perzentil nicht einschließen (Gleichung (A.1), Bild (A.1)).

Die Zahlen in Klammern in der Tabelle sind aber das eigentliche Ziel unserer Auswertung, da sie es uns ermöglichen, Aussagen über die Sicherheit der Ergebnisse zu machen. In der Einführung haben wir uns bewusst gemacht, dass eine einzige Messung gar keine Aussage ermöglicht, weder über Messwertbereiche noch über Sicherheiten. Jetzt haben wir aber Informationen über den zu erwartenden Messwertbereich ( $\hat{q}_{0,90}$ ,  $\hat{q}_{0,95}$ ) und darüber, wie sicher diese Aussagen sind. Aufgrund der Varianz in der Stichprobe ergeben sich die Standardabweichungen der Parameter- und damit auch der Perzentilschätzer in der Tabelle. Diese Sicherheit der Aussagen nimmt selbstverständlich mit der Größe der Stichprobe zu, wie man an den Standardabweichungen für  $\hat{\mu}$  und  $\hat{\sigma}$  sehen kann.

Diese fließen auch in die Berechnung der 70 %-Obergrenze  $\theta_{0,70}$  für Stichproben mit zensierten Werten ein (Gleichung (A.8)). Für Stichproben ohne zensierte Werte kann man  $\theta_{0,70}$  direkt berechnen [9] (Gleichung (A.7)). Dabei ist  $\theta_{0,70}$  so gewählt, dass sie bei derselben Verteilung in 70 % der Stichproben mit demselben  $n$ , das wahre  $q_{0,90}$  oder  $q_{0,95}$  überdeckt. Die Wahl von 70 % ist eine Konvention, die auf die EN 689 [2] zurückgeht.

Man kann in beiden Fällen sehen, dass bei der Lognormalverteilung  $\theta_{0,70}$  weiter vom  $\hat{q}_{0,95}$  entfernt ist als vom  $\hat{q}_{0,90}$ . Die Schätzer der hohen Perzentile haben wie erwartet eine höhere Unsicherheit als die der niedrigeren.

## 4 Diskussion

Damit können wir nun die Frage vom Anfang für die beiden Stichproben aus der Tabelle beantworten. Wir gehen davon aus, dass Stichproben unter denselben Bedingungen auch an einem anderen Messtag wieder dazu kommen, dass Lognormalverteilungen vorliegen und dass die Perzentile  $q_{0,90}$  oder  $q_{0,95}$  dieser Verteilungen von den  $\theta_{0,70}$  überdeckt werden. Man kann also die  $\theta_{0,70}$  mit Beurteilungswerten vergleichen.

Zugegeben, 30 bis 50 Messwerte, die für einen sinnvollen Umgang mit Messwerten hoher Streuung nötig sind, bedeuten einen gewissen Aufwand. Dieser muss aber bei Expositionsmessungen in Kauf genommen werden. Daher sollte die Strategie sein, sinnvolle Messreihen mit Zufallsauswahl und ausreichender Stichprobengröße an Arbeitsplätzen mit gleichen Expositionsbedingungen vorzunehmen. Bei Erfolg lassen sich die so gewonnenen Ergebnisse auf andere Arbeitsplätze mit denselben Expositionsbedingungen übertragen, die nicht Teil der Stichprobe gewesen sind.

Man mag aus vielerlei Gründen versucht sein, mit weniger Messwerten auszukommen. Die Überlegung ist dabei, dass auch einfachere Analysen ausreichen, wenn man weit genug vom Beurteilungswert entfernt ist. Die Technische Regel für Gefahrstoffe (TRGS) 402 [10] hält zum Beispiel den Beurteilungswert für eingehalten, wenn ein einzelner, repräsentativ ermittelter Wert weniger als ein Zehntel des Beurteilungswerts beträgt. Leider macht ein Blick auf Gleichung (2) deutlich, dass ein Faktor 10 für diesen Schluss bei weitem nicht ausreicht. Dort ist zu sehen, wie weit das 95. Perzentil von einem „mittleren“ oder „repräsentativen“ Wert wie dem Median  $e^\mu$  entfernt ist: Der Faktor beträgt  $e^{1,645\sigma}$ . Selbst für sehr kleine Werte für  $\sigma$ ,  $\sigma = 1,5$  [11] ist der Faktor 10 schon erreicht. Ein 95. Perzentil liegt also nicht mehr unterhalb des Beurteilungswertes. Für  $\sigma = 2$  beträgt der Faktor schon 26, für  $\sigma = 4$  beträgt er 720.

Die Methode in der TRGS 420 [12] erlaubt kleinere Stichprobengrößen, „wenn niedrige Bewertungsindices mit geringer Streuung der Einzelwerte vorliegen oder wenn sich die Arbeitsbedingungen langfristig wenig ändern“. Dies hat zur Folge, dass meist Stichproben mit  $n = 20$  bis 30 Werten für die Analysen verwendet werden. Dazu muss man wissen, dass die TRGS 420 annimmt, die Stichprobe sei schon die gesamte Verteilungsfunktion. Das ist aber nur sinnvoll, wenn tatsächlich alle Werte identisch verteilt sind, was nicht überprüft wird. Außerdem ist es notwendig, dass die Stichprobe auch das 95. Perzentil abdeckt. Dies ist bei der empfohlenen Bedingung von  $n \geq 72$  mit hoher Wahrscheinlichkeit so (Anhang A.1).

Der Fehler, der gemacht wird, wenn man zu kleineren Stichproben übergeht, ist in Gleichung (A.1) und Bild A.1 zu sehen. Für  $n = 30$  ist in 21 % der Stichproben das 95. Perzentil nicht in der Stichprobe: Das 95. Perzentil wird also falsch und zu klein geschätzt. Bei  $n = 20$  ist das bei 36 % der Stichproben so. Dies ist immer der Fall, unabhängig von der Streuung der Einzelwerte oder der Schwankung der Arbeitsbedingungen.

Dies sind nur zwei Beispiele, bei denen die Probleme zu kleiner Stichproben deutlich werden. Ganz prinzipiell verzichten beide Regelwerke darauf, zu prüfen, ob eine Verteilung vorliegt. Daher können Fälle wie die Stichprobe 2 nicht ausgeschlossen werden, sodass die Regelwerke billiger verfälschte Ergebnisse in Kauf nehmen. Da die Nutzenden dieser Regeln nicht wissen, ob

eine Verteilung vorliegt, wissen sie auch nicht, welche Verteilung vorliegt. Damit können sie unmöglich etwas über Perzentile aussagen. Ebenso unmöglich ist es, mit diesen Methoden die Gretchenfrage im Titel dieser Arbeit zu beantworten.

Je nach Fragestellung muss man nicht das 90. oder 95. Perzentil für die Beurteilung nutzen. Der Median einer Lognormalverteilung ist deutlich einfacher zu beschreiben und vielleicht „repräsentativer“. Es ist auch möglich, mit kleineren Stichproben zu leben, wenn man über die Risiken aufklärt, wie oben beschrieben. Schließlich kann man auch ganz auf Beurteilungswerte und damit auf Messwerte verzichten.

Wenn eine Aufsichtsperson sicherstellen kann, dass die Expositionsbedingungen während einer Messung so sind, dass der eine Messwert, der dabei erzeugt wird, repräsentativ für den Grad der getroffenen Schutzmaßnahmen ist [10], dann liegen genug Informationen vor, um auch ohne diese eine Messung sagen zu können, ob die Schutzmaßnahmen ausreichend sind. Wenn man also die Wahl hat, sollte man die nicht-messtechnische Beurteilung einem Beurteilungsversuch mit großen Unsicherheiten bei zu kleiner Stichprobengröße vorziehen. Der Befund ist damit sicherlich nicht schlechter, und es werden messtechnischen Ressourcen frei, die an anderer Stelle sinnvoller eingesetzt werden können. Zum Beispiel könnten mit den freigewordenen Ressourcen gezielt Messreihen mit ausreichender Stichprobengröße zur sicheren Beurteilung durchgeführt werden.

## Mathematischer Anhang

### A.1 Ordnungsstatistiken

Die gemessenen Konzentrationen  $c'$  müssen dieselbe Einheit haben. Man bildet einheitenlose Werte ( $c_i = c' / \text{unit}$  ( $\text{unit} = \left\{ \frac{\text{mg}}{\text{m}^3}, \frac{\mu\text{g}}{\text{L}}, \dots \right\}$ )), um Logarithmen bilden zu können  $x_i = \ln(c_i)$ .

Man kann die  $x_i$  (oder  $c_i$ ) der Größe nach sortieren, um die Ordnungsstatistiken  $x_{(i)}$  (oder  $c_{(i)}$ ) zu erhalten. Ordnungsstatistiken, die unabhängig und identisch verteilte Zufallsgrößen mit der kumulierten Verteilungsfunktion  $F(x)$  sind, haben für den höchsten Stichprobenwert die Verteilungsfunktion

$$F_{x_{(n)}}(x) = F(x)^n \quad (\text{A.1})$$

Gleichung (A.1) hängt nicht von der Form von  $F(x)$  ab. Sie gilt also für alle Verteilungsfunktionen. Damit kann man bestimmen, mit welcher Wahrscheinlichkeit das 95. Perzentil  $F(q_{0,95}) = 0,95$  kleiner ist als  $x_{(n)}$ . In **Bild A.1** ist zu sehen, dass für  $n = 72$   $q_{0,95}$  nur zu 2,5 % außerhalb der Stichprobe liegt. Das ist für Methoden wichtig, die  $\hat{q}_{0,95}$  aus den  $x_{(i)}$  schätzen [12].

Eine Methode, um  $p(x_{(i)})$  für normalverteilte  $x_{(i)}$  anzugeben [13], die sich auch in der Normung [14] durchgesetzt hat, ist

$$\hat{p}(x_{(i)}) = \frac{i - \frac{3}{8}}{n + \frac{1}{4}} \quad (\text{A.2})$$

Wenn  $F(x) = F_N(x)$  die kumulierte Verteilungsfunktion für die Standardnormalverteilung ist (Gleichung (A.5) mit  $\mu = 0$ ,  $\sigma = 1$ ), dann ist  $F_N^{-1}(p)$  die entsprechende Umkehrfunktion oder Quantilfunktion. Mit Gleichung (A.2) ergeben sich die theoretisch zu erwartenden Quantile in einem QQ-Plot  $\hat{z}_i$ . Außerdem wird klar, warum eine Regressionsanalyse eine Schätzmethode für  $\hat{\mu}$  und  $\hat{\sigma}$  sein kann:

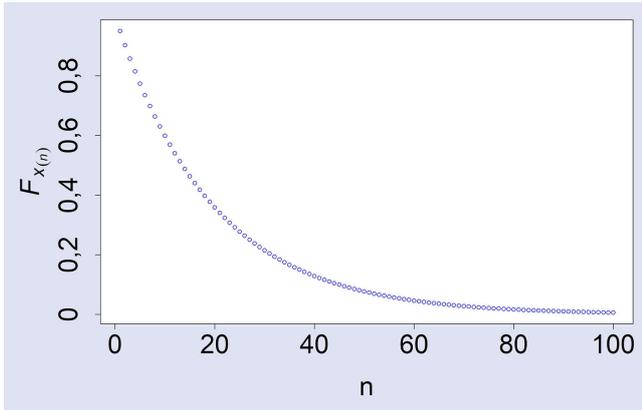


Bild A.1.  $F_{x^{(n)}}(q_{0,95})$  als Funktion von  $n$ . Grafik: Autor

$$F_N^{-1}(\hat{p}(x_{(i)})) = \hat{z}_i$$

$$\hat{z}_i = \frac{x_{(i)} - \hat{\mu}}{\hat{\sigma}} \Leftrightarrow x_{(i)} = \hat{\sigma} \hat{z}_i + \hat{\mu} \tag{A.3}$$

### A.2 Die Lognormalverteilung

Die Wahrscheinlichkeitsdichte  $f_{LN}(c)$  und die kumulierte Verteilungsfunktion der Lognormalverteilung  $F_{LN}(c)$ ,

$$f_{LN}(c) = \frac{1}{\sqrt{2\pi\sigma^2}c} \exp\left[-\frac{(\ln(c) - \mu)^2}{2\sigma^2}\right], F_{LN}(c) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\ln(c) - \mu}{\sigma\sqrt{2}}\right)\right], \tag{A.4}$$

sind dadurch gekennzeichnet, dass sie denselben Lageparameter  $\mu$  und  $\sigma$  Streuparameter haben wie die Wahrscheinlichkeitsdichte  $f_N(x)$  und die kumulierte Verteilungsfunktion  $F_N(x)$  der  $x = \ln(c)$

$$f_N(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right], F_N(x) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right]. \tag{A.5}$$

Da die Logarithmusfunktion eine streng monotone Transformation ist, lassen sich Punktschätzer, etwa Quantile ( $F_{\{LN, LN\}}(q_{\{LN, LN\}, p}) = p$ ), leicht ineinander umrechnen:

$$\hat{q}_{N,p} = \hat{\mu} + z_p \hat{\sigma}$$

$$\hat{q}_{LN,p} = \exp\{\hat{q}_{N,p}\} \tag{A.6}$$

Hier ist  $z_p$  das  $p$ -Quantil der Standardnormalverteilung (Gleichung (A.4) mit  $\mu = 0, \sigma = 1$ ).

Bei Stichproben ohne zensierte Werte wird Gleichung (A.6) mit den Schätzern aus Gleichung (1) gebildet. Die  $0_{0,70}$  für  $\hat{q}_{N,p}$  wird mit dem 0,70-Pezentil der nicht-zentralen t-Verteilung  $t_{0,70, n-1}(z_p \sqrt{n})$  für  $n - 1$  Freiheitsgrade und dem Nicht-Zentralitätsparameter  $z_p \sqrt{n}$  gebildet [9]:

$$0_{0,70}(\hat{q}_{N,p}) = \hat{\mu} + \frac{t_{0,70, n-1}(z_p \sqrt{n})}{\sqrt{n}} \hat{\sigma} \tag{A.7}$$

Auch hier lassen sich die Punktschätzer ineinander umrechnen  $0_{0,70}(\hat{q}_{LN,p}) = \exp\{0_{0,70}(\hat{q}_{N,p})\}$ .

Für Stichproben mit zensierten Werten ergibt sich die Standardabweichung  $s_p = \sqrt{s_\mu^2 + z_p^2 s_\sigma^2}$  für  $\hat{q}_{N,p}$  bei unabhängigen  $\hat{\mu}$  und  $\hat{\sigma}$  aus den Standardfehlern  $s_\mu, s_\sigma$  des numerischen Anpassungsverfahrens der ML-Methode (Abschnitt A.3). Damit ist

$$0_{0,70}(\hat{q}_{N,p}) = \hat{q}_{N,p} + t_{0,70, n-1} s_p \tag{A.8}$$

wobei  $t_{0,70, n-1}$  das 0,70-Pertzentil der zentralen t-Verteilung mit  $n - 1$  Freiheitsgraden ist. Der Schätzer in der Lognormalverteilung ergibt sich wie bei Gleichung (A.7).

### A.3 ML-Methode

Für unabhängige Messwerte ist die Likelihood-Funktion die Wahrscheinlichkeitsdichte der gesamten Stichprobe, also das Produkt der einzelnen Wahrscheinlichkeitsdichten.

$$L = \prod_{i=1}^n f(x_i) \tag{A.9}$$

Im Fall der Normalverteilung  $f(x) = f_N(x)$  hat  $L$  nur zwei Parameter:  $\mu$  und  $\sigma$ , die durch die Bedingungen  $\frac{\delta L}{\delta \hat{\mu}} = 0$  und  $\frac{\delta L}{\delta \hat{\sigma}} = 0$  zu den Gleichungen (1) führen.

Für Stichproben mit zensierten Werten unterhalb einer Bestimmungsgrenze  $x_j < BG_j$  ersetzt man die Wahrscheinlichkeitsdichte durch die kumulierte Verteilungsfunktion der Normalverteilung.

$$L_{BG} = \prod_{j=1}^{n_j} \int_{-\infty}^{BG_j} f(x_j) dx_j \prod_{i=1}^{n_i} f(x_i) = \prod_{j=1}^{n_j} F_N(BG_j) \prod_{i=1}^{n_i} f(x_i) \tag{A.10}$$

Die Stichprobe hat  $n = n_i + n_j$  Werte, wobei  $n_j$  Werte unterhalb von Bestimmungsgrenzen  $BG_j$  liegen, die für jedes  $j$  verschieden sein können. Auch  $L_{BG}$  hängt wieder nur von den beiden Parametern  $\mu$  und  $\sigma$  ab, die nun aber nicht mehr analytisch darstellbar sind, da  $F_N$  nur noch numerisch zugänglich ist. Ansonsten stellt das Maximieren von  $L_{BG}$  aber nur ein Anpassungsproblem in den beiden Parametern dar, das etwa mit der Newton-Raphson-Methode [15] gelöst werden kann.

### A.4 Efron-Test

Falls  $F_{LN}(c|\hat{\mu}, \hat{\sigma})$  die kumulierte Verteilungsfunktion einer Lognormalverteilung ist, deren Parameter  $\hat{\mu}, \hat{\sigma}$  aus einer ML-Schätzung stammen, ist die Teststatistik des Efron-Tests [4] gegeben durch

$$\hat{C} = \int_0^\infty F_{LN}(c|\hat{\mu}, \hat{\sigma}) d\hat{F}^0 = \int_0^\infty G(c|\hat{\mu}, \hat{\sigma}) \hat{f}^0(c) dc \tag{A.11}$$

Der nicht-parametrische Schätzer  $\hat{F}^0$  für die Verteilungsfunktion der Stichprobe wird analog zum Kaplan-Meier-Schätzer angegeben. Dabei wird das Wahrscheinlichkeitsgewicht  $1/n$  eines zensierten Wertes gleichmäßig auf alle kleineren zensierten und nicht-zensierten Werte verteilt:

$$\hat{F}^0 = \sum_k \hat{f}^0(\xi_k) \text{ mit}$$

$$\hat{f}^0(\xi_k(\delta_k = 1)) = \frac{1}{n} \prod_{i=k+1}^n \left(\frac{i}{i-1}\right)^{1-\delta_i} \tag{A.12}$$

Dabei sind die Stichprobenwerte  $\xi_i = \max\{c_i, BG_i\}$ . Mathematisch wird das durch die (zufällige) Reihe  $\delta_1, \delta_2, \dots, \delta_n$  angegeben mit

$$\delta_i = \begin{cases} 1, & \text{nicht-zensiert, } \xi_i = c_i \\ 0, & \text{zensiert, } \xi_i = BG_i \end{cases}$$

In Bild 2 ist rechts  $\hat{F}^0$  aus Gleichung (A.12) gegen die  $\xi_i$  aufgetragen. Die  $BG_i$  spielen für  $\hat{F}^0$  keine Rolle ( $\delta_k = 1$  in Gleichung (A.12)), nur die niedrigste Bestimmungsgrenze muss berücksich-

tigt werden, um einen selbstkonsistenten Schätzer zu erhalten. Das bedeutet, dass immer  $\delta_1 = 1$  gesetzt ist, und daher auch  $\xi_1 = BG_1$  in Bild 2 aufgetragen ist.

---

#### L I T E R A T U R

- [1] EN 482:2021. Workplace exposure – Procedures for the determination of the concentration of chemical agents – Basic performance requirements.
- [2] EN 689:2018. Workplace exposure – Measurement of exposure by inhalation to chemical agents – Strategy for testing compliance with occupational exposure limit values.
- [3] *D. M. Diez, C. D. Barr, M. Çetinkaya-Rundel*: OpenIntro Statistics.
- [4] *L. Sachs*: Angewandte Statistik. 9. Auflage. Berlin, Springer 1999.
- [5] *S. S. Shapiro, M. B. Wilk*: An analysis of variance test for normality (for complete samples), *Biometrika* 52 (1965) no. 3/4, p. 591-611.
- [6] *B. Efron*: The Two-Sample Problem with Censored Data. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4 (1967) p. 831-852.
- [7] *P. Hewett, G. H. Ganzer*: A comparison of several methods for analyzing censored data. *Ann Occup Hyg* 51 (2007) no. 7, p. 611-32.
- [8] *D. R. Helsel*: Statistics for Censored Environmental Data Using Minitab and R, 2. Auflage. John Wiley & Sons, Hoboken/USA 2012.
- [9] *K. Krishnamoorthy, T. Mathew*: Statistical Tolerance Regions – Theory, Applications and Computation. Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken/USA 2009.
- [10] Technische Regel für Gefahrstoffe: Ermitteln und Beurteilen der Gefährdungen bei Tätigkeiten mit Gefahrstoffen: Inhalative Exposition (TRGS 402). GMBI. (2023), Nr. 42, S. 898-920.
- [11] *F. Clerc, R. Vincent*: Assessment of Occupational Exposure to Chemicals by Air Sampling for Comparison With Limit Values: The Influence of Sampling Strategy. *The Annals of Occupational Hygiene* 58 (2014) no. 4, p. 437-449.
- [12] Technische Regel für Gefahrstoffe: Verfahrens- und stoffspezifische Kriterien (VSK) für die Ermittlung und Beurteilung der inhalativen Exposition (TRGS 420). GMBI. (2014) Nr. 48, S. 997-1002; zul. geänd. GMBI. (2020) Nr. 9-10, S. 199.
- [13] *G. Blom*: Statistical estimates and transformed beta-variables. John Wiley & Sons, New York/USA 1958.
- [14] DIN ISO 5479:2004: Statistische Auswertung von Daten – Tests auf Abweichung von der Normalverteilung.
- [15] *T. Strutz*: Data Fitting and Uncertainty – A practical introduction to weighted least squares and beyond, 2. Auflage, 2016, Springer Vieweg.

---

**Dr. rer. nat. Jörg Rissler**

Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung (IFA), Sankt Augustin.