ISSN 1013-4506 FS-2022-183-AKSIGMA

Fachverband für Strahlenschutz e. V.

Mitgliedsgesellschaft der
International Radiation
Protection Association
(IRPA)
für Deutschland
und die Schweiz

Publikationsreihe
FORTSCHRITTE
IM STRAHLENSCHUTZ
Publication Series
PROGRESS IN RADIATION
PROTECTION

A PRIMER IN METROLOGY AND STATSTICS

R. Michel
J. Birkhan

A Primer in Metrology and Statistics

Rolf Michel

Institute for Radioecology and Radiation Protection
Leibniz University Hannover, Germany

Jonny Birkhan

Institute for Nuclear Physics
Technical University of Darmstadt

Abstract

This primer is intended for students or newcomers in the fields of metrology and statistics. Though it lays emphasis on measurements of ionizing radiation, it offers a generally applicable methodology. It summarizes some basics of the underlying ideas and tries to build a bridge from the theoretical background to the practical application.

Statistics plays an essential role when interpreting measurements of ionizing radiation for the purpose of radiological protection. It is needed for the quantification of measurement uncertainties, it provides the tools to decide whether or not the measurement result exceeds the background, it allows to assess whether or not a measurement procedure fulfils sensitivity requirements, and, last not least, to decide whether or not a result conforms to requirements, e.g. from regulations. Due to recent developments in metrology a number of international standards and guides exist by which all these aspects can be treated in a consistent and internationally accepted manner.

The Joint Committee for Guides in Metrology (http://www.bipm.org/en/committees/jc/jcgm/) has issued a number of guides which can be downloaded free of charge under http://www.bipm.org/en/publications/guides/. There is the International Vocabulary in Metrology (JCGM 2012a), VIM for short, which defines the general terminology in order to have an unambiguous language for addressing metrological issues. The Guide to the Expression of Uncertainty in Measurement (JCGM 2008a), GUM for short, and its Supplement 1 (JCGM 2008b, 2009), GUM S1 for short, and Supplement 2 (JCGM 2011), GUM S2 for short, provide the basis for quantification of measurement uncertainties. Finally, the guide "Evaluation of measurement data – The role of measurement uncertainty in conformity assessment" (JCGM 2012b) gives guidance for conformity assessments. A methodology for assessing detectability is given in ISO 11929 (ISO 2019a, 2019b, 2019c, 2020).

Content

List of Figures

Foreword

Statistics plays an essential role when interpreting measurements of ionizing radiation for the purpose of radiological protection. It is needed for the quantification of measurement uncertainties, it provides the tools to decide whether or not the measurement result exceeds the background, it allows to assess whether or not a measurement procedure fulfills sensitivity requirements, and, last not least, to decide whether or not a result conforms to requirements, e.g. from regulations. Due to recent developments in metrology a number of international standards and guides exist by which all these aspects can be treated in a consistent and internationally accepted manner. The applications of this methodology extend far beyond measurements of ionizing radiation; it can be applied universally.

The Joint Committee for Guides in Metrology (http://www.bipm.org/en/committees/jc/jcgm/) has issued a number of guides which can be downloaded free of charge under http://www.bipm.org/en/publications/guides/. There is the International Vocabulary in Metrology (JCGM 2012a), **VIM** for short, which defines the general terminology in order to have an unambiguous language for addressing metrological issues. The Guide to the Expression of Uncertainty in Measurement (JCGM 2008a), **GUM** for short, and its Supplement 1 (JCGM 2008b), **GUM S1** for short, and Supplement 2 (JCGM 2011), **GUM S2** for short, provide the basis for quantification of measurement uncertainties. Finally, the guide "Evaluation of measurement data – The role of measurement uncertainty in conformity assessment" (JCGM 2012b) gives guidance for conformity assessments. A methodology for assessing detectability is given in ISO 11929 (ISO 2019 a, 2019b, 2019c, 2020).

The methodology provided by these guides and ISO 11929 has been taken into account also in a number of ISO standards dealing with environmental radiation measurements; see e.g. Calmet (2014) and Calmet et al. (2016) for surveys.

This primer extends previous papers which deal with the statistical and practical aspects of measuring, estimating, and deciding under uncertainty (Michel 2016, 2017, 2019). It is intended for students of natural sciences who are dealing with any kind of measurements. The purpose is to provide them with a theoretical basis enabling them to critically interpret the results of their measurements.

Acknowledgements

The author (R. Michel) is greatly indebted to Klaus Weise (retired professor from the Physikalisch Technische Bundesanstalt, PTB, at Braunschweig). His ideas and work provided the basis for a Bayesian theory of uncertainty in measurement and the development of ISO 11929. Klaus Weise served for more than 30 years in this working group and laid the foundation of ISO 11929.

There are more colleagues from PTB from which we benefitted as for instance Bernd Siebert and Rolf Behrens.

Since the end of 1981 a working group of the German Swiss Association for Radiation Protection (Deutsch-Schweizerischer Fachverband für Strahlenschutz, FS, www.fs-ev.org), called AK SIGMA, was developing standards for the determination of characteristic limits. To mention are the gone secretaries Prof. Heinrich Schultz (University Hannover) and Prof. Klaus Kirchhoff (Niedersächsisches Landesamt für Ökologie, today NLWKN). The author (R. Michel) participated in the work of this group since 1987 and became the secretary in 2000 after the death of Klaus Kirchhoff.

The authors thank the members of the working group AK SIGMA of the Swiss-German Radiation Protection Association (FS) for their cooperation and helpful discussions.

Klaus Weise was a member of this group from the beginning until 2019 and, practically, he was the brain of the group. Starting from a frequentist approach, the standard series DIN 25482 was developed which later became a corresponding series of ISO 11929. New developments in metrology, in particular a Bayesian theory of uncertainty in measurements by Klaus Weise and Wolfgang Wöger (1993, 1999), led the way to a new approach. Using Bayesian statistics, the last three parts of DIN 25482 and the corresponding ISO standards were formulated. In 2010, a new and one-part ISO 11929 (ISO 2010) was published basing the entire standard on Bayesian foundations. It came back to Germany as DIN ISO 11929: 2011 replacing the previous DIN 25482 series.

The routine revision of ISO 11929:2010 extended the applicability of ISO 11929 to cases where the classical GUM is no longer applicable and the GUM Supplement 1 or 2 have to be applied. It now consists of 4 parts, the first three of which were published in 2019 and the 4th part of DIN ISO 11929 is ready for the final voting in summer of 2020.

In particular, we want to thank Günter Kanisch and Dieter Schrammel for their continuous efforts to move ISO 11929 ahead. Günter Kanisch developed a software package, UncertRadio, which covers all aspects of the GUM and ISO 11929. Dieter Schrammel scrutinized all calculations using commercial software packages. Over the years, Dieter Schrammel, Günter Kanisch, and one of the authors (R.M.) offered training courses on ISO 11929 at the FTU of KIT, Karlsruhe.

The ISO 11929 series of standards was handled by ISO/TC85/SC2/WG17 “Radioactivity Measurements”. My thanks (R. Michel) are due to the continuous support by the WG17 members, in particular due to the Convenor Dominique Calmet.

The authors gave a number of lectures about the topics dealt with in this primer at the Institute for Radiation Protection and Radioecology of the Leibniz University Hannover, Germany and the Institute for Nuclear Physics of the Technical University of Darmstadt, Germany. These lectures and the discussions with the attending students helped very much to make the issues of this primer clearer. We hope that it provides some assistance for the education of students of natural sciences for the future.

Quantities and symbols

Quantities are denoted by capital letters. They have to be carefully distinguished from their (measured or true) values which are denoted by the respective lower case letters. In this document quantities and symbols are used according to ISO 11929.

In this standard, a quantity is considered to have a true value for a particular measurement problem which is unknown and unknowable. In some applications, one needs to assume a true value.

Y	physical quantity of interest, the **measurand.** The symbol Y is also used for a random variable serving as an estimator of the measurand. For measurements of ionizing radiation, the measurand is assumed to be non-negative.
G	**model of evaluation**, set of mathematical relationships connecting the input quantities to the measurand, e.g. $Y = G(X_i, i = 1,...,n)$
c_i	**sensitivity coefficient** $c_i = \left.\dfrac{\partial G}{\partial X_i}\right\vert_{X_1=x_1,...,X_m=x_m}$
W	**calibration factor**
w	estimate of the calibration factor
$u_{\mathrm{rel}}(w)$	**relative standard uncertainty** of a quantity W associated with the estimate w
Y_0	**random variable** serving as an estimator of Y without taking into account that Y is non-negative
X_i	physical quantities, **input quantities.** The symbols are also used for random variables serving as estimators of the quantities.
$\tilde{x}_i$	possible or assumed **(true) quantity values** of X_i , unknown and unknowable
x_i	**primary measurement results** of X_i, known and fixed
$\tilde{y}$	possible or assumed **(true) quantity value**, unknown and unknowable. It holds $\tilde{y} = G(\tilde{x}_i, i = 1,...,n)$.
y	**primary measurement result** $y = G(x_i, i = 1,...,n)$, known and fixed. The primary result of a measurement may be negative.
$u(y)$	**standard uncertainty** associated with the primary measurement result y
$\hat{y}$	**best estimate** of the true value of the measurand. The non-negativity of the measurand is taken into account in the calculation of the best estimate.
$u(\hat{y})$	**standard uncertainty** associated with the best estimate $\hat{y}$
a	**set of information**. It comprises the $x_i (i = 1,...,n)$ and their associated standard uncertainties $u(x_i)\,(i = 1,...,n)$ and potentially existing covariances[1], as well as

[1] For simplicity, the input quantities $X_i (i = 1,...,n)$ are supposed not to be correlated in this primer, so that covariances need not to be considered. For some advice on how to handle covariances, see ISO 11929-3:2019.

any other available information about the quantities involved. It includes also the values and associated uncertainties of the input quantities **not** taking, however, into account that the measurand is non-negative.

$\Im$ **information** available before a measurement is performed and taken into account for the calculation of the best estimate and its associated standard uncertainty

y^* **decision threshold**

$y^\#$ **detection limit**

y_r **guideline value**

$P(A|B)$ **conditional probability** for event A being true given B is true

$f(y|x)$ **probability density function** (PDF) for a value y given a value x

$f_Y(\tilde{y}|y, Y \geq 0)$ **posterior probability density function** (PDF) that the true value is $\tilde{y}$ given the estimate y taking into account the condition that the measurand Y is non-negative.

$f_{Y,0}(\tilde{y}|y)$ **posterior probability density function** (PDF) that the true value is $\tilde{y}$ given the estimate y **not** taking into account the condition that the measurand Y is non-negative.

$f(\tilde{y})$, $f(\tilde{y}|\Im)$ **prior probability density functions** that the true value is $\tilde{y}$ taking into account the information $\Im$ available before a measurement is performed.

$f_Y(y|\tilde{y})$ **predictive probability density function** (PDF) to obtain a measured value y if a true value $\tilde{y}$ of the measurand Y is assumed.

$f_Y(y|\tilde{y}=0)$ **predictive probability density function** (PDF) to obtain a measured value y if a true value $\tilde{y}=0$ of the measurand Y is assumed.

$f_Y(y|\tilde{y}=y^\#)$ **predictive probability density function** (PDF) to obtain a measured value y if a true value $\tilde{y}$ of the measurand Y equal to the detection limit $y^\#$ is assumed.

y_j values y from different measurements $(j = 0, 1, 2, ...)$

$\tilde{y}_i$ intermediate values for approximations of the detection limit $y^\#$

y_low, y_up **lower and upper limit of an unspecified coverage interval**, respectively, of the measurand

$y^\triangleleft$, $y^\triangleright$ lower and upper **limit of the probabilistically symmetric coverage interval**, respectively. The non-negativity of the measurand is taken into account.

$y^<$, $y^>$ lower and upper **limit of the shortest coverage interval**, respectively. The non-negativity of the measurand is taken into account.

ω, κ auxiliary quantities

n_g, n_0 **number of counted pulses** of the gross effect (index g) and of the background effect (index 0), respectively

n_i	**number of counted pulses** obtained from the measurement of the count rate R_i during a certain measurement duration
t_g, t_0	**measurement duration** of the measurement of the gross effect (index g) and of the background effect (index 0), respectively
t_i	**measurement duration** of the measurement of the count rate R_i
R_g, R_0	**gross count rate (index g) and background count rate (index 0)**, respectively
r_g, r_0	**estimate of the gross count rate (index g) and of the background count rate (index 0)**, respectively
τ_g, τ_0	**relaxation time constant of a ratemeter** used for the measurement of the gross effect (index g) and of the background effect (index 0), respectively
α, β	**probability of a false positive and false negative decision**, respectively
$1-\gamma$	**probability for the coverage interval** of the measurand
k_p, k_q	**quantiles of the standard normal distribution** for the probabilities p and q, respectively (for instance $p = 1-\alpha$, $1-\beta$ or $1-\gamma/2$)
$\Phi(t)$	**distribution function of the standard normal distribution**; $\Phi(k_p) = p$ applies.

1 Introduction

1.1 General considerations about human knowledge and physical truth

The problem of human knowledge is still today nicely explained by Plato's Allegory of the Cave (Platon, 427 BC – 347 BC, Politeia 7th book, about 370 BC). Plato describes some people, who are tied up in a sub-surface cave over their entire life in a way, that they cannot move their heads or their bodies. Therefore, they can look only at the wall of the cave in front of them. They have only light from a fire burning behind them. Between the fire and their backs figures and objects are carried around, which cast their shadows at the wall. The „prisoners" can see only these shadows, their own ones and those of their co-prisoners. If the carriers speak, the echo comes back from the wall as if the shadows were speaking themselves. Since the world of the prisoners is all about these shadows, they take them for the real world.

So in the world of the prisoners the shadows are the objects present which science has to deal with. About 2000 years later Descartes specifies how such a scientific endeavour should be performed: „*It must be the final goal of scientific endeavour to guide the intellect in a way that it gives reasonable and true judgements about the objects present*.“ Descartes (1596 – 1650) Regulae ad Directionem Ingenii, Regula 1.

However, Descartes' reasoning – except for a proof of the existence of god – just resulted in the only true statement given the uncertainty of human knowledge about the "real world": „*Cogito ergo sum*.“ Descartes (1596 – 1650) Discours de la méthode pour bien conduire sa raison et chercher la vérité dans les sciences.

However, what are the "objects present"? The sentence „*Cogito ergo sum*“ does not give an answer to this question. So one ends up with the answer of a poet: "… *and a dream is all life and the dreams themselves are dream* …" Calderon de la Barca (1635) La vida es sueño.

However, since physicists are pragmatic people, the only meaningful advice can be: Don't care about what is real and what is dream! All our observations are the "objects present"; i.e. the phenomena of our world. This is a positivistic approach. Positivism is a philosophical theory stating that knowledge is based on natural phenomena and their properties and relations. Data received from the senses provide empirical evidence. Information derived from experience, interpreted through reason and logic, forms the exclusive source of all knowledge. The quoted "*interpreted through reason and logic*" means the use of a scientific method to obtain knowledge.

The scientific method can be defined as a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge. To be termed scientific, a method of inquiry is commonly based on empirical or measurable evidence subject to specific principles of reasoning ("Rules for the study of natural philosophy", Newton transl. 1999, pp. 794–96, after Book 3, The System of the World). The Oxford Dictionaries Online (Oxford Dictionaries: British and World English, 2016, obtained 28 May 2016) define the scientific method as "*a method or procedure that has characterized natural science since the 17th century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses*". Experiments need to be designed to test hypotheses. The most important part of the scientific method is the experiment.

Thus, physical science is a systematic enterprise that builds and organizes knowledge in the form of testable explanations and predictions about the universe; i.e. the "objects present".

It tries to explain and predict nature's phenomena, based on empirical evidence. Hypotheses must be verified scientifically to be regarded as scientific theory.

To describe the relations between the objects present one needs reason and logic and causality: *nihil fit sine causa*. Gottfried Wilhelm Leibniz formulated the Principle of Sufficient Reason (Satz vom zureichenden Grund): „*Im Sinne des zureichenden Grundes finden wir, dass keine Tatsache als wahr oder existierend gelten kann und keine Aussage als richtig, ohne dass es einen zureichenden Grund dafür gibt, dass es so und nicht anders ist, obwohl uns diese Gründe meistens nicht bekannt sein mögen*." G.W. Leibniz: Monadologie, § 32; cited from German. French Suhrkamp edition 1998, p. 27.

From causality and the Principle of Sufficient Reason it is just a short step to speak of scientific determinism. Causal determinism, which in physics is known as cause-and-effect, is the concept that events within a given paradigm are bound by causality in such a way that any state (of an object or event) is completely determined by prior states. In the history of science, Laplace' demon was the first published articulation of causal or scientific determinism by Pierre-Simon Laplace in 1814. According to determinism, if someone (the Demon) knows the precise location and momentum of every atom in the universe, their past and future values for any given time are entailed; they can be calculated from the laws of classical mechanics.

"We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. Une intelligence... Rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux." Pierre Simon Laplace (1814) A Philosophical Essay on Probabilities.

We do not need to start here a discussion whether Laplace' sentence is right or wrong, because, unfortunately, we do not have such an intelligence as his demon. And even worse, the question is whether or not the world is deterministic or in other words whether or not God plays dice. Regarding this latter question we quote Albert Einstein "*God does not play dice with the universe*." and Stephen Hawking: "*So God does play dice with the universe. All the evidence points to him being an inveterate gambler, who throws the dice on every possible occasion"* Stephen Hawking (https://www.hawking.org.uk/in-words/lectures/does-god-play-dice, obtained 2021.03.21)

But it does not matter whether we have not enough intelligence to understand nature or whether nature is ruled by chance. In any case, there is uncertainty due to lack of information and so we end up saying that human knowledge is limited. But, if everything – even the world – is uncertain, one has to get involved with probabilities. The success decides about the truth. In order to live we have to infer and decide on the basis of common sense on the basis of incomplete information. And we conclude:

Uncertainty originates from lack of information, regardless of whether this lack is due to nature itself or simply due to ignorance.

1.2 Radioactive decay as a stochastic process

A major blow against determinism came from the observation of the radioactive decay and the development of quantum mechanics. We shall only shortly discuss these aspects here for the case of the radioactive decay.

The spinthariscope by Crookes (1903) allowed seeing the unpredictable flashes and their random occurrences. The Law of the Radioactive Decay was formulated with a decay probability λ, N being the number of radioactive nuclei

$$A = -\frac{\mathrm{d}N}{\mathrm{d}t} = \lambda \cdot N \quad A(t) = A(0) \cdot \exp(-\lambda t) \tag{1.1}$$

If the decay probability is small, then the number of nuclei N will practically not change during the time of observation. Then, the probability $p(n)$ to observe per unit time exactly n decays is given by a Poisson distribution $\mathrm{Po}(n|\lambda, N)$. If the decay probability is not small compared to the time of observation, the process of radioactive decay can be described by a Binomial distribution; see below in this chapter.

The Poisson distribution is a discrete probability distribution that expresses the probability of a given number (n = 0, 1, 2, …) of events occurring in a fixed interval of time t_m if these events occur with a constant rate, here $r = \lambda \cdot N$. The Poisson distribution applies in cases where a small probability acts on a large number of items, as e.g. in the decay of a long-lived radionuclide.

The Poisson distribution (Fig. 1) has only one parameter $\mu > 0$ and is given by

$$\mathrm{Po}(n|\mu) = \frac{\mu^n \cdot \mathrm{e}^{-\mu}}{n!} \tag{1.2}$$

For the radioactive decay this means

$$\mathrm{Po}(n|\lambda, N, t_\mathrm{m}) = \frac{(\lambda \cdot N \cdot t_\mathrm{m})^n \cdot \mathrm{e}^{-\lambda \cdot N \cdot t_\mathrm{m}}}{n!} \tag{1.3}$$

It holds

$$\mathrm{Var}(n) = \mathrm{E}(n) = \lambda \cdot N \cdot t_\mathrm{m} \tag{1.4}$$

The parameter $\lambda \cdot N$ of the Poisson distribution can be estimated from the mean value $< n >$ of the number of decays observed per unit time in experiments by $\lambda \cdot N \cdot t_\mathrm{m} =< n >$.

$$\mathrm{Po}(n|\lambda, N, t_\mathrm{m}) = \frac{< n >^n \cdot \mathrm{e}^{-<n>}}{n!} \tag{1.5}$$

For large $\lambda \cdot N \cdot t_\mathrm{m}$ the Poisson distribution becomes a Gaussian or normal distribution which has two parameters, $\mu \in \Re$ and $\sigma > 0$ (see Fig 1.). It is given by

$$\mathrm{N}(x\,|\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{1.6}$$

For the radioactive decay this reads

$$\mathrm{N}(n\,|\,\lambda\cdot N\cdot t_\mathrm{m}) = \frac{1}{\sqrt{2\pi\cdot\lambda\cdot N\cdot t_\mathrm{m}}}\exp\left[-\frac{(x-\lambda\cdot N\cdot t_\mathrm{m})^2}{2\lambda\cdot N\cdot t_\mathrm{m}}\right], \text{ i.e. } \sigma = \sqrt{\mathrm{Var}(n)}\,. \tag{1.7}$$

For this Gaussian distribution holds also $\mathrm{Var}(n) = \mathrm{E}(n)$.

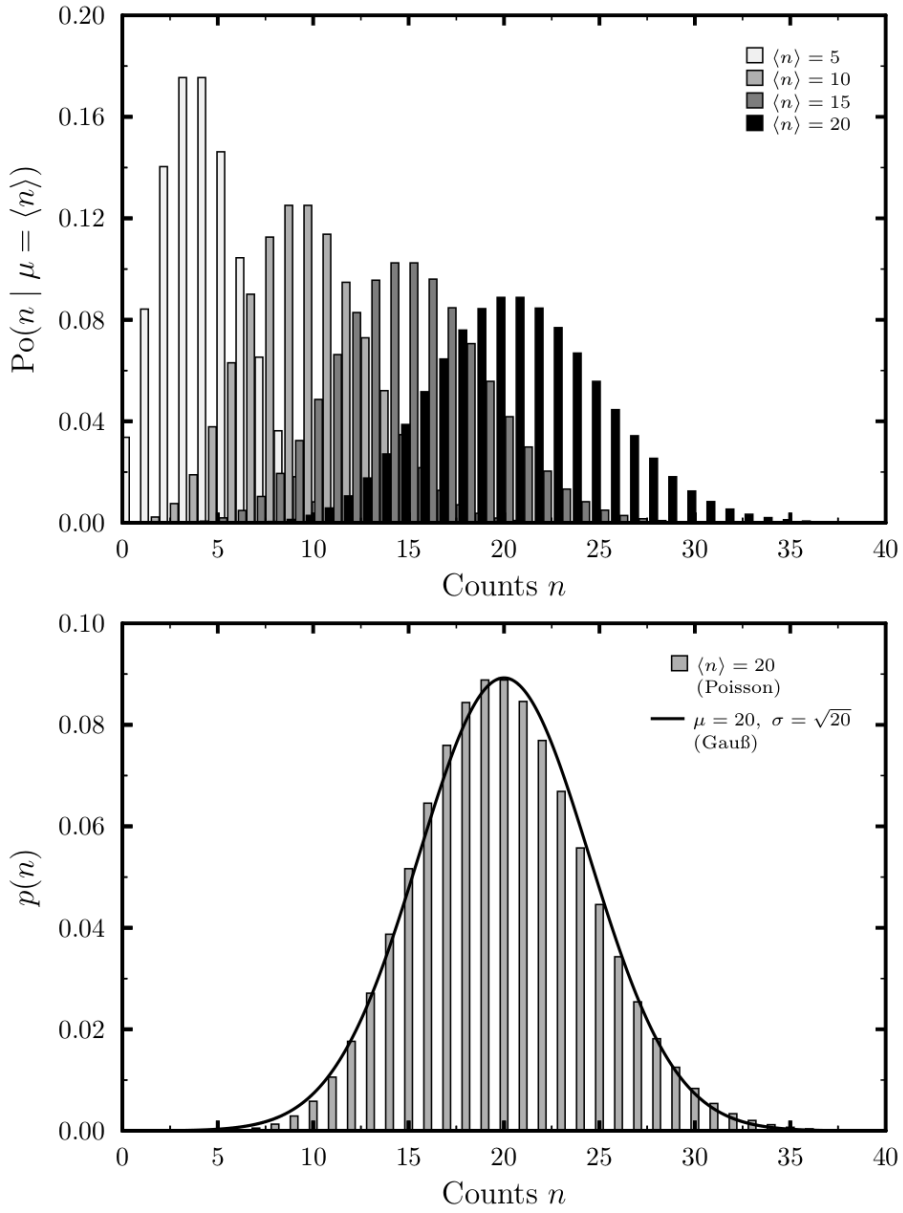


Fig. 1: Examples of Poisson distributions (top) and comparison of a Poisson with a normal distribution (bottom).

If the decay probability is not small, then the number of nuclei N will change during the time of observation. Then, the probability to observe per unit time exactly n decays is given by a Binomial distribution $\mathrm{Bi}(n\,|\,\lambda, N)$ (Fig. 2).

The Binomial distribution applies to so-called Bernoulli trials. Independent repeated trials of an experiment with exactly two possible outcomes are called Bernoulli trials. Call one of the outcomes "success" and the other outcome "failure". Let p be the probability of success in a Bernoulli trial, and q be the probability of failure, then a Binomial distribution describes the probability $\mathrm{Bi}(k|\,p,n)$, to have success exactly k times in n Bernoulli trials. With $q = 1 - p$ the Binomial distribution reads

$$\mathrm{Bi}(k|\,p,n) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \tag{1.8}$$

$$\mathrm{E}(n) = n \cdot p \quad \mathrm{Var}(n) = n \cdot p \cdot (1-p) \tag{1.9}$$

Applying this to the radioactive decay means

$$\mathrm{Bi}(n|\,\lambda,N,t_\mathrm{m}) = \binom{N}{n} (\lambda \cdot t_\mathrm{m})^n (1-\lambda \cdot t_\mathrm{m})^{N-n} \text{ , i.e. } \lambda \text{ is something like } \lambda = \frac{\mathrm{d}\rho}{\mathrm{d}t} \tag{1.10}$$

with

$$\mathrm{E}(n) = \lambda \cdot N \cdot t_\mathrm{m} \quad \mathrm{Var}(n) = \lambda \cdot t_\mathrm{m} \cdot (1-\lambda \cdot t_\mathrm{m}) \cdot N \tag{1.11}$$

The functions $\mathrm{Po}(n|\,\lambda,N,t_\mathrm{m})$, $\mathrm{N}(n|\,\lambda \cdot N \cdot t_\mathrm{m})$ and $\mathrm{Bi}(n|\,\lambda,N,t_\mathrm{m})$ are probability density functions (PDFs) describing the probability distribution of outcomes of counting experiments as a consequence of the fact that the radioactive decay is a stochastic process. The counting efficiency is assumed to be unity in this simple example. These PDFs are conditional probabilities to obtain n counts in the measurement time t_m given the decay probability λ and the number N of radioactive nuclei.

For larger number of counts, both the Poisson and the Binomial distribution can be well approximated by a normal distribution (Fig. 3) if the product of the decay probability and the measurement time is markedly less than 1.

What does all that then mean for uncertainty and probability? To answer this question, we shall look into these two terms, uncertainty and probability, more generally.

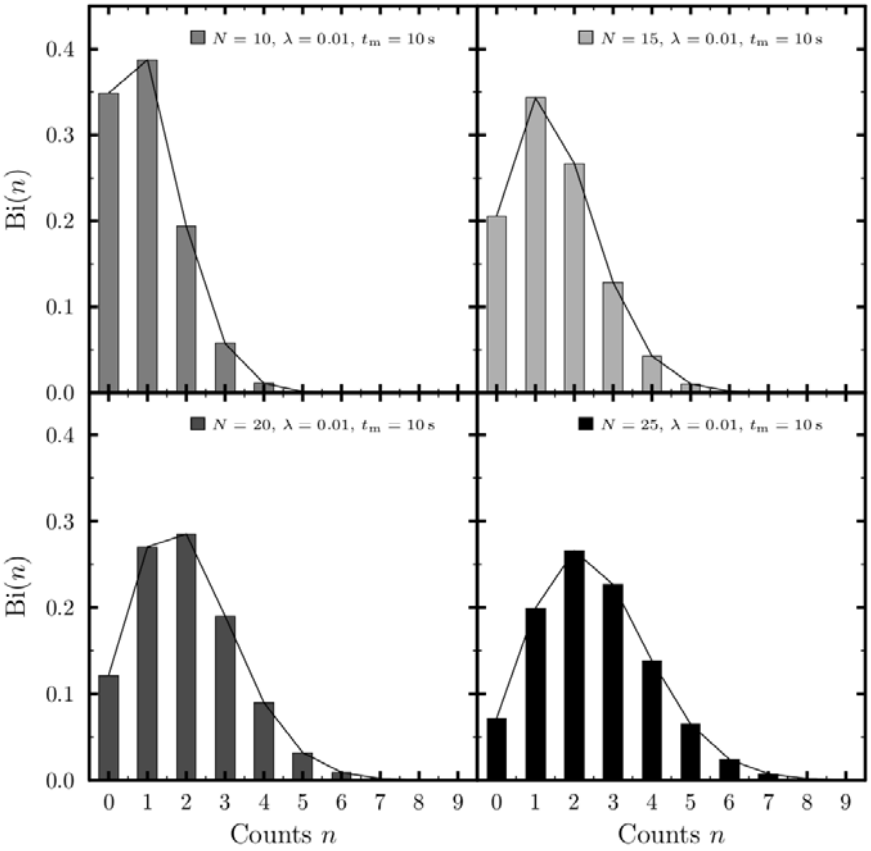


Fig. 2: Examples of Binomial distributions.

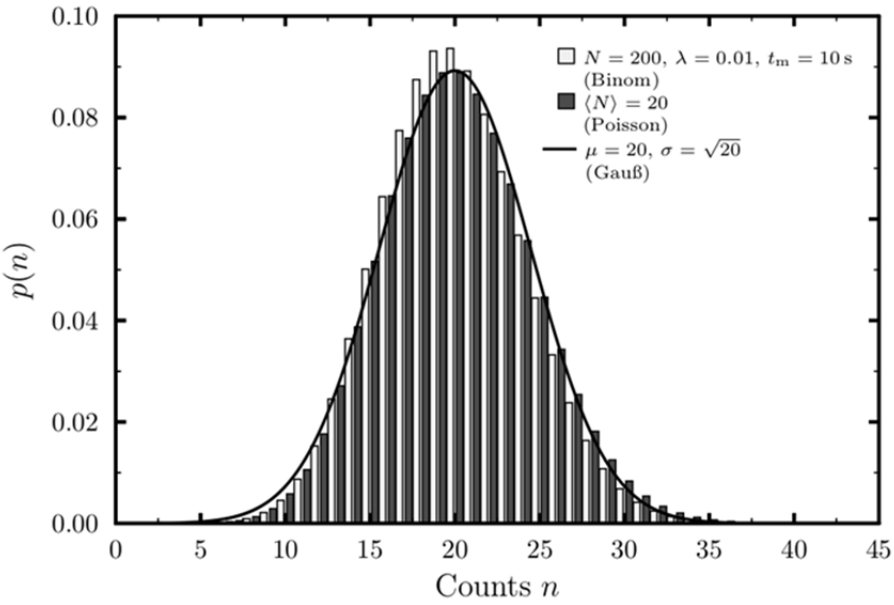


Fig. 3: Comparison of Poisson, Binomial and normal distributions.

1.3 Uncertainty and probability

Uncertainty is a general characteristic of human existence. It originates from ignorance. The ignorance, that one does not know what the dice will show, whether it will be a girl or a boy, whether a wing beat of a butterfly is responsible for today's weather, when a nucleus will de-

cay, what the future will be, what the truth is, which quantities influence the results of an experiment, whether there is causal connection between two quantities, whether a system is deterministic, stochastic or chaotic, or whether chance is ruling the world.

Uncertainties are important characteristics of human reasoning, decision making and action and, in the end, they are a consequence of limited and incomplete information. Humans always have to decide and to act under uncertainty, i.e. on the basis of incomplete information.

In the case of ignorance one can only rely on probabilities. Uncertainty can be quantified by probabilities. James Clerk Maxwell said in this context „*The true logic for this world is the calculus of probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind*" (Maxwell 1850). Success decides about the truth.

A complete description of the uncertainty can be obtained by deriving a probability density function (PDF) over the space of possibilities. Probability theory and probability calculus provide the tools to establish and propagate probability density functions. Just a few principles are sufficient for a given problem to derive from the available information the desired PDF. Fundamental is the Principle of Indifference, also called Principle of Insufficient Reason (Laplace 1812). Given $n > 1$ distinguishable, mutually exclusive and collectively exhaustive events, the Principle of Indifference states that without further information each event should be assigned a probability equal to $1/n$. The Principle of Indifference is closely related with the Principle of Maximum Entropy (PME) (Jaynes 1982) and the Bayes Theorem (Bayes 1763).

In metrology, these principles are used to quantify uncertainty in measurement. In the 1990ties the quantification of measurement uncertainties was standardized by the ISO Guide to the Expression of Uncertainty in Measurement (GUM) (ISO 1993) and the standard series DIN 1319 (DIN 1996, 1999). A Bayesian theory of measurement uncertainties (Weise and Wöger 1993) provided a theoretical basis for the GUM. This basis makes use of the Bayes Theorem, the Principle of Maximum Information-Entropy and the Product Rule to establish and propagate the desired PDFs (see chapter 4). After initial problems of acceptance and manifold discussions the GUM was newly published (JCGM 2008a) and extended (JCGM 2008b, 2011) by the Joint Committee on Guides in Metrology and today represents the internationally accepted methodology for the quantification of measurement uncertainties.

Uncertainty manifests itself in metrology as follows. By a measurement one obtains an uncertain estimate y (measurement result) of the unknown and unknowable true quantity value $\tilde{y}$ of a measurand Y. So, the conditional probability $f_Y(\tilde{y}|y)$, i.e. the probability that the true value of the measurand Y is $\tilde{y}$ given the measurement result y (Fig. 4), provides the complete description of the uncertainty associated with the measurement result y. $f_Y(\tilde{y}|y)$ is the PDF of a random variable serving as an estimator of Y; as a PDF it is normalized $\int_{-\infty}^{\infty} f_Y(\tilde{y}|y)\,\mathrm{d}\tilde{y} = 1$.

If further information $\mathfrak{I}$ is available before a measurement is performed, it is taken into account in the PDF $f_Y(\tilde{y}|y,\mathfrak{I})$ instead of $f_Y(\tilde{y}|y)$ which then completely describes the uncertainty. The uncertainty can also be described by a coverage interval $[y^{\triangleleft}, y^{\triangleright}]$ which contains the true value of the measurand with a preselected probability $(1 - \gamma)$ or by the best estimate

and its associated standard uncertainty. With the PDF $f_Y(\tilde{y}|y)$ the best estimate $\hat{y}$ of the true value $\tilde{y}$ is the expectation

$$\hat{y} = \mathrm{E}(f_Y(\tilde{y}|y,\Im)) = \int_{-\infty}^{+\infty} \tilde{y} \cdot f_Y(\tilde{y}|y,\Im)\,\mathrm{d}\tilde{y} \tag{1.12}$$

and the variance

$$u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}|y,\Im)) = \int_{-\infty}^{+\infty} (\tilde{y}-\hat{y})^2 \cdot f_Y(\tilde{y}|y,\Im)\,\mathrm{d}\tilde{y} \tag{1.13}$$

gives the standard uncertainty $u(\hat{y})$ associated with the best estimate.

The PDF depends on the available information. The GUM - not so clearly - and the GUM Supplement 1 (GUM S1) explicitly make use of the PME and the Bayes Theorem to derive various PDFs depending on the available information. If there is more information available than y only, e.g. any other available prior information $\Im$, the PDF completely describing the desired probability is $f_Y(\tilde{y}|y,\Im) = C \cdot f_Y(\tilde{y}|y) \cdot f_Y(\tilde{y}|\Im)$ (compare chapter 5). Then, the best estimate of the true value $\tilde{y}$ is $\hat{y} = \mathrm{E}(f_Y(\tilde{y}|y,\Im))$ and its associated squared standard uncertainty $u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}|y,\Im))$.

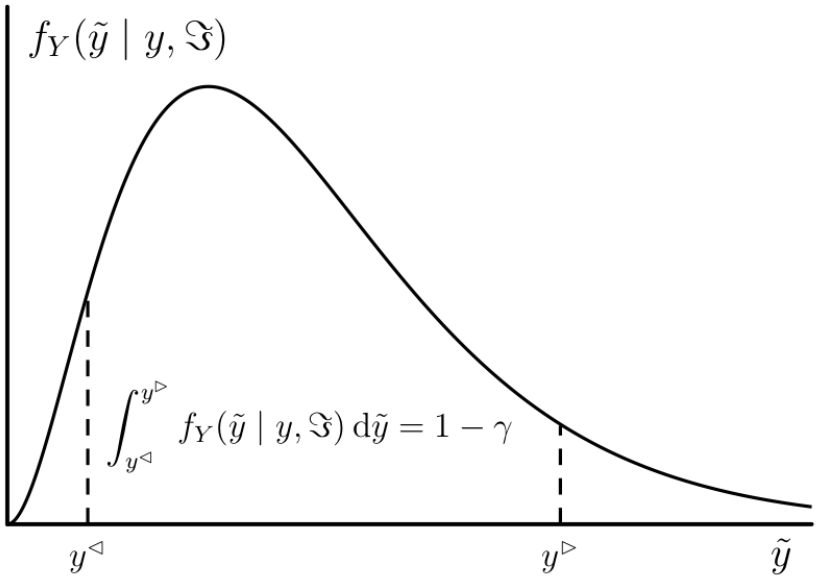


Fig. 4: Schematic of a PDF $f_Y(\tilde{y}|y,\Im)$ and of the limits of a coverage interval $[y^\triangleleft, y^\triangleright]$.

2 Statistical basics

2.1 General aspects

According to a dictionary (http://www.dictionary.com/browse/statistics), statistics is the science that deals with the collection, classification, analysis, and interpretation of numerical

facts or data, and that, by use of mathematical theories of probability, imposes order and regularity on aggregates of more or less disparate elements.

The following fields of statistics have to be distinguished:

- descriptive statistics (also empirical statistics),
- explorative statistics (also hypothesis-generating statistics, data-mining),
- inductive statistics (also mathematical statistics, inferring statistics).

In **inductive statistics** one infers from a random sample (data) the characteristics (value) of a basic population (quantity). Such inference needs plausibility which – as explained in the next chapter – can be quantified by probability. Probability theory provides the basis for the methods of testing, estimating, inferring and deciding. Inductive statistics is the topic with which we deal in this primer.

2.2 Plausibility and probability

There are two types of inference (German: Schließen) to be distinguished:

- **deductive reasoning** (German: deduktives logisches Denken) whenever enough information is at hand to permit it;
- **inductive or plausible reasoning** (German: induktives oder plausibles logisches Denken) when – as is almost invariably the case in real problems – the necessary information is not available.

The quote partially already mentioned above „*The actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have a reason on. Therefore the true logic for this world is the calculation of probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind*" (James Clerk Maxwell 1850) clarifies exactly the distinction between deductive reasoning, i.e. *the science of logic …*, and inductive reasoning, i.e. *the true logic for this world …*

Another quote dealing with this distinction is from E.T. Jaynes (1982) in the reprint of 2002 edited by G. Larry Bretthorst**: "***If a problem can be solved by deductive reasoning, probability theory is not needed for it; thus our topic is the optimal processing of incomplete information.*"

The contrast between deductive reasoning and plausible reasoning can be explained according to Jaynes (2003) by the following examples. As is generally credited to the *Organon* of Aristotle (4^{th} century b.c.) **deductive reasoning** (*apodeixis: Beweis*) can be analyzed ultimately into the repeated application of two strong syllogisms (logische Schlüsse)

1. If A is true, then B is true. Given A is true; therefore, B is true: $A \Rightarrow B$.

and its inverse

2. If A is true, then B is true, given B is false, therefore A is false: $\neg A \Leftarrow \neg B$.

These are the tools of deductive reasoning which allow assuring with certainty that a proposition is either true or false.

In **plausible reasoning** (*epagogue***:** *induction*) weaker syllogisms are used

1. If A is true, then B is true, Given B is true; therefore, A becomes more plausible.

   The evidence does not prove that A is true, but verification of one of its consequences does give us more confidence in A.

Another weak syllogism, still using the same major premise, is

2. If A is true, then B is true. Given A is false; therefore, B becomes less plausible.

A still weaker syllogism is

3. If A is true, then B becomes more plausible. Given B is true; therefore, A becomes more plausible.

As a consequence of allowing for plausibility, one needs a quantitative measure of plausibility. What are the requirements for such a measure? The basic requirements were formulated by R.T. Cox (1946) who wanted his system of plausibility to satisfy the following conditions

I. Divisibility and comparability – The plausibility of a proposition is a real number and is dependent on information we have related to the proposition.

II. Common sense – Plausibilities should vary sensibly with the assessment of plausibilities in the model, i.e. qualitative correspondence with common sense

III. Consistency – If the plausibility of a proposition can be derived in many ways, all the results must be equal.

See also Jaynes (2003) for more explanations of the "*basic desiderata for plausibility*" (desideratum = das Erwünschte).

2.3 The quantitative rules: from plausibility to probability

Probability theory is nothing but common sense reduced to calculation. Laplace, 1819

The following implications of Cox' postulates lead the way from plausibility to probability. Let $A|B$ be the plausibility of the proposition A given B satisfying Cox' postulates. Then, from the laws of probability follows that there exists a function w mapping the plausibility to the interval $[0,1]$ and a positive number m and that the following three statements hold; see also E.T. Jaynes (2003) for further details.

1. Certainty is represented by $w(A|B)=1$
2. $w^m(A|B)+w^m(\neg A|B)=1$
3. $w(A,B|C)=w(A|C)\cdot w(B|A,C)=w(B|C)\cdot w(A|B,C)$

These postulates imply only these general properties. One may recover the usual laws of probability by setting a new function, conventionally denoted P or Pr, equal to w^m. Then one obtains the laws of probability in a more familiar form

1. **Certain truth** is represented by $P(A|B)=1$, certain falsehood by $P(A|B)=0$.

2. **rule for negation**

$$P(A|B)+P(\overline{A}|B)=1 \tag{2.1}$$

3. **product rule**, rule for **conjunction** (logical „and“: $AB=A, B=A\wedge B$)

$$P(AB|C)=P(A|C)\cdot P(B|AC)=P(B|C)\cdot P(A|BC) \tag{2.2}$$

4. **sum rule**, **rule for disjunction** (logical „or“: $A+B=A\vee B$)

$$P(A+B|C)=P(A|C)\cdot P(B|C)-P(AB|C) \tag{2.3}$$

Since any proposition containing conjunction, disjunction, and negation can be equivalently formulated using conjunction and negation alone (the so-called conjunctive normal form), one can handle any compound proposition with the above 4 rules. “*It is (ought to be) possible, by repeated applications of the product rule and sum rule, to arrive at the plausibility of any proposition in a Boolean algebra*” (Jaynes 1983).

As soon as we recognize this, it is clear that, instead of saying that $p(x)=P(A_i|B)$ is an arbitrary monotonic function of *x*, it is much more to the point to turn this around and say that

***The plausibility x ≡ A|B is an arbitrary monotonic function of p, defined in* (0 ≤ *p* ≤ 1).**

The original mathematical *definition* of probability was given by James Bernoulli (1713) and was used by most writers for the next 150 years as given for example in Laplaces *Théorie Analytique des Probabilités* (1812)

The Probability for an event is the ratio of the number of cases favorable to it, to the number of all cases possible when nothing leads us to expect that any one of these cases should occur more than any other, which renders them, for us, equally possible.

Now we can introduce the notation to be used here further. As a formal probability symbol the capital *P* will be used

$$P(A|B)$$

which signifies that the arguments are *propositions*. Probabilities whose arguments are numerical values are generally denoted by other functional symbols such as

$$f(r|n,p) \tag{2.4}$$

which denote ordinary mathematical functions.

The Kolmogorov axioms presented an approach to probability theory phrased in the language of set theory and measure theory (Kolmogorov, 1933).

1. First axiom: the probability $P(E)$ of an event E from a total event space Ω is a non-negative real number

$$P(E) \geq 0 \quad \forall E \in \Omega \tag{2.5}$$

2. Second axiom: the probability that at least one of the elementary events in the entire sample space will occur is 1.

$$P(\Omega) = 1 \tag{2.6}$$

3. Third axiom (σ-additivity): for any countable sequence of disjoint sets (mutually exclusive events $E_1, E_2, \ldots$) holds

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i) \tag{2.7}$$

As pointed out by Jaynes (2003), the system of probability used here, differs conceptually from that of Kolmogorov in that we do not interpret propositions in terms of sets, but we do interpret probability distributions as carriers of incomplete information. Partly as a result, our system has analytical resources not present at all in the Kolmogorov system. This enables us to formulate and solve many problems – particularly the so-called ill posed problems and generalized inverse problems – that would be considered outside the scope of probability theory according to the Kolmogorov system.

2.4 Two schools of statistics

Given the theoretical basis for plausibility and probability described in the preceding chapter, we have to deal with a problem which occurs when using the concept of probability in statistics. There are two basic schools in statistics; that of **Bayesian statistics** and that of **conventional or frequentist statistics**. The two schools are contradictive since the term probability does not have the same meaning, though many but not all results obtained by the two statistics are practically equal. The conventional or frequentist view is that probability is the stochastic limit of relative frequencies. The Bayesian view is that probability is a measure of the degree of belief an individual has in an uncertain proposition. Due to this fundamental difference the two statistics must not be confused with each other.

Bayesians follow the principle that the mathematical theory of probability is applicable to the degree to which a person believes a proposition. The Bayes Theorem can be used as the basis for a rule for updating beliefs in the light of new information. Such updating is known as Bayesian inference. In his "Essay towards solving a problem in the doctrine of chances", Thomas Bayes (* 1702, † 1761) invented the "Bayesian inference", i.e. calculating the probability of the validity of a proposition on the basis of a prior estimate of its probability and new relevant evidence (Bayes 1763). Bayesian inference is the natural way of human learning: incorporating new experience into the available set of prior assumptions. This is also applied in a Bayesian theory of measurement uncertainties (Weise and Wöger 1993). For a general introduction to Bayesian statistics see e.g. references (Lee 1989, Bernardo and Smith 1994, Robert 2001, Bernardo 2003, Gelman et al. 2003, Gregory 2005).

Probability theory and probability calculus provide the tools to establish and propagate probabilities. Just a few principles are sufficient for a given problem to derive from the available information the desired probability density function (PDF). Fundamental is the Principle of Indifference, also called Principle of Insufficient Reason (Laplace 1812). Given $n > 1$ distinguishable, mutually exclusive and collectively exhaustive events, the Principle of Indifference states that without further information each event should be assigned a probability equal to 1/n. The Principle of Indifference is closely related with the Principle of Maximum Entropy (PME) Jaynes 1982) and the Bayes Theorem (Bayes 1763).

In metrology, these principles are used to quantify uncertainty in measurement. In the 1990ties the quantification of measurement uncertainties was standardized by the ISO Guide to the Expression of Uncertainty in Measurement (GUM) (ISO 1993). After a decade-long controversy about the statistical foundation of the GUM, the GUM S1 made the clear statement that the GUM can only work on the basis of Bayesian statistics as formulated in a Bayesian theory of measurement uncertainties (Weise and Wöger 1993). This basis makes use of the Bayes Theorem, the Product Rule, and the Principle of Maximum (Information-) Entropy (PME) to establish and propagate the desired PDFs. After initial problems of acceptance and manifold discussions the GUM was newly published (JCGM 2008a) and extended (JCGM 2008b, 2011, 2012) by the Joint Committee on Guides in Metrology and today represents the internationally accepted methodology for the quantification of measurement uncertainties.

There is a persistent problem, namely that people using the GUM are still living in two different worlds: the worlds of Bayesian statistics and of conventional or frequentist statistics. Though many results obtained by the two statistics are practically equal, the statistics themselves must not be confused with each other. The term probability does not have the same meaning in the two worlds of statistics. The conventional or frequentist view is that probability is the stochastic limit of relative frequencies. The Bayesian view is that probability is a measure of the degree of belief an individual has in an uncertain proposition. This meaning of probability in Bayesian statistics is the same as in the statement that “the probability to get a six, when tossing a 6-sided dice, is 1/6”. If asked “What is the probability of tossing a six?”, a frequentist would answer “I do not know; I did not yet toss the dice.”

The GUM Suppl. 1 (JCGM 2008b) makes the clear statement that the GUM can only work on the basis of Bayesian statistics. Frequentist statistics cannot take into account type B uncertainties. For the definition if type B uncertainties see chapter 5.1. Further it only allows establishing the conditional probability $f_Y(y|\tilde{y})$ but not $f_Y(\tilde{y}|y)$.

2.5 Probability: an ambiguous term in statistics

The confusion about probability in the two types of statistics leads to a lot of misunderstandings and misinterpretations. Let be μ an estimate of the true value of a quantity X with

$$\mu = \overline{x} \pm \frac{\sigma}{\sqrt{n}} \tag{2.8}$$

Then the statement of conventional or frequentist statistics

$$P(\mu - \frac{\sigma}{\sqrt{n}} \le \bar{x} \le \mu + \frac{\sigma}{\sqrt{n}}) = 0,68 \tag{2.9}$$

has a completely different meaning than that of Bayesian statistics

$$P(\bar{x} - \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + \frac{\sigma}{\sqrt{n}}) = 0,68 \tag{2.10}$$

In frequentist statistics, Equation (2.9) makes a statement about the probability to obtain an average $\bar{x}$ of data in experiments given the fixed parameters μ and σ, while the Bayesian statistics statement (2.10) quantifies the probability[2] that the parameter μ lies within an interval $\left[\bar{x} - \sigma / \sqrt{n}, \bar{x} + \sigma / \sqrt{n}\right]$ given the fixed value of the data $\bar{x}$. D'Agostini (2003) in his book "Bayesian Reasoning in Data Analysis" gave a very clear description of the problem: "*However, conventional statistics says only that $\bar{X}$ is a probabilistic statement about $\bar{x}$, given μ, σ and n. Probabilistic statements concerning μ are not foreseen by the theory (μ is a constant of unknown value), although this is what we are, intuitively, looking for: Having observed the effect x we are interested in stating something about the possible true value responsible for it. In fact, when we do an experiment, we want to increase our knowledge about μ and, consciously or not, we want to know which values are more or less believable. A statement concerning the probability that an observed value falls within a certain interval around μ is meaningless if it cannot be turned into an expression which states the quality of the knowledge about μ itself. Since the usual probability theory does not help, the probability inversion is performed intuitively. In routine cases it usually works, but there are cases in which it fails*."

The confusion about probability in the two types of statistics manifests itself also in the nomenclature. Some examples are listed below.

- While frequentists speak of confidence intervals and confidence levels, Bayesians use the terms coverage interval (sometimes also called credible interval) and coverage probability.

- Frequentist use hypothesis testing to answer the question whether or not a null-hypothesis H_0 has to be rejected given an alternative hypothesis H_1. The test of a hypothesis can result in false results which are quantified by type 1 and type 2 errors. However, if a null-hypothesis H_0 is to be rejected that does not mean that the alternative hypothesis H_1 can be accepted.

  The hypothesis testing is particularly problematic with respect to the use of the *p*-value because it does not make a statement about the probability of a true value but rather just a statement about how probable it is to obtain a similar result as the actual one in a repetition of a measurement.

  Bayesians use a decision theory in combination with an appropriate loss function to decide between alternative hypotheses. The choice which hypothesis is accepted is the

[2] Actually, the Bayesian Statistics Formula (2.10) is wrong as we shall see in chapter 5.6.1, though its interpretation given here is correct.

result of a comparison of their respective probabilities (compare chapter 8 on decision thresholds)

- In epidemiology, the term “bias” is used to describe a systematic influence on the outcome of estimation. In metrology, the term “influencing factor” is used, which may be taken into account by a correction, which itself adds to the uncertainty associated with a measurement result.

- In the literature, in particular in epidemiology and risk assessments, one frequently finds the term “error”. The error concept, which defines an error as the difference between a measured value and the true value of a measurand, is outdated. Since the true value of a measurand is unknown and unknowable, this error cannot be quantified. The GUM therefore uses the term of “remaining error” which again is unknown and unknowable and consequently speaks only of uncertainty in measurement.

Though many results – but by far not all – of frequentist and Bayesian statistics are numerically identical, they must not be confused which each other because of the different definitions of probability.

What is the connection between uncertainty and metrology? This will be dealt with in the next chapter.

3 Metrology and statistics

3.1 Quantities, units and values

The International Vocabulary of Metrology – Basic and General Concepts and Associated Terms – (called VIM for short) defines a **quantity** as a property of a phenomenon, body, or substance, to which a number can be assigned with respect to a reference standard. By a measurement the value of a quantity is estimated. The **quantity value**, value of a quantity or simply value is a number and a measurement unit together expressing the magnitude of a quantity.

A quantity is characterized by quantity X = quantity value $\{X\}$ × measurement unit $[X]$, i.e.

$$X = \{X\} \cdot [X] \tag{3.1}$$

The **measurement unit, unit of measurement or simply unit** is a scalar quantity; defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number. A unit is a scalar quantity with the value 1. The unit of a quantity is according to the international SI-system expressed by up to 7 basic units.

$$Q: \quad [Q] = \mathrm{m}^{\alpha}\ \mathrm{kg}^{\beta}\ \mathrm{s}^{\gamma}\ \mathrm{A}^{\delta}\ \mathrm{K}^{\varepsilon}\ \mathrm{mol}^{\varsigma}\ \mathrm{cd}^{\eta} \text{ with rational numbers } \alpha, \beta, \ldots \tag{3.2}$$

The original SI-system consisted of prototypes representing the units. Consequently, the basic units were subject to changes with time and influenced by environmental conditions.

Under https://en.wikipedia.org/wiki/2019_redefinition_of_the_SI_base_units one finds the historical development. In 2019, the SI base units were redefined in agreement with the International System of Quantities, effective on the 144th anniversary of the Metre Convention, 20 May 2019. The redefinition followed an old idea which was proposed already by Max Planck in the year 1900. When formulating his radiation law, he pointed to the possibility to use "*natural measurement units*" which would be valid "*for all times and for all, even extra-terrestrial and extra-human cultures*".

Seven natural constants received fixed values in the new SI-system. The numerical values originated from equalization calculus by CODATA in the summer of 2017 (CODATA 2017 special adjustment). As natural constants are believed to have unique true values, the system of seven natural constants should be unambiguous.

Note that other quantities dealt with in metrology may have many true values as for instance the mass of different objects. However, in a given ideal measurement of a given object, however, the quantity "mass" is believed to have one unambiguous true value.

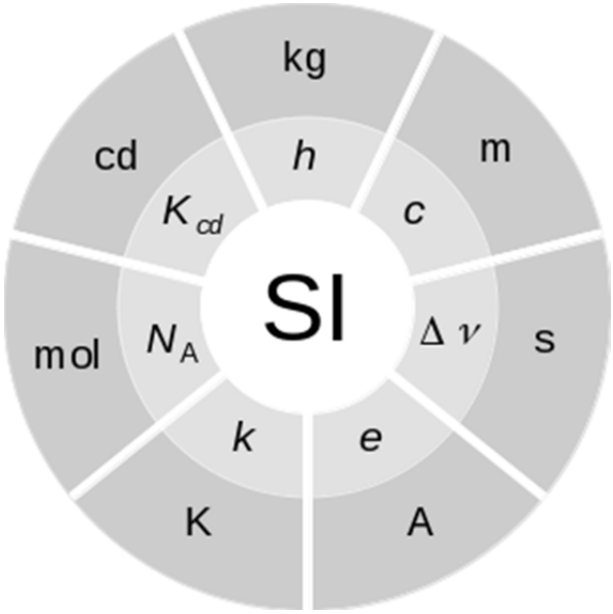


Fig. 5: Context of the new and the old SI-units (PTB 2017).

The seven natural constants are:

- Frequency of the **hyperfine structure transition** of the ground state in the ^{133}Cs-Atom, $\Delta\boldsymbol{\nu} = 9\ 192\ 631\ 770\ \mathrm{s}^{-1}$
- **Velocity of light** in vacuum, $\boldsymbol{c} = 299\ 792\ 458\ \mathrm{m\ s}^{-1}$
- **Planck constant**, $\boldsymbol{h} = 6{,}626\ 070\ 15 \cdot 10^{-34}\ \mathrm{J\ s}$ ($\mathrm{J\ s} = \mathrm{kg\ m^2\ s^{-1}}$)
- **Elementary charge**, $\boldsymbol{e} = 1{,}602\ 176\ 634 \cdot 10^{-19}\ \mathrm{C}$ ($\mathrm{C} = \mathrm{A\ s}$)
- **Boltzmann constant**, $\boldsymbol{k} = 1{,}380\ 649 \cdot 10^{-23}\ \mathrm{J\ K^{-1}}$ ($\mathrm{J\ K^{-1}} = \mathrm{kg\ m^2\ s^{-2}\ K^{-1}}$)

- **Avogadro constant**, $N_A = 6{,}022\ 140\ 76 \cdot 10^{23}\ \mathrm{mol}^{-1}$

- The **photometric radiation equivalent** K_{cd} of a monochromatic radiation with the frequency $540 \cdot 10^{12}$ Hz is exactly 683 Lumen per Watt.

The basic units of the old system are defined in the new SI-system as:

- **Second (s)**, $1\ \mathrm{s} = 9\ 192\ 631\ 770/\Delta\nu$

- **Meter (m)**, $1\ \mathrm{m} = (c/299\ 792\ 458)\ \mathrm{s} = 30{,}663\ 318\ldots\ c/\Delta\nu$

- **Kilogram (kg)**, $1\ \mathrm{kg} = (h/6{,}626\ 070\ 15 \cdot 10^{-34})\ \mathrm{m}^{-2}\ \mathrm{s} = 1{,}475\ 521... \cdot 10^{40}\ h\ \Delta\nu/c^2$

- **Ampere (A)**, $1\ \mathrm{A} = e/(1{,}602\ 176\ 634 \cdot 10^{-19})\ \mathrm{s}^{-1} = 6{,}789\ 686... \cdot 10^{8}\ \Delta\nu\ e$

- **Kelvin (K)**, $1\ \mathrm{K} = (1{,}380\ 649 \cdot 10^{-23}/k)\ \mathrm{kg\ m^2\ s^{-2}} = 2{,}266\ 665...\ \Delta\nu\ h/k$

- **Mol (mol)**, $1\ \mathrm{mol} = 6{,}022\ 140\ 76 \cdot 10^{23}/N_A$

- **Candela (cd)**, $1\ \mathrm{cd} = (K_{cd}/683)\ \mathrm{kg\ m^2\ s^{-3}\ sr^{-1}} = 2{,}614\ 830... \cdot 10^{10}\ (\Delta\nu)^2\ h\ K_{cd}$

3.2 Measuring the value of a quantity

One has to start with the definition of the measurand Y, i.e. the physical quantity of interest, and then to set-up the measurement procedure. It is anticipated here that the measurement procedure is suitable to meet the measurement objective and that it is fit for purpose. Together with the measurement procedure all possible influence quantities have to be identified which might affect the measurement and could be relevant with respect to the estimate y obtained as a measurement result. The setting-up of the measurement procedure and the consideration of the influencing phenomena or quantities comprise the establishment of the model of evaluation

$$Y = G(X_i, i = 1,...,n)\,, \tag{3.3}$$

connecting all quantities X_i relevant for the measurement mathematically to the measurand Y. A measurement model relates output quantities, about which information is required, to input quantities, about which information is available.

If the model of evaluation is correct, Equation (3.3) holds also for the true values of the input quantities and the measurand

$$\tilde{y} = G(\tilde{x}_i, i = 1,...,n) \tag{3.4}$$

By measurements one obtains estimates x_i $(i = 1,\ldots,n)$ of the input quantities and with Equation (3.3) an estimate of the value of the measurand.

$$y = G(x_i, i = 1,...,n) \tag{3.5}$$

This modelling of the measurement is the crucial step of any experiment. The model function G represents the scientific theory connecting the phenomena associated with the input quanti-

ties with that quantified by the measurand. The measurand has to be of relevance with the measurement objective and the model of evaluation has to describe the scientific theory or at least a meaningful approximation of it. In setting up a model of evaluation Ockham's razor should be applied.

Ockham's razor means today the principle of economy in science.

- It says that the simplest theory has to be preferred to other theories, explaining the same data.
- A theory has to be built up in a way that the internal connections are as simple as possible.
- Ockham's principle of economy in science demands, that one does not introduce more assumptions into the hypotheses, than actually needed in order to describe the data and to produce predictions which can be empirically verified.
- A hidden agenda is in this context, that hypotheses with few assumptions are easier to falsify than complex hypotheses and therefore to be preferred.

The simpler model is the better one of two models – until it is disproved by data.

The JCGM (2019) has issued a committee draft for comments which gives some advice for developing and using measurement models and also how to approach the problem of model uncertainties if different models appear to be meaningful.

3.3 Estimating the true value of a measurand

For the purpose of making a statement about the true value $\tilde{y}$ of a measurand Y a **random variable** is assigned to the quantity as an **estimator**. Frequently, no distinction is made between the symbol of the quantity and of the estimator; the character Y is used for both of them.

To work with such an estimator, some rules about random variables have to be looked at.

A set of elementary events may be describable in the way that a quantity X under random conditions shows values x from the set of real numbers $\mathbb{R}$. Then all events of this trial can be described by the variable X, called a random quantity or a random variable.

Consists $\mathbb{R}$ of a finite or infinite countable number of values, X is called a discrete random variable or quantity; consists $\mathbb{R}$ of all real numbers or of partial intervals X is called a continuous random variable or quantity.

The distribution density $f(x)$ and distribution function $F(x)$ (also called probability density function (PDF) and probability function, respectively) are defined by

$$\begin{aligned} &f(x)\,\mathrm{d}x = P(x < X < x + \mathrm{d}x)\,\mathrm{d}x \\ &P(a < X \le b) = \int_a^b f(x)\,\mathrm{d}x \quad F(x) = P(X \le x) = \int_{-\infty}^{x} f(x')\,\mathrm{d}x' \end{aligned} \tag{3.6}$$

An estimate of the true value of a measurand can be derived by calculating the PDF of the estimator. The best estimate is the expectation of the PDF and the squared standard uncertainty is its variance as explained in detail in the next chapter.

In metrological practice this means that the output quantity Y is connected to a number of in input quantities X_i $(i = 1,...,m)$ by a model of evaluation $Y = G(X_i, i = 1,...,m)$. The available information consists of the estimates x_i $(i = 1,...,m)$ of the input quantities and their associated standard uncertainties. The $x_i (i = 1,...,m)$ are given and fixed and from them one calculates the primary measurement result via $y = G(x_i, i = 1,...,m)$. The associated uncertainties $u(x_i)$ $(i = 1,...,m)$ and $u(y)$ have to be derived on the basis of the GUM or the GUM S1. This involves the PDFs $f_{X_i}(\tilde{x}_i | x_i, \Im)$ and, finally, $f_Y(\tilde{y} | \mathbf{x}, \Im)$ from which the best estimate $\hat{y}$ of Y and its associated standard uncertainty $u(\hat{y})$ are calculated as $\hat{y} = \mathrm{E}\left(f_Y(\tilde{y} | \mathbf{x}, \Im)\right)$ and $u^2(\hat{y}) = \mathrm{Var}\left(f_Y(\tilde{y} | \mathbf{x}, \Im)\right)$.

4 Available information and probability density functions (PDFs)

4.1 General aspects

There are basically three tools for establishing and updating PDFs; i.e. the Principle of Maximum Entropy (PME), the Bayes Theorem and the Product Rule. The GUM S1 makes use of the PME as well as of the Bayes Theorem.

According to E.T. Jaynes "Probability Theory: The Logic of Science" (Jaynes 1982) the PDF sought can be derived by the **Principle of Maximum Entropy (PME)** which requires

$$S = -\int f_Y\left(\tilde{y} | y\right) \cdot \ln\left(f_Y\left(\tilde{y} | y\right)\right) \mathrm{d}y = \max \tag{4.1}$$

The PME consists in choosing as $f_Y\left(\tilde{y} | y\right)$ the most likely density by taking into account relevant information available as so-called constraints and maximizing the entropy S by applying a variational method. If nothing was known before, then the prior is uniform according to the Bernoulli principle (see definition of the prior below). Only this case is assumed by the PME application of the GUM supplement. It is stressed once more that this solution of the PME depends essentially on the nearly always incomplete information available and taken into account. The GUM S1 explains regarding the PME:

"*GUM S1: 6.3 Principle of maximum entropy*

6.3.1 When using the principle of maximum entropy, introduced by Jaynes [25], a unique PDF is selected among all possible PDFs having specified properties, e.g. specified central moments of different orders or specified intervals for which the PDF is non-zero. This method is particularly useful for assigning PDFs to quantities for which a series of indications is not available or to quantities that have not explicitly been measured at all.

6.3.2 In applying the principle of maximum entropy, to obtain a PDF $f_Y\left(\tilde{y} | y\right)$ that adequately characterizes incomplete knowledge about a quantity X according to the information avail-

able, the functional $S = -\int f_Y\left(\tilde{y}\middle|y\right)\cdot\ln\left(f_Y\left(\tilde{y}\middle|y\right)\right)\mathrm{d}y$ *the "information entropy", introduced by Shannon [48], is maximized under constraints given by the information*."

Note that in the clause above the original functional notations have been changed in order to conform to the notation of this booklet.

Another tool is provided by the **Bayes Theorem**

$$f_Y(\tilde{y}|\,y,\mathfrak{I}) = \frac{f_Y(y|\,\tilde{y})\cdot f_Y(\tilde{y}|\,\mathfrak{I})}{f_Y(y|\,\mathfrak{I})} \tag{4.2}$$

or by naming the different terms $Posterior\ f_Y(\tilde{y}|\,y,\mathfrak{I}) = \frac{Likelihood\ f_Y(y|\,\tilde{y})}{Evidence\ f_Y(y|\,\mathfrak{I})} \times Prior\ f_Y(\tilde{y}|\,\mathfrak{I})$.

The Bayes Theorem was published posthumously in an „Essay towards solving a problem in the doctrine of chances" Bayes (1763). It introduces the so-called "„Bayesian estimation", i.e. calculating the probability of the validity of a proposition on the basis of a prior estimate of its probability and new relevant evidence." In other words, the Bayes Theorem provides the mathematical form of human learning by starting from available knowledge and updating this according to new relevant information.

In the Bayes Theorem $f_Y(\tilde{y}|\,y,\mathfrak{I}) = C\cdot f_Y(y|\,\tilde{y})\cdot f_Y(\tilde{y}|\,\mathfrak{I})$ the prior knowledge is described by the so-called prior $f_Y(\tilde{y}|\,\mathfrak{I})$ which considers all information available before an experiment is performed. The term $f_Y(y|\,\tilde{y})$ is the so-called likelihood. By multiplication of the prior with the likelihood one obtains after normalization with the normalization constant C the posterior $f_Y(\tilde{y}|\,y,\mathfrak{I})$ sought.

A third tool is the **Product rule**[3]

$$f_Y(\tilde{y}|\,y,\mathfrak{I}) = f(\tilde{y}|\,y)\cdot f(\tilde{y}|\,\mathfrak{I}) \tag{4.3}$$

which simply considers a logical "and" for two disjoint conditions for the probability of a result by multiplying the respective probabilities. If the conditions are not disjoint the first equation in footnote 3 applies. Some applications of these tools are described in the next chapter. Further examples may be found in the GUM S1.

All the PDFs discussed in this chapter play important roles in measurements for radiation protection purposes, such as dosimetry and environmental radioactivity. Table A1 in the Appendix gives a survey on these applications.

Alternatively to these PDFs the PME, the Bayes' Theorem or the product rule of the probability theory can be applied to include known frequency or parameter distributions, respectively. If more information about the input quantity involved is available than assumed in ISO 11929, it can be used in form of suitable model priors considered in the Bayes theorem.

[3] Product Rule: $f(A,B) = f(A\wedge B) = f(A)\cdot f(B|\,A)$; if A and B are independent, the Product Rule simply reads $f(A,B) = f(A\wedge B) = f(A)\cdot f(B)$.

A wealth of further distributions exists and may be used depending on the available information. The GUM S1 gives advice for quite cases. In addition, also empirically frequency distributions from earlier assessments can provide a basis to estimate the uncertainty associated with a quantity value. See for instance the paper by Cox et al. (2006) regarding neutron dosimetry.

Table A1 in the Appendix summarizes various PDFs frequently used in metrology, in general, and in radiation protection in particular. Here some of them are dealt with in detail.

Thomas Bayes
* 1702 † 7.4.1761

Andrei Andrejewitsch Markov
* 2.6.1856; † 20.7.1922

Edwin Thompson Jaynes
* 5.7.1922 – † 30.4.1998

Fig. 6: Famous persons in the context of a Bayes Theory of uncertainty in metrology (photos taken from Wikipedia).

4.2 The normal or Gaussian distribution: only y and $u(y)$ are known

If only y and $u(y)$ are known, they are the best estimate and its associated standard uncertainty. If a best estimate y and its associated standard uncertainty are the only information available regarding the quantity Y, then according to the PME the PDF assigned to Y is a Gaussian distribution $\mathrm{N}(y,u^2(y))$

$$\mathrm{N}(y,u^2(y)) = f_Y(\tilde{y}\,|\,y,u(y)) = \frac{1}{\sqrt{2\pi}\cdot u(\mathrm{y})}\exp\left(-\frac{(\tilde{y}-y)^2}{2u^2(y)}\right) \tag{4.4}$$

with the expectation value $\mathrm{E}(f_Y(\tilde{y}\,|\,y,u(y))) = \hat{y}$ and the variance $\mathrm{Var}(f_Y(\tilde{y}\,|\,y,u(y))) = u^2(\hat{y})$.

4.2.1 Some features of the normal distribution

There are some remarkable special features of the standard normal distribution which are explained below.

First, the sum of independent Gaussian-distributed random variables is a Gaussian-distributed random variable (Fig. 7).

$$X_1 \sim N(\mu_1,\sigma_1); X_2 \sim N(\mu_2,\sigma_2) \Rightarrow Y = X_1 + X_2 \sim N(\mu = \mu_1 + \mu_2, \sigma = (\sigma_1^2 + \sigma_2^2)^{1/2}) \quad (4.5)$$

Second, the **Central Limiting Theorem of Statistics** reads: Every sum of independent random variables with arbitrary distributions goes for a Gaussian or normal distribution, if the number of random variables goes for ∞.

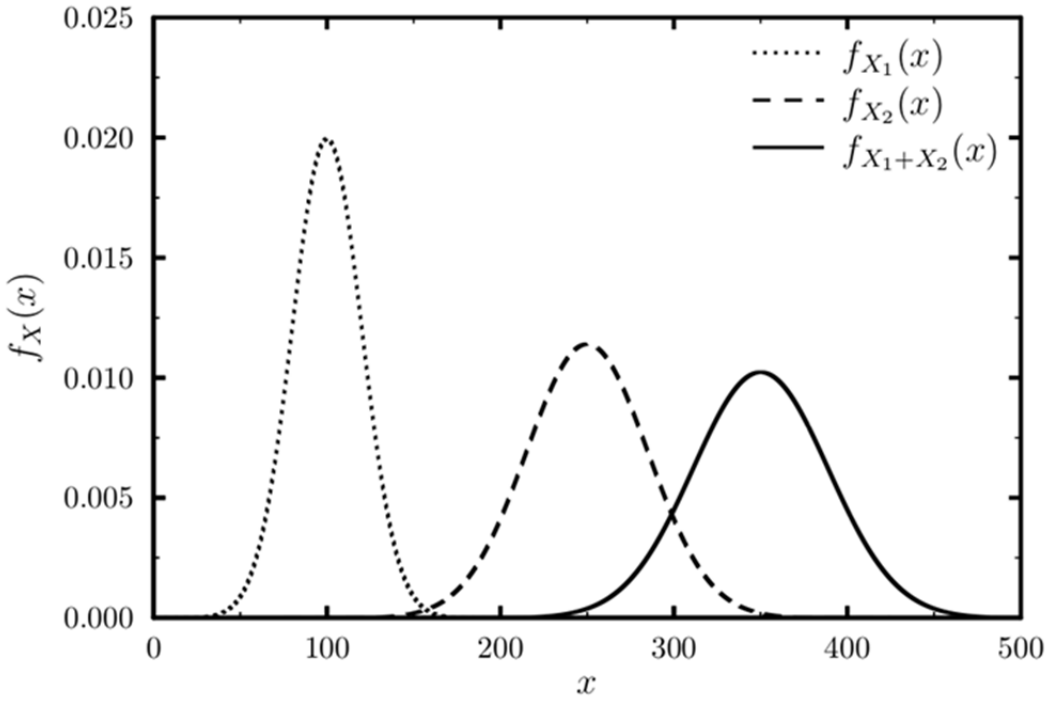


Fig. 7: Example for the sum of independent Gaussian distributions.
It holds $\mu_1 = 100$, $\sigma_1 = 20$, $\mu_2 = 250$, $\sigma_2 = 35$, $\mu_{1+2} = 350$, $\sigma_{1+2} = 39$.

Third, given the conditional PDFs of two random variable from two normal distributions, distinguished as cases 1 and 2, the PDF of a random variable given case 1 and case 2 is also a normal PDF (Fig. 8):

$$\begin{aligned} &P_1 = P(x_1 | case\,1) \sim N(\mu_1,\sigma_1) = f_{X_1}(x_1); P_2 = P(x_2 | case\,2) \sim \mathrm{N}(\mu_2,\sigma_2) = f_{X_2}(x_2) \\ &\Rightarrow P = P(x_{1\wedge 2} | case\,1 \wedge case\,2) = P_1 \cdot P_2 \sim \mathrm{N}(\mu_{1,2},\sigma_{1,2}) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \end{aligned} \quad (4.6)$$

4.2.2 Quantiles of a PDF

For the calculation of the characteristic limits according to ISO 11929 quantiles of suitable PDFs are important. A quantile k of a PDF $f(x)$ for a probability α is defined by $F(k)=\alpha$ with $F(x)$ being the distribution function $F(x)$ of the PDF $f(x)$.

Particularly important is the standard normal distribution N(0,1)

$$\mathrm{N}(0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(-y^2/2\right) \quad (4.7)$$

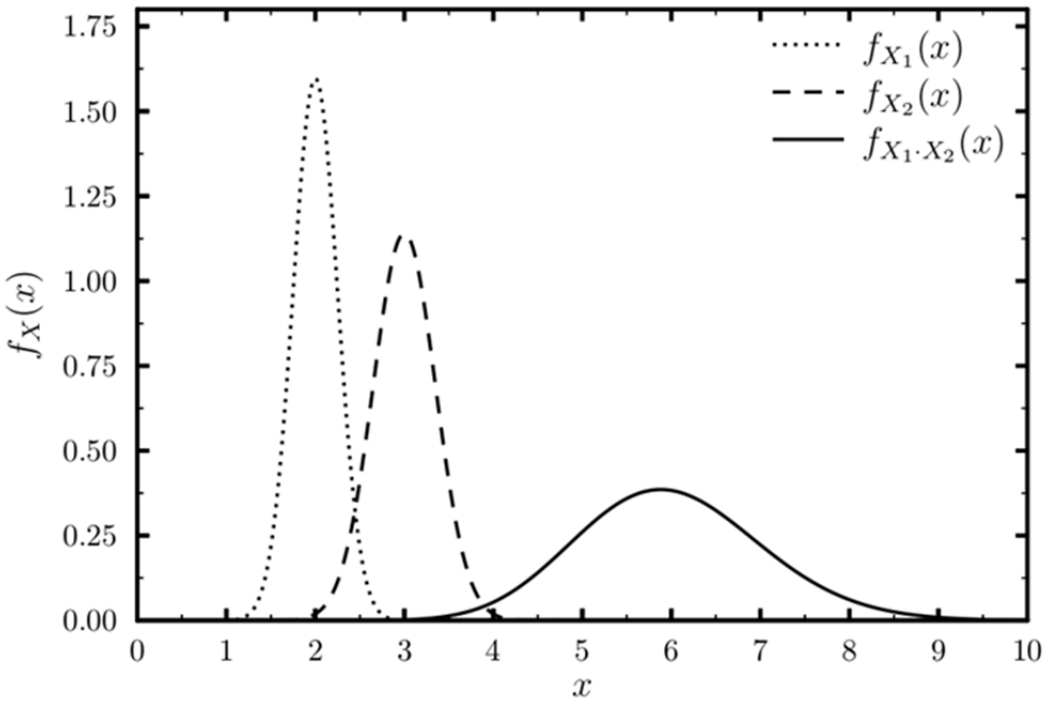


Fig. 8: Example for the folding of two independent Gaussian distributions. It holds $\mu_1 = 90$, $\sigma_1 = 20$, $\mu_2 = 150$, $\sigma_2 = 30$.

It allows sampling from $\mathrm{N}(y,\sigma^2)$ by making a draw z from the standard normal distribution $\mathrm{N}(0,1)$ and calculating $y' = y + u(y)\cdot z$.

A normal distribution $\mathrm{N}(\mu,\sigma^2)$ with any parameters μ and σ, in our case y and $u(y)$, and the distribution function $F(y)$ has the following relationship to the standard normal distribution $\mathrm{N}(0,1)$:

$$F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right) \tag{4.8}$$

with Φ being the distribution function of the standard normal distribution.

If a random variable $Y \sim \mathrm{N}(\mu,\sigma^2)$, then the standardization $z = \frac{y-\mu}{\sigma}$ leads to a standard normally distributed random variable Z because

$$P\left(\frac{y-\mu}{\sigma} \le z\right) = P(y \le \sigma\cdot z + \mu) = F(\sigma\cdot z + \mu) = \Phi(z) \tag{4.9}$$

Of particular importance are the quantiles of the standard normal distribution (Fig. 9). The distribution function of the standard normal distribution $\mathrm{N}(0,1)$ is defined by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{v^2}{2}\right)\mathrm{d}v = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\sum_{j=0}^{\infty}\frac{x^{2j+1}}{1\cdot 3\cdots(2j+1)} \tag{4.10}$$

and its quantile k_p for the probability p by $\Phi(k_p) = p$. The relations $\Phi(-x) = 1 - \Phi(x)$ and $k_{1-p} = -k_p$ apply.

The one-sided lower quantile k_α of the standard normal distribution for the probability α is defined by $\Phi(k_\alpha) = \alpha$. The one-sided upper quantile $k_{1-\alpha}$ of the standard normal distribution for the probability α is defined by $\Phi(k_{1-\alpha}) = 1-\alpha$. The two-sided quantiles of the standard normal distribution $k_{1-\gamma/2}$ of the standard normal distribution for the probability γ are defined by $\Phi(k_{1-\gamma/2}) = \Phi(k_{\gamma/2}) = \gamma/2$. The standard normal distributions as well as its frequently used one-sided and two-sided quantiles are shown in Fig. 9.

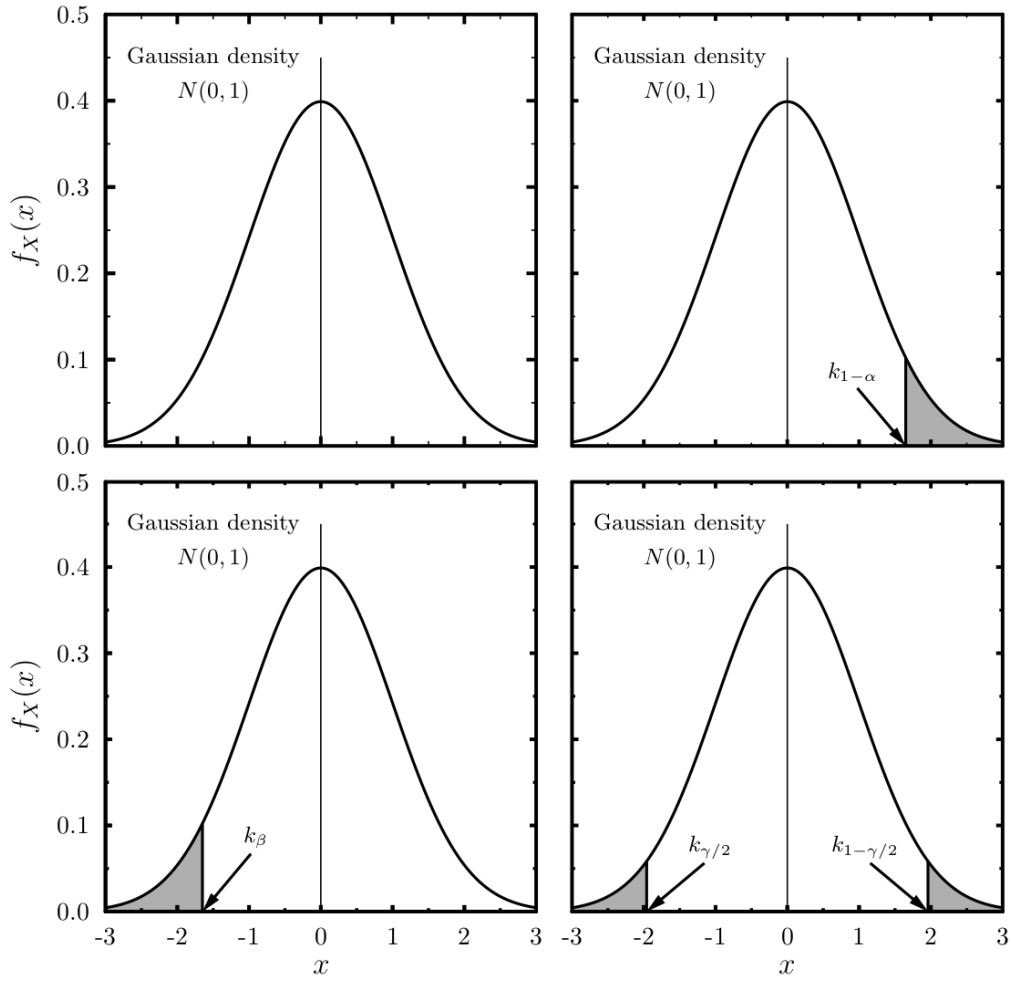


Fig. 9: The standard normal distribution N(0,1) and their one-sided and two-sided quantiles for the probabilities α, β, and γ.

The particular feature of the standard normal distribution $\mathrm{N}(0,1)$ is that the quantiles of any normal distribution $\mathrm{N}(\mu,\sigma)$ can be calculated using the quantiles of the standard normal distribution. The one-sided lower quantile is $\mu + k_\alpha \cdot \sigma$ with $P(x < \mu + k_\alpha \cdot \sigma) = \alpha$, the one-sided upper quantile is $\mu + k_{1-\alpha} \cdot \sigma$ with $P(x > \mu + k_{1-\alpha} \cdot \sigma) = \alpha$, and the two-sided quantiles are $\mu \pm k_{1-\gamma/2} \cdot \sigma$ with $P(x > \mu + k_{1-\gamma/2} \cdot \sigma) = \gamma/2$ and $P(x < \mu - k_{1-\gamma/2} \cdot \sigma) = \gamma/2$.

The numerical values of the quantiles of the standard normal distribution are tabulated, e.g. in ISO 11929-1:2019 and DIN ISO 11929-1:2020. Frequently used values are the one-sided quantiles for the probability $\alpha = 0,05$ with $k_{1-\alpha} = -k_{\alpha} = 1,65$, respectively $\beta = 0,05$ with $k_{1-\beta} = -k_{\beta} = 1,65$, and the two-sided quantiles for the probability $\gamma = 0,05$ with $k_{1-\gamma/2} = -k_{\gamma/2} = 1,96$.

4.3 The Gamma distribution: count rate measurements

The measurand is the count rate *R* with the true value $\tilde{r}$. In order to obtain an estimate of the count rate a counting measurement is performed. Let *n* ionizing-radiation events be recorded in a counting measurement of a fixed duration *t* in order to measure a count rate. The number *n* of counts is assumed to be drawn from an underlying **Poisson frequency distribution of a random variable *N* with an unknown parameter** $\tilde{r} \cdot t \geq 0$. $\tilde{r}$ is the true value of the count rate *R* which is the measurand.

R is estimated by counting of *n* random decay events during a measurement time *t*. The Bayes Theorem gives then for the posterior PDF:

$$f_R(\tilde{r}|n,1/t) = C \cdot f_N(n|\tilde{r},t) \cdot f_R(\tilde{r}) \tag{4.11}$$

With this one obtains the likelihood $f_N(n|\tilde{r},t) = e^{-\tilde{r}\cdot t}(\tilde{r}\cdot t)^n / n!; \quad n = 0,1,2,...$

In order to establish the PDF of the count rate *R* Bayes Theorem is applied and yields

$$f_R(\tilde{r}|n,1/t) = \frac{f_N(n|\tilde{r}\cdot t) \cdot f_R(\tilde{r})}{f_N(n)} \text{ with } f_N(n) = C^{-1} \text{ according to equation (4.11)} \tag{4.12}$$

$$f_R(\tilde{r}|n,1/t) \propto f_N(n|\tilde{r},t) \cdot f_R(\tilde{r}) = f_R(\tilde{r}) \cdot e^{-\tilde{r}\cdot t}(\tilde{r}\cdot t)^n / n! \tag{4.13}$$

The **likelihood** $f_N(n|\tilde{r},t)$ considers the fact that the counts are random draws from a stationary Poisson process with the parameter $\tilde{r} \cdot t \geq 0$ and hence is written

$$f_N(n|\tilde{r},t) = e^{-\tilde{r}\cdot t}(\tilde{r}\cdot t)^n / n!; \quad n = 0,1,2,... \tag{4.14}$$

According to Jaynes (1968) the **non-informative prior** $f_R(\tilde{r})$ of Jeffreys (1946) given in Equation (4.15) should be used for a stationary Poisson process; e.g.

$$f_R(\tilde{r}) = C / \tilde{r}; \quad (\tilde{r} \geq 0) \tag{4.15}$$

By inserting this prior into Equation (4.13) and normalization one obtains the **gamma density function**

$$f_R(\tilde{r}|n,1/t) = t \cdot (\tilde{r}\cdot t)^{n-1} \cdot e^{-\tilde{r}\cdot t} / (n-1)!; \quad (\tilde{r} \geq 0) \tag{4.16}$$

The expectation $\mathrm{E}(f_R(\tilde{r}|n,1/t)) = r = n/t$ is the measurement result and the variance $\mathrm{Var}(f_R(\tilde{r}|n,1/t)) = r/t = n/t^2$ leads to the associated standard uncertainty. This information is used in ISO 11929-1:2019. See Fig. 10 for some examples of gamma distributions.

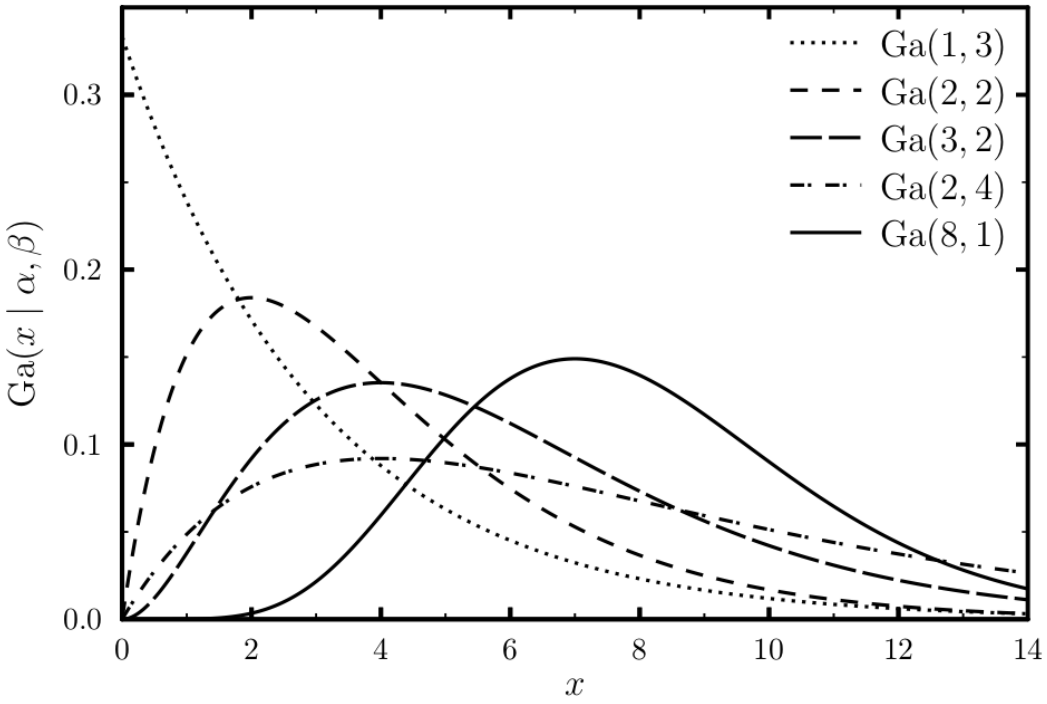


Fig. 10: Examples of Gamma distributions.

With the Bayes Theorem $f(\tilde{y}|y) \propto f(y|\tilde{y}) \cdot f(\tilde{y})$ and the Likelihood $f(y|\tilde{y})$, the non-informative prior according to Jeffreys (1946) together with the so called Fisher Information, $I_{\tilde{y}}$, $f(\tilde{y}) \propto \sqrt{E_{\tilde{y}}\left(\frac{\partial f(y|\tilde{y})}{\partial \tilde{y}}\right)^2} = \sqrt{I_{\tilde{y}}}$, fulfils the requirement of invariance under transformation or under parameterization. Jeffreys' priors are in general a class of priors $f(\tilde{r}) \propto 1/\tilde{r}^x$ with $x \in [0,1]$ which satisfy this requirement.

The Fisher information $I_{\tilde{y}}$ is an indicator for the amount of information about $\tilde{y}$ which is added by the prior. The request to minimize $I_{\tilde{y}}$ and thereby to obtain the reference prior is equivalent to minimize the influence of the prior and to maximize information entropy.

As discussed in detail by Weise et al. (2013) and in ISO 11929-2:2019 Annex A, the particular case $n = 0$ must be treated separately: $\mathrm{E}\left(f_R(\tilde{r}|n,1/t)\right)$ and $\mathrm{Var}\left(f_R(\tilde{r}|n,1/t)\right)$ vanish, and $f_R(\tilde{r}|n,t) = \delta(\tilde{r})$ and a zero uncertainty follows. This is not reasonable since one can never be sure that exactly $R = 0$ if no event happens to be recorded in a measurement of finite duration. Thus, no reasonable statement can be made on the count rate R if $n = 0$. With any more realistic prior $f_R(\tilde{r})$, one should always obtain $\mathrm{E}\left(f_R(\tilde{r}|n,t)\right) > 0$ and $\mathrm{Var}\left(f_R(\tilde{r}|n,t)\right) > 0$.

To avoid this shortcoming it is assumed that the counting measurement is carried out with a duration t chosen suitably large according to the experience of former, similar measurements, so that for any reasonable $\tilde{r} > 0$ at least a few counts can be expected. The duration t is therefore no longer arbitrary. This knowledge can justify another non-informative prior out of the class of non-informative Jeffreys priors (Jeffreys 1961) with

$$f_R(\tilde{r}) = C / \tilde{r}^{\,v} \text{ with } 0 \le v < 1 \tag{4.17}$$

One can argue that *R* will be bounded for physical or experimental reasons, although a sufficiently large upper bound need not be specified explicitly. This knowledge is represented by an equally likely *R* between zero and the upper bound, thus, by a uniform prior

$$f_R(\tilde{r}) = C \text{, i.e. } v = 0 \tag{4.18}$$

The Bayes theorem and normalization then yield the **gamma density function** (for examples see Fig. 10).

$$f_R(\tilde{r}\,|\,n, 1/t) = t \cdot (\tilde{r} \cdot t)^n \cdot e^{-\tilde{r} \cdot t} / n!; \quad (\tilde{r} \ge 0) \tag{4.19}$$

The expectation $\mathrm{E}\left(f_R(\tilde{r}\,|\,n,t)\right) = r = (n+1)/t$ is the measurement result and the variance $\mathrm{Var}\left(f_R(\tilde{r}\,|\,n,1/t)\right) = u^2(r) = (n+1)/t^2$ leads to the associated standard uncertainty. This information is used in ISO 11929-2:2019 Annex A for small count numbers.

This result is more reasonable for $n = 0$ since the standard uncertainty $u(r) = \sqrt{n+1}\,/\,t$ does not vanish, and the interval with limits $r \pm u(r)$ of reasonable estimates of the measurand according to GUM turns out to also contain the value $\tilde{r} = 0$. Asymptotically for large n, both approaches lead to the same results. The main differences only occur for very small n.

By introducing a so-called **reference prior** (Bernardo 1979), which minimizes $I_{\tilde{y}}$, a prior is preferred that maximizes the missing information in an experiment. In the case of a Poisson process the reference prior according to Yang and Berger (1996) is $f(\tilde{y}) \propto 1/\sqrt{\tilde{y}}$. The reference prior $f_R(\tilde{r}\,|\,t) = C/\sqrt{\tilde{r}}$ is the prior which minimizes the so-called Fisher-information and maximizes the Shannon-Jaynes-information-entropy and thereby adds the least to the available information (Bernardo 1979, Berger et al 2009).

With a non-informative reference prior $f_R(\tilde{r}\,|\,t) = C/\sqrt{\tilde{r}}$ one obtains the posterior

$$f_R(\tilde{r}\,|\,n, 1/t) = \frac{t \cdot e^{-\tilde{r} \cdot t} (\tilde{r} \cdot t)^{n-1/2}}{(n-1/2)!} \quad \tilde{r} \ge 0\,. \tag{4.20}$$

This is the Gamma-distribution $\mathrm{G}(\tilde{r}; n+1/2, 1/t)$ with

$$\mathrm{E}\left(f_R(\tilde{r}\,|\,n,t)\right) = r = (n+1/2)/t \text{ and } \mathrm{Var}\left(f_R(\tilde{r}\,|\,n,t)\right) = r/t = (n+1/2)/t^2 \tag{4.21}$$

$$\Rightarrow \quad u(r) = \sqrt{r/t} = \sqrt{(n+1/2)/t^2}\,.$$

Though the discussions about the right non-informative prior are epistemologically very interesting, the results are of little practical relevance. Whether $u(r)$ is equal to $\sqrt{n/t^2}$, $\sqrt{(n+1/2)/t^2}$ or $\sqrt{(n+1)/t^2}$, is only relevant for measurements with extremely small numbers of counted events. Whether such measurements can meet any reasonable measurement objective can be doubted – at least for the purpose of radiological protection.

4.4 The *t*-distribution: repeated measurements

The Student's *t*-distribution has the PDF

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\cdot\pi}\cdot\Gamma\left(\nu/2\right)}\cdot\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \tag{4.22}$$

with ν degrees of freedom and Γ being the Gamma function (Fig. 11). With $t = \frac{\mu - \overline{x}}{s/\sqrt{n}}$ it finds its application in repeated measurements.

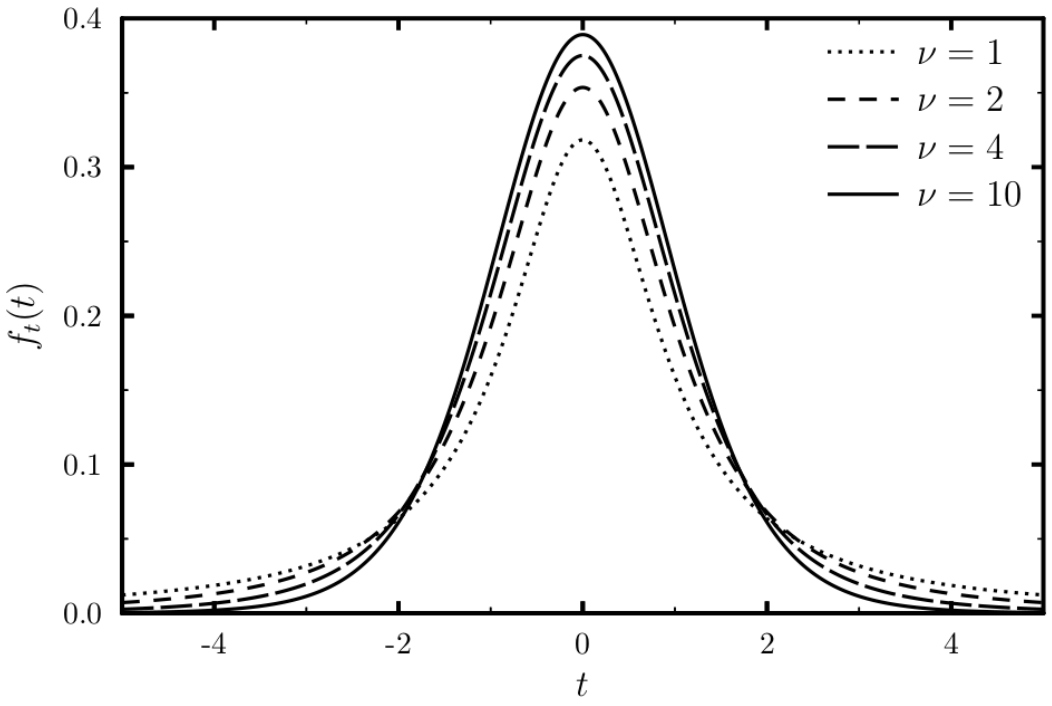


Fig. 11: Example of a *t*-distribution for different degrees of freedom ν.

Given n repeated measurements with the indications $x_1, x_2, ..., x_n$ for a quantity with unknown expectation μ_0 and variance σ_0 which is normally distributed according to $\mathrm{N}(\mu_0, \sigma_0^2)$ the resulting PDF is the scaled and shifted *t*-distribution $t_\nu(\overline{x}, s^2/n)$ with $\nu = n-1$ degrees of freedom (GUM S1 chapter 6.4.9). $\overline{x}$ is the average and s^2 the variance of the indications.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ and } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{4.23}$$

The PDF for X is

$$f_X(x|\overline{x}, s) = \frac{\Gamma(n/2)}{\sqrt{(n-1)\cdot\pi}\cdot\Gamma((n-1)/2)} \cdot \frac{1}{s/\sqrt{n}} \cdot \left(1 + \frac{1}{n-1}\cdot\left(\frac{x-\overline{x}}{s/\sqrt{n}}\right)^2\right)^{-n/2} \tag{4.24}$$

where

$$\Gamma(z) = \int_0^\infty t^{z-1}\cdot e^{-t}\,\mathrm{d}t \quad (z > 0)\,. \tag{4.25}$$

has the expectation and the variance

$$\mathrm{E}(X) = \overline{x} \text{ and } \mathrm{Var}(X) = \frac{n-1}{n-3}\cdot\frac{s^2}{n} \tag{4.26}$$

and the best estimate and its associated standard uncertainty are therefore

$$\hat{x} = \overline{x} \text{ and } u^2(\hat{x}) = \frac{n-1}{n-3}\cdot\frac{s^2}{n} \tag{4.27}$$

An application of this case are repeated or "black-box" measurements in which primary results y_i of the measurand Y are directly indicated without any information for the user about the measurement technique and the model of evaluation involved. There are just series of observations y_i. The measurement problem is to compare a series of indications $y_{\mathrm{b},i}$ $(i = 1,...,n_\mathrm{b})$ which are judged by the user to represent a "normal", "background" or "blank" situation with a series of indications $y_{\mathrm{g},i}$ $(i = 1,...,n_\mathrm{g})$ for another situation, called "gross" situation. To this end, a net quantity Y is investigated

$$Y = \overline{Y}_\mathrm{g} - \overline{Y}_\mathrm{b} \tag{4.28}$$

and the characteristic limits are evaluated for the measurand Y.

Primary estimates of $\overline{Y}_\mathrm{g}$ and $\overline{Y}_\mathrm{b}$ are obtained as the arithmetic means

$$\overline{y}_\mathrm{g} = \frac{1}{n_\mathrm{g}}\sum_{i=1}^{n_\mathrm{g}} y_{\mathrm{g},i} \text{ and } \overline{y}_\mathrm{b} = \frac{1}{n_\mathrm{b}}\sum_{i=1}^{n_\mathrm{b}} y_{\mathrm{b},i}\,, \tag{4.29}$$

respectively, with their respective sample standard deviations

$$s_\mathrm{g} = \left(\frac{1}{n_\mathrm{g}-1}\sum_{i=1}^{n_\mathrm{g}}(y_{\mathrm{g},i} - \overline{y}_\mathrm{g})^2\right)^{1/2} \text{ and } s_\mathrm{b} = \left(\frac{1}{n_\mathrm{b}-1}\sum_{i=1}^{n_\mathrm{b}}(y_{\mathrm{b},i} - \overline{y}_\mathrm{b})^2\right)^{1/2} \tag{4.30}$$

Since no other information is available the $y_{\mathrm{b},i}$ $(i = 1,\ldots,n_{\mathrm{b}})$ and $y_{\mathrm{g},i}$ $(i = 1,\ldots,n_{\mathrm{g}})$ are assumed to be samples from Gaussian distributions with unknown expectations and variances. According to GUM S1 Clause 6.4.9 the arithmetic means $\overline{y}_{\mathrm{g}}$ and $\overline{y}_{\mathrm{b}}$ according to Equation (4.29) are the best estimates and the standard uncertainties associated with $\overline{y}_{\mathrm{g}}$ and $\overline{y}_{\mathrm{b}}$ are given by

$$u^2(\overline{y}_{\mathrm{g}}) = \frac{n_{\mathrm{g}}-1}{n_{\mathrm{g}}-3}\cdot\frac{s_{\mathrm{g}}^2}{n_{\mathrm{g}}}\ ,\ u^2(\overline{y}_{\mathrm{b}}) = \frac{n_{\mathrm{b}}-1}{n_{\mathrm{b}}-3}\cdot\frac{s_{\mathrm{b}}^2}{n_{\mathrm{b}}} \quad \text{and} \quad u^2(y) = u^2(\overline{y}_{\mathrm{g}}) + u^2(\overline{y}_{\mathrm{b}})\ . \tag{4.31}$$

In the context of the methods described in ISO 11929-2:2019, the PME yields the scaled and shifted *t*-distributions $t_{n_{\mathrm{g}}-1}(\overline{y}_{\mathrm{g}}, s_{\mathrm{g}}^2/n_{\mathrm{g}})$ and $t_{n_{\mathrm{b}}-1}(\overline{y}_{\mathrm{b}}, s_{\mathrm{b}}^2/n_{\mathrm{b}})$ as the respective PDFs (GUM S1 Clause 9.2.3).

Evidently $n_{\mathrm{b}}, n_{\mathrm{g}} > 3$ are required. The choice of suitable and meaningful numbers of indications $n_{\mathrm{b}} > 3$ depends on the judgement of the user, the measurement objective and the prevailing circumstances.

4.5 The negative binomial distribution: repeated counting measurements with random influences

The **negative binomial distribution** (Fig. 12) represents a continuous mixture of a Poisson distribution and a gamma distribution. It is of practical value in all those cases where sample treatments, such as taking aliquots, positioning before the detector, chemical separations, etc., add further (mostly unknown) uncertainties to those of the Poisson process. This leads to a overdispersion of the measured counts and $\mathrm{Var}(f(n)) > \mathrm{E}(f(n))$.

Because of this, the negative binomial distribution is also known as the **gamma–Poisson (mixture) distribution**. "*It is especially useful for discrete data over an unbounded positive range whose sample variance exceeds the sample mean. In such cases, the observations are overdispersed with respect to a Poisson distribution, for which the mean is equal to the variance. Hence a Poisson distribution is not an appropriate model. Since the negative binomial distribution has one more parameter than the Poisson, the second parameter can be used to adjust the variance independently of the mean* (Wikipedia: negative binomial distribution, 15.4.2021)".

The negative binomial distribution is a Poisson distribution, where the parameter λ is itself a random variable, distributed as a gamma distribution with the shape r and the scale $\theta = p/(1-p)$ or correspondingly the rate $\beta = (1-p)/p$.

By folding the Gamma distribution with a Poisson distribution with the ingredients in Equation (4.32) one obtains the distribution density $f(X=k)$

$$f(X=k) = \frac{\lambda^k}{k!}\cdot e^{-\lambda} \quad \text{and} \quad f(\lambda) = \frac{x^{a-1}\cdot e^{-\lambda/b}}{\Gamma(a)\cdot b^a} \quad \text{with } \lambda \geq 0, a, b > 0 \tag{4.32}$$

$$f(X=k) = \int_0^{\infty} \frac{\lambda^k}{k!}\cdot e^{-\lambda}\cdot f(\lambda)\,\mathrm{d}\,\lambda \tag{4.33}$$

and after some longer derivation (Mende 1990)

$$f(X=k)=\frac{\Gamma(r+k)}{\Gamma(r)\cdot k!}\cdot(1+b)^{-a}\cdot\left(\frac{b}{b+1}\right)^{k} \tag{4.34}$$

With $r=a$ and $p=\frac{1}{b+1}$ one obtains with Equation (4.34) the negative binomial distribution

$$f(X=k)=\frac{\Gamma(r+k)}{\Gamma(r)\cdot k!}\cdot p^{r}\cdot(1-p)^{k}\quad k=0,1,2,... \tag{4.35}$$

The expectation and the variance of the negative binomial distribution can be calculated with the parameters given in Equation (4.36).

$$\mathrm{E}(X)=\frac{r\cdot(1-p)}{p}\ \text{and}\ \mathrm{Var}(X)=\frac{r\cdot(1-p)}{p^{2}}=\mathrm{E}(X)+\frac{\mathrm{E}(X)^{2}}{r} \tag{4.36}$$

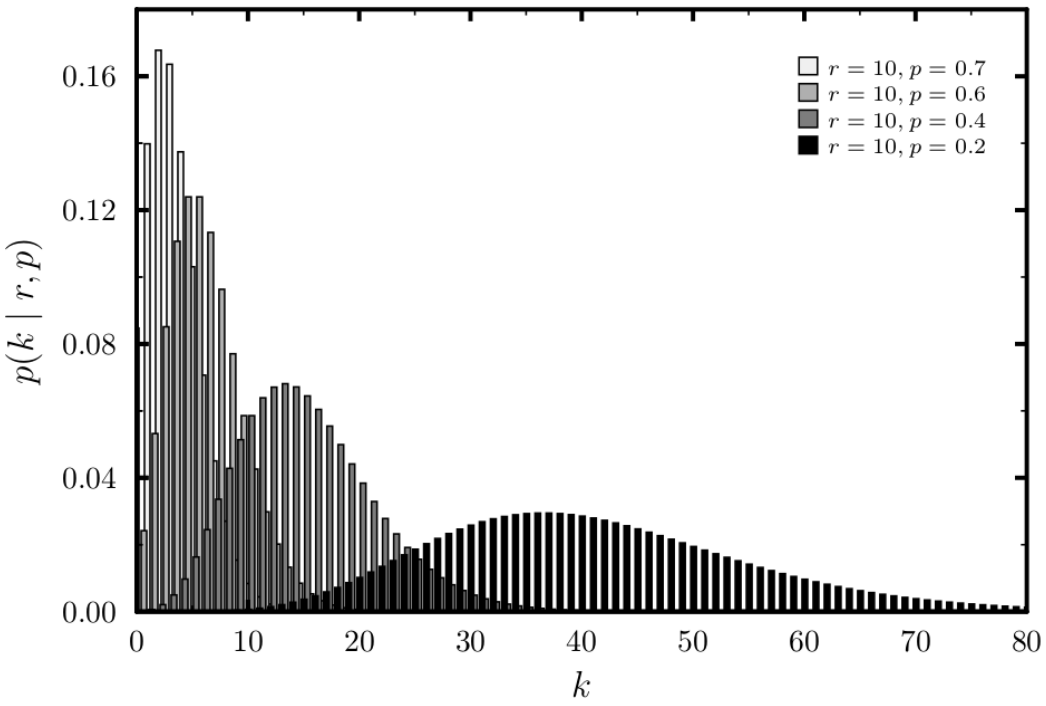


Fig. 12: Examples of negative binomial distributions.

The parameters r and p result from the expectation and the variance of the distribution by

$$r=\frac{\mathrm{E}(X)^{2}}{\mathrm{Var}(X)-\mathrm{E}(X)}=\frac{\mu^{2}}{\sigma^{2}-\mu}\ \text{and}\ p=\frac{\mathrm{E}(X)}{\mathrm{Var}(X)}=\frac{\mu}{\sigma^{2}} \tag{4.37}$$

with the mean value μ and the standard deviation σ obtained from repeated measurements.

A particular property of the negative binomial distribution follows from Equation (4.38) if one calculates the relative standard deviation, i.e. the coefficient of variation σ/μ. One obtains

$$\frac{\sigma}{\mu} = \frac{\sqrt{\mathrm{Var}(X)}}{\mathrm{E}(X)} = \sqrt{\frac{1}{E(X)} + \frac{1}{r}} \tag{4.38}$$

The first term under the radix of Equation (4.38) describes the influence on the coefficient of variation caused by the uncertainty of the Poisson process. For a pure Poisson distribution without random influences, e.g. by sample treatment or instability of equipment, it holds

$$\frac{\sigma}{\mu} = \frac{\sqrt{\mathrm{Var}(X)}}{\mathrm{E}(X)} = \sqrt{\frac{1}{E(X)}} \tag{4.39}$$

The second term $1/r$ in Equation (4.38) quantifies the additional influences of sample preparations and instrument instabilities. $1/r = \vartheta^2$ is used in ISO 11929-1:2019 for the parameterization of uncertainties caused by known random influences.

4.6 The rectangular distribution: very limited information

If the only information regarding a quantity X is that it is behaving randomly between a lower limit $x_{\min}$ and an upper limit $x_{\max}$, the PME or the Principle of Indifference yields a rectangular distribution (Fig. 13) over the interval $[x_{\min}, x_{\max}]$

$$f_X(x) = \begin{cases} 1/(x_{\max} - x_{\min}), & x_{\min} \le x \le x_{\max} \\ 0 & \text{elsewhere} \end{cases} \tag{4.40}$$

with the expectation $\mathrm{E}(f_X(x)) = \dfrac{x_{\min} + x_{\max}}{2}$ and the variance $u^2(x) = (x_{\max} - x_{\min})^2 / 12$.

To sample from $\mathrm{R}(x_{\min}, x_{\max})$ make a draw r from the standard rectangular distribution $\mathrm{R}(0,1)$ and calculate $x = x_{\min} + (x_{\max} - x_{\min}) \cdot r$.

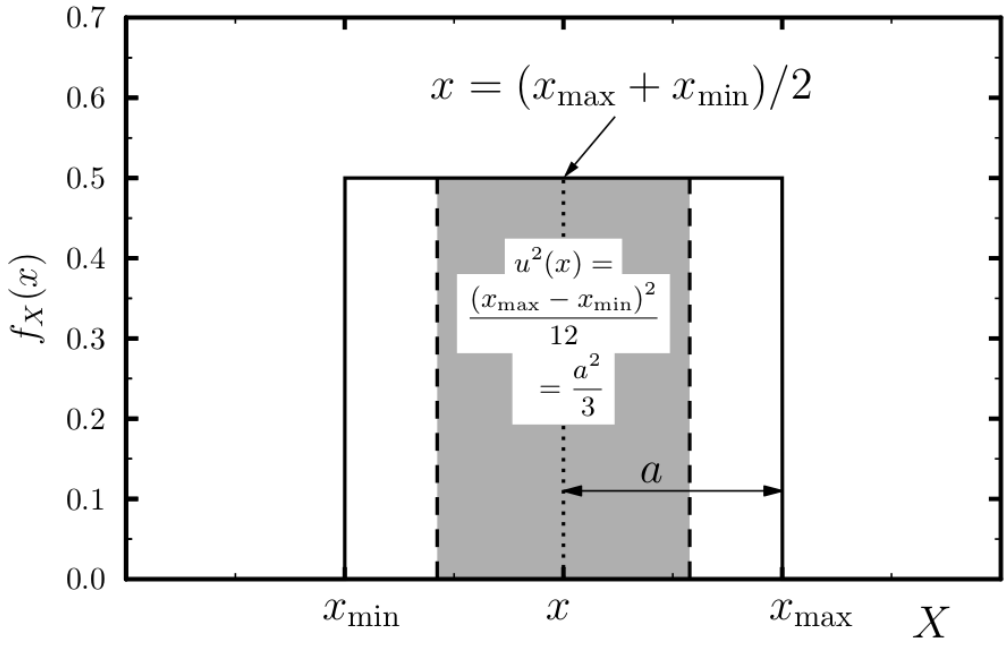


Fig. 13: Example of a rectangular distribution

4.7 The Beta distribution: efficiency measurements

If a series of n efficiency measurements in a counting experiment $(\varepsilon_i, i = 1,...,n)$ is available and if the prior information is used that the efficiency ε must fulfill $0 \le \varepsilon \le 1$, it can be shown (Ullrich and Xu, 2007) that the Bayes Theorem leads to a beta distribution $\mathrm{B}(\tilde{\varepsilon},\alpha,\beta)$ with the parameters α and β (not to be confused with the same symbols in the chapter on quantities and symbols as the PDF for the efficiency). The beta distribution (Fig. 14) is given by

$$f_\varepsilon\left(\tilde{\varepsilon}\middle|\varepsilon_i, i = 1,...,n\right) = \mathrm{B}(\tilde{\varepsilon},\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\cdot\tilde{\varepsilon}^{\alpha-1}\cdot(1-\tilde{\varepsilon})^{\beta-1} = \frac{1}{B(\alpha,\beta)}\cdot\tilde{\varepsilon}^{\alpha-1}\cdot(1-\tilde{\varepsilon})^{\beta-1} \tag{4.41}$$

with the expectation $\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n)) = \dfrac{\alpha}{\alpha+\beta}$ (4.42)

and the variance $\mathrm{Var}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n)) = \dfrac{\alpha\cdot\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ (4.43)

From this, the parameters of the Beta distribution are calculated as

$$\alpha = \left[\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n))\right]^2\cdot\left(\frac{1-\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n))}{\mathrm{Var}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n))} - \frac{1}{\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\varepsilon_i, i = 1,...,n))}\right) \tag{4.44}$$

and

$$\beta = \alpha \cdot \left(\frac{1}{\mathrm{E}(f_\varepsilon(\tilde{\varepsilon} \,|\, \varepsilon_i, i = 1,...,n))} - 1 \right) \tag{4.45}$$

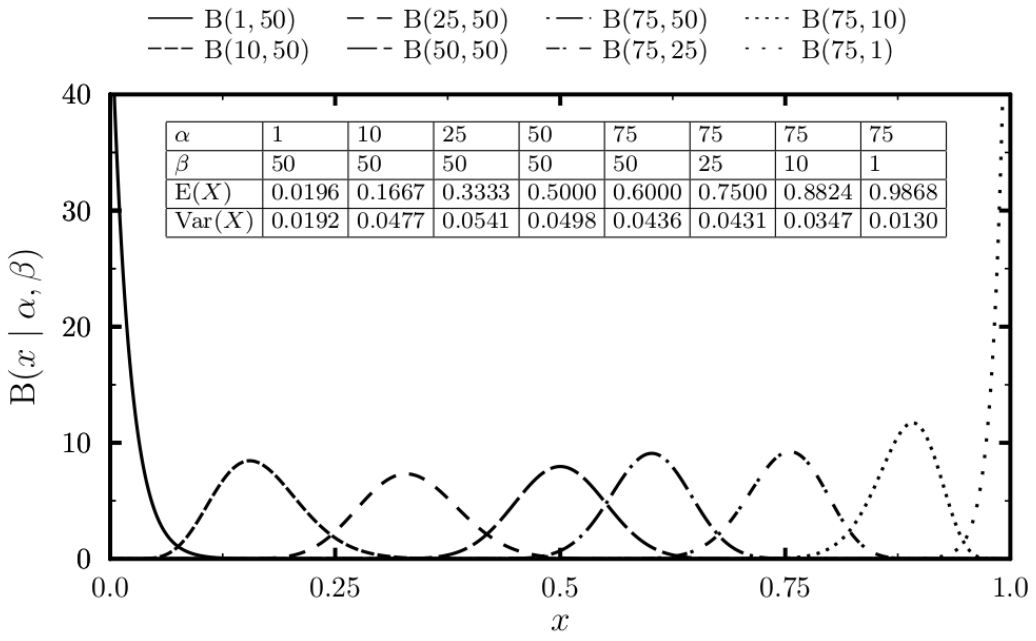
In the case of a single counting calibration measurement α corresponds approximately to the number of detected particles and β corresponds approximately to the number of missed particles (e. g. derived from a calibration standard or a second detector with known efficiency). So, in Equation (4.46) for a reasonable number *n* of events the detection efficiency is calculated and in Equation (4.47) the corresponding statistical variance. For low event numbers see the corrections derived in (Ullrich and Xu, 2007).

With the mean $\bar{\varepsilon} = \frac{1}{n-1} \sum_{i=1}^{n} \varepsilon_i$ and the squared standard deviation $s_\varepsilon^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2$ of the measured efficiency data the parameters of the Beta distribution are calculated by

$$\alpha = \bar{\varepsilon}^2 \left(\frac{1-\bar{\varepsilon}}{s_\varepsilon^2} - \frac{1}{\bar{\varepsilon}} \right) \text{ and } \beta = \alpha \cdot \left(\frac{1}{\bar{\varepsilon}} - 1 \right) \tag{4.46}$$

If the efficiency is not limited to the interval $[0,1]$ but to $[\varepsilon_\mathrm{L}, \varepsilon_\mathrm{U}]$ the Beta distribution can be expanded to the rescaled four parameter Beta distribution

$$f\left(\tilde{\varepsilon} \,|\, (\varepsilon_i, i = 1,...,n), \varepsilon_\mathrm{L} \le \tilde{\varepsilon} \le \varepsilon_\mathrm{U}\right) = \mathrm{B}(\tilde{\varepsilon}, \alpha, \beta, \varepsilon_\mathrm{L}, \varepsilon_\mathrm{U}) = \frac{1}{\varepsilon_\mathrm{U} - \varepsilon_\mathrm{L}} \cdot \mathrm{B}\left(\frac{\tilde{\varepsilon} - \varepsilon_\mathrm{L}}{\varepsilon_\mathrm{U} - \varepsilon_\mathrm{L}}, \alpha, \beta \right) \tag{4.47}$$



α	1	10	25	50	75	75	75	75
β	50	50	50	50	50	25	10	1
E(X)	0.0196	0.1667	0.3333	0.5000	0.6000	0.7500	0.8824	0.9868
Var(X)	0.0192	0.0477	0.0541	0.0498	0.0436	0.0431	0.0347	0.0130

Fig. 14: Examples of beta distributions.

4.8 The logarithmic normal distribution: multifactorial influences

The **logarithmic normal distribution** (log-normal for short) (Fig. 15) is the exact solution of the PME if the measurand is a multi-factorial quantity and the available information consists of a mean value $m(\ln(x))$ of the logarithms of measured values x and their standard deviation $s(\ln(x))$.

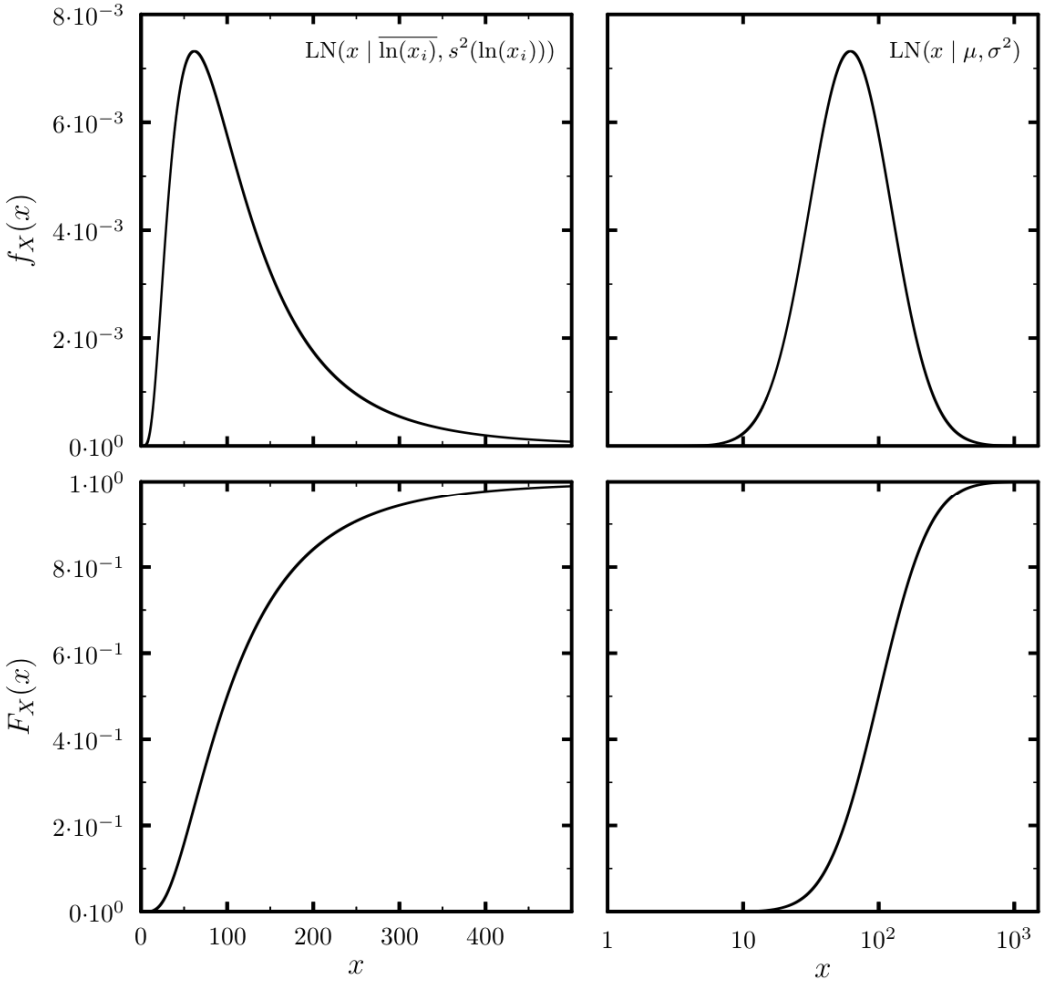


Fig. 15: Example of a logarithmic normal distribution: Densities (top) and distribution function of LN(ln(100);ln(2)^2) (bottom).

$$f_X(\tilde{x}\,|\,m(\ln(x)), s(\ln(x))) = \frac{1}{\sqrt{2\pi}\cdot\tilde{x}\cdot s(\ln(x))}\cdot\exp\left(\frac{-(\ln(\tilde{x})-m(\ln(x)))^2}{2\cdot s^2(\ln(x))}\right) \tag{4.48}$$

This log-normal distribution has the expectation

$$\begin{aligned}\mathrm{E}\left(f_X(\tilde{x}\,|\,m(\ln(x)), s(\ln(x)))\right) &= \int_0^\infty \xi\cdot f_X\left(\xi\,|\,m(\ln(x)), s(\ln(x))\right)\mathrm{d}\,\xi \\ &= \exp\left(m(\ln(x)) + \frac{s^2(\ln(x))}{2}\right)\end{aligned} \tag{4.49}$$

and the variance

$$\mathrm{Var}\left(f_X(\tilde{x}\,|m(\ln(x)),s(\ln(x)))\right)=\int_a^b(\xi-x)^2\cdot f_X(\xi\,|m(\ln(x)),s(\ln(x)))\,\mathrm{d}\,\xi= \\ =\exp\left(2\cdot m(\ln(x))+s^2(\ln(x))\right)\cdot\left(\exp(s^2(\ln(x)))-1\right) \quad . \quad (4.50)$$

For products of random variables holds the **Multiplicative Limiting Theorem of Statistics**: The product of *n* independent random variables with arbitrary distribution densities goes for a logarithmic normal distribution, if the number *n* of random variables goes for ∞.

4.9 Inverse probability density functions

The general model of evaluation according to Equation (20) of ISO 11929-1:2019 deals with counting measurements of ionizing-radiation events. Typically, a net count rate r_n shall be converted to the quantity activity $y=r_\mathrm{n}\cdot w$ by use of a calibration factor w.

The calibration factor itself is calculated from one or more input quantities $(X_i, i\geq 5)$ by $w=\frac{x_6\cdot x_8\cdots}{x_5\cdot x_7\cdots}$. For the evaluation of the calibration factor according to ISO 11929-2:2019, the free choice of arranging the PDF of an input quantity either in the numerator or in the denominator can have a significant impact on the resulting characteristic values due to the nonlinearity, typically if the relative uncertainty of the calibration factor $u_\mathrm{rel}(w)$ becomes larger than about 0,3. In the following this is elaborated for the example where the calibration depends on the detection efficiency ε only, i. e. $w=1/\varepsilon$.

Other nonlinearities than those caused by divisions may occur in the calibration factor, e. g. due to corrections of the inverse-square law with uncertain distance or Beer-Lambert's law with uncertain attenuation coefficient or distance. In such cases similar considerations apply.

This case occurs if either the efficiency ε or the calibration factor $w=\frac{1}{\varepsilon}$ is used and if a meaningful PDF shall be assigned to both quantities. In this case not the same PDF can be assigned to ε and w. The respective models of evaluation in the simplest form are $y=\left(r_\mathrm{g}-r_0\right)\cdot\frac{1}{\varepsilon}$ and $y=\left(r_\mathrm{g}-r_0\right)\cdot w$. Given an efficiency ε with the PDF $f_\varepsilon\left(\tilde{\varepsilon}\,|\,\mathfrak{I}\right)$ depending on the available information $\mathfrak{I}$, a transformation $w=g(\varepsilon)=\frac{1}{\varepsilon}$ is performed when shifting from the first model of evaluation $y=\left(r_\mathrm{g}-r_0\right)\cdot\frac{1}{\varepsilon}$ to the second one $y=\left(r_\mathrm{g}-r_0\right)\cdot w$ and $g^{-1}(w)=-\frac{1}{w}$ follows. Then the PDF $f_W\left(\tilde{w}\,|\,\mathfrak{I}\right)$ is calculated by

$$f_w(\tilde{w}\,|\,\mathfrak{I})=f_\varepsilon(g^{-1}(w)\,|\,\mathfrak{I})\cdot\left|\frac{d(g^{-1}(w))}{dw}\right|=f_\varepsilon\left(\frac{1}{\varepsilon}\,\middle|\,\mathfrak{I}\right)\cdot\frac{1}{w^2} \quad (4.51)$$

For large relative uncertainties of ε, the mean of w $\mathrm{E}(f_w(\tilde{w}|\Im)) = \mathrm{E}(1/f_\varepsilon(\tilde{\varepsilon}|\Im))$ can take values larger than $1/\mathrm{E}(f_\varepsilon(\varepsilon|\Im))$ and also the variance and quantiles can change significantly (see Fig. 16).

If normal distributions are assumed for the PDFs $f_w(\tilde{w}|\Im)$ or $f_\varepsilon(\tilde{\varepsilon}|\Im)$, the transformations $w = g(\varepsilon) = \frac{1}{\varepsilon}$ respectively $\varepsilon = g(w) = \frac{1}{w}$ result in just minor distortions of the alternative PDFs as long as $u_{\mathrm{rel}}(w) \leq 0{,}25$ respectively $u_{\mathrm{rel}}(\varepsilon) \leq 0{,}25$ holds. Results based on both PDFs will be in agreement and $\mathrm{E}(f_w(\tilde{w}|\Im)) \approx 1/\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\Im))$ applies. Consequently, it is possible to use ISO 11929-1:2019 and to determine the calibration factor w either in the numerator or in the denominator and the calibration can be done according to Annex A.1. The application of ISO 11929-1:2019 and ISO 11929-2:2019 will show just minor differences in the calculated characteristic limits.

For larger relative uncertainties significant differences the PDFs with important implications for the resulting characteristic limits occur. For very large relative uncertainties even the detection limit according to ISO 11929-1:2019 may not exist. The results obtained by ISO 11929-1:2019 and ISO 11929-2:2019 will differ significantly because of the important non-linearities of the model of evaluation.

If ISO 11929-2 is applied, the use of a normal distribution for detection efficiencies ε with large relative uncertainties should be avoided and it should be taken into account that ε is restricted to the interval $[0,1]$ or smaller.

The information that the detection efficiency ε is typically restricted to an interval $[\varepsilon_\mathrm{L}, \varepsilon_\mathrm{U}]$ can be taken into account by using the rectangular distribution if further information is not available. The choice, whether a quantity is in the nominator or the denominator of the calibration factor is arbitrary, as the reciprocal values of the interval boundaries could be used instead. It simply depends on whether the upper and lower bounds are given for which quantity, either ε or w, and to which of the quantities a rectangular PDF has to be assigned. For the other quantity the proper transformation of the PDF has to be used according to Equation (4.55).

According to Equation (4.52) the original rectangular distribution

$$f_\varepsilon(\tilde{\varepsilon}|\Im) = \begin{cases} 1/(\varepsilon_\mathrm{U} - \varepsilon_\mathrm{L}) & \text{for } \varepsilon_\mathrm{L} \leq \tilde{\varepsilon} \leq \varepsilon_\mathrm{U} \\ 0 & \text{otherwise} \end{cases} \tag{4.52}$$

with the expectation

$$\mathrm{E}(f_\varepsilon(\tilde{\varepsilon}|\Im)) = \frac{1}{2}(\varepsilon_\mathrm{U} + \varepsilon_\mathrm{L}) \tag{4.53}$$

and the variance

$$\mathrm{Var}(f_\varepsilon(\tilde{\varepsilon}|\Im)) = \frac{1}{12}(\varepsilon_\mathrm{U} - \varepsilon_\mathrm{L})^2 \tag{4.54}$$

transforms to the distribution

$$f_w(\tilde{w}|\Im) = \begin{cases} 1/\left\{\tilde{w}^2 \cdot (\varepsilon_U - \varepsilon_L)\right\} & \text{for } \varepsilon_U^{-1} \le \tilde{w} \le \varepsilon_L^{-1} \\ 0 & \text{otherwise} \end{cases} \tag{4.55}$$

with the expectation

$$\mathrm{E}(f_w(\tilde{w}|\Im)) = \frac{\ln(1/\varepsilon_L) - \ln(1/\varepsilon_U)}{\varepsilon_U - \varepsilon_L} \tag{4.56}$$

and the variance

$$\mathrm{Var}(f_w(\tilde{w}|\Im)) = \frac{1}{\varepsilon_L \cdot \varepsilon_U} - \left(\frac{\ln(1/\varepsilon_L) - \ln(1/\varepsilon_U)}{\varepsilon_U - \varepsilon_L}\right)^2 \tag{4.57}$$

Since the difference between both distributions is apparently significant and the choice between a numerator or denominator is arbitrary or determined by convention, the rectangular distribution is to be used with care in practice, despite its repeated use in the previous examples for the sake of convenience. Nevertheless, the benefit of the rectangular distribution either for the calibration factor or the efficiency is the warranted existence of the detection limit $y^{\#}$ due to the limited range of this distribution.

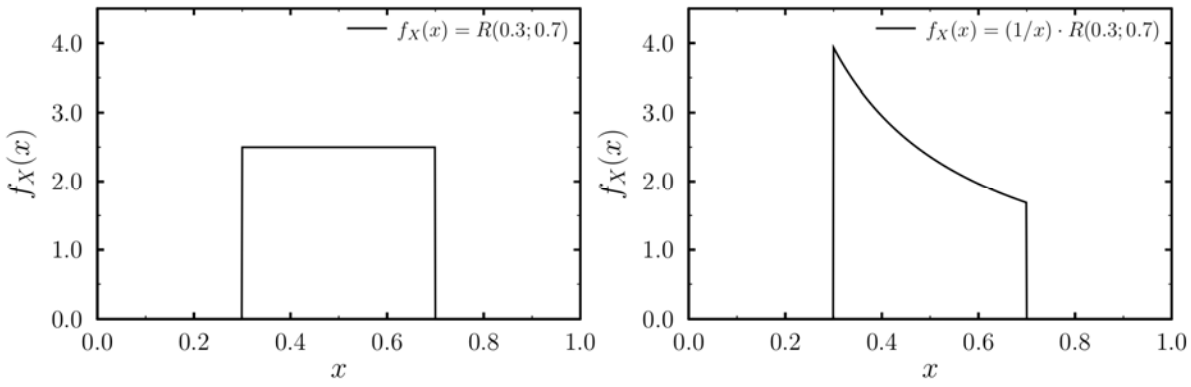


Fig. 16: Example for the problem of inverse distributions in the case of a rectangular PDF.

4.10 Functions of random variables

Is X a random variable, then $Y = g(X)$ is also a random variable. Is to each value x of the random variable X a value y of a random variable Y uniquely assigned and is the assignment described by a function $y = g(x)$, the following Equations hold

$$F_Y(y) = P(Y < y) = P(g(X) < y) \quad F_Y(y) = \int_{g(x)<y} f_X(x)\,\mathrm{d}x \tag{4.58}$$

If for a function $y = g(x)$ the inverse function $x = g^{-1}(y)$ exists, then the relation between $F_Y(y)$ and $F_X(x)$ can be given explicitly. One has, however, to distinguish, whether $g(x)$ increases or decreases monotonously. The following holds:

$$\begin{aligned} F_Y(y) &= F_X(x) \qquad \text{if } g'(x) \geq 0 \\ F_Y(y) &= 1 - F_X(x) \quad \text{if } g'(x) \leq 0 \end{aligned} \tag{4.59}$$

4.11 Propagating probability density functions

After establishing the probability distributions of all input quantities, a joint probability distribution, $f_{\mathbf{X}}(\tilde{\mathbf{x}}|\mathbf{a})$, has to be formed which in the case of independent input quantities is given as

$$f_{\mathbf{X}}(\tilde{\mathbf{x}}|\mathbf{a}) = \prod_{i=1}^{m} f_{X_i}(\tilde{x}_i|a_i) \tag{4.60}$$

with a_i being the subset of information available for X_i. A joint probability distribution has to be assigned to those X_i that are not independent and inserted in Equation (4.61) for the respective input quantities; see JCGM (2008b) for details.

The posterior probability distribution $f_Y(\tilde{y}|\mathbf{a})$ is calculated from the joint probability distribution $f_{\mathbf{X}}(\tilde{\mathbf{x}}|\mathbf{a})$ using the model Equation $Y = G(\mathbf{X})$ by the so-called Markov Formula

$$f_Y(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{+\infty} f_{\mathbf{X}}(\xi|\mathbf{a}) \cdot \delta(\tilde{y} - G(\xi))\, \mathrm{d}\xi\,. \tag{4.61}$$

The ISO/IEC Guide 98-3-1 recommends the application of Monte Carlo techniques to solve Equation (4.61) and to derive $f_Y(\tilde{y}|\mathbf{a})$. Suitable numeric is described in detail by Weise et al. (2009).

Using the Monte Carlo approach, $i = 1,...,n_{\mathrm{M}}$ Monte Carlo trials are performed for propagating the probability distributions by drawing sets $x_{1,i},...,x_{n,i}$ from the probability distributions $f_{X_i}(\tilde{x}_i|\mathbf{a})$. For each of these sets, one calculates $y_i = G(x_{1,i},...,x_{n,i})$. The vector $\boldsymbol{y}_{\mathrm{M}} = \{y_1,...,y_{n_{\mathrm{M}}}\}$, ordered ascendingly and afterwards assigning cumulative probabilities i/n_{M} to the y_i of $\boldsymbol{y}_{\mathrm{M}}$, is a discrete representation of the distribution function $F_Y(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{\tilde{y}} f_Y(\eta|\mathbf{a})\,\mathrm{d}\eta$ of Y.

In establishing $f_{\mathbf{X}}(\tilde{\mathbf{x}}|\mathbf{a})$, in performing the Monte Carlo trials and in calculating $f_Y(\tilde{y}|\mathbf{a})$ one does not take into account that the input quantities X_i and the measurand Y are non-negative. The fact that the measurand Y is non-negative is only taken into account when calculating the

limits of the coverage interval $y^{\triangleleft}, y^{\triangleright}$ or $y^{<}, y^{>}$ and the best estimate $\hat{y}$ and its associated standard uncertainty $u(\hat{y})$.

If the probability density $f_{\mathbf{X}}(\tilde{\mathbf{x}}|\mathbf{a})$ is explicitly available, the integral of Equation (4.61) can also have an explicit solution.

5 A Bayesian theory of uncertainty in measurement

5.1 General aspects

"Bayesian estimation", a reasoning method published in 1763 posthumously by Thomas Bayes (* 1702 † 1761) as an „Essay towards solving a problem in the doctrine of chances", means calculating the probability $f(A|B)$ of the validity of a proposition A on the basis of a prior estimate $f(A)$ of its probability and new relevant evidence B by using the so-called Bayes Theorem

$$f(A|B) \cdot f(B) = f(B|A) \cdot f(A) \tag{5.1}$$

The Bayes Theorem represents the natural form of learning from experience.

In metrology one has a **measurement result** y which is known and fixed[4]. The probability for a particular $\tilde{y}$ being the **true value**, which is unknown and unknowable, has to be determined. It is to emphasize here again that a statement about $\tilde{y}$ can only be derived on the basis of Bayesian statistics. In order to make probability statements about $\tilde{y}$ given y, one starts with a model providing a joint probability distribution for $\tilde{y}$ and y; see e.g. Gelman et al. (2014). The **joint probability density function (PDF)** of $\tilde{y}$ and y can be written as a product of two densities.

$$f_Y(\tilde{y}, y) = f_Y(\tilde{y}) \cdot f_Y(y|\tilde{y}) \tag{5.2}$$

$f_Y(\tilde{y})$ is called the **prior distribution** or **prior** for short comprising the information existing before a measurement is performed, and

$f_Y(y|\tilde{y})$ is called the **likelihood** and is the probability to obtain a measurement result y given a true value $\tilde{y}$.

Applying Bayes Theorem[5] to the right side of Equation (5.2) yields the posterior density

$$f_Y(\tilde{y}|y) = f_Y(\tilde{y}) \cdot f_Y(y|\tilde{y}) / f_Y(y) \text{ with } f_Y(y) = \int f_Y(\tilde{y}) \cdot f_Y(y|\tilde{y}) \, \mathrm{d}\tilde{y} \,. \tag{5.3}$$

[4] This applies also to repeated measurements where the individual results as well as the mean value and the standard deviation are given as fixed.

[5] Bayes Theorem: $f(A|B) \cdot f(B) = f(B|A) \cdot f(A)$. In this case the Bayes Theorem reads: $f_Y(\tilde{y}|y) \cdot f_Y(y) = f_Y(y|\tilde{y}) \cdot f_Y(\tilde{y})$

The probability for $\tilde{y}$ is the proposition to be updated from a prior estimate given the new relevant information represented by a measurement result y. The probability density function (PDF) $f_Y(\tilde{y}|y)$ according to Equation (5.3) is the updated PDF for $\tilde{y}$ given y and the prior information $f_Y(\tilde{y})$. It is the **posterior probability density function** or called **posterior** for short. It provides a complete description of the uncertainty regarding the true value $\tilde{y}$ of *Y* associated with y. It can be derived from the Bayes Theorem in equation (5.4).

$$f_Y(\tilde{y}|y) = \frac{f_Y(y|\tilde{y}) \cdot f_Y(\tilde{y})}{f_Y(y)} \tag{5.4}$$

$f_Y(y)$ is the probability to obtain a measurement result y which does not depend on $\tilde{y}$ and, with fixed *y*, can thus be considered a constant. A constant probability $f_Y(y)$ to obtain a result *y* means nothing more than that the measurement procedure yields reproducible results.

The assumption that $f_Y(y)$ is constant in metrology due to the requirement of a measurement procedure being reproducible does not hold for all applications of the Bayes Theorem in fields outside metrology. See for instance the example in chapter 3 of Bernardo (2003).

Therefore, in metrology the Bayes Theorem frequently is simply written in the form

$$f_Y(\tilde{y}|y) \propto f_Y(\tilde{y}) \cdot f_Y(y|\tilde{y}) \tag{5.5}$$

or

$$f_Y(\tilde{y}|y) = C \cdot f_Y(\tilde{y}) \cdot f_Y(y|\tilde{y}) \tag{5.6}$$

with C being a normalization constant.

Equation (5.5) is the basis of the **Bayesian theory of uncertainty** in measurement according to Weise and Wöger (1993) which is used in ISO 11929:2010 and ISO 11929:2019 (ISO 2019a, ISO 2119b, ISO 2019c, and ISO 2020).

There are two equivalent ways to determine $f_Y(\tilde{y}|y)$; either by determining the left side of Equation (5.5) using the Principle of Maximum (Information-) Entropy (PME) (Jaynes 1982, 1989) or by calculating the right side of Equation (5.5) thereby applying the Bayes Theorem (Bayes 1763). Both methods need additional information in order to solve the problem: the PME needs constraints and the Bayes Theorem needs information on the likelihood and a prior. Finally, the posterior $f_Y(\tilde{y}|y)$ has to be normalized.

The prior $f_Y(\tilde{y})$ considers all information available before the measurement is performed. If no particular prior information is at hand a so-called non-informative prior has to be applied. According to Jeffreys (1946) a non-informative prior should be invariant under translations. This is fulfilled by so-called Jeffreys prior with

$$f_Y(\tilde{y}) = C / \tilde{y} \quad (\tilde{y} \geq 0) \tag{5.7}$$

The **best estimate** $\hat{y}$ of the measurand *Y* is the expectation of $f_Y(\tilde{y}|y)$

$$\hat{y} = \mathrm{E}(f_Y(\tilde{y}|y)) = \int \tilde{y} \cdot f_Y(\tilde{y}|y)\,\mathrm{d}\tilde{y} \tag{5.8}$$

The **squared standard uncertainty** $u^2(\hat{y})$ associated with the best estimate $\hat{y}$ is the variance of $f_Y(\tilde{y}|y)$

$$u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}|y)) = \int (\tilde{y} - \hat{y})^2 \cdot f_Y(\tilde{y}|y)\,\mathrm{d}\tilde{y} \tag{5.9}$$

The VIM distinguishes two ways by which measurement uncertainties can be derived: type A and type B. Type A uncertainties are derived from repeated or counting measurements. Type B uncertainties result from other sources. For short we call them type A and type B uncertainties. But they are equivalent and the „types“ distinguish only the ways the uncertainties are obtained.

5.2 The methodology according to the GUM S1

The methodology of the GUM S1 is used in ISO 11929 by starting from a model of evaluation $Y = G(X_i, i = 1,...,n)$ connecting input quantities X_i mathematically to the measurand *Y*. It proceeds in two steps.

First, the **prior knowledge that the measurand is non-negative**[6] is separated from $f_Y(\tilde{y}|y)$ applying the product rule and describing this prior knowledge by the **Heaviside step function**

$$f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y}) \tag{5.10}$$

$$f_Y(\tilde{y}) = \mathrm{H}(\tilde{y}) = \begin{cases} \text{const} & (\tilde{y} \geq 0) \\ 0 & (\tilde{y} < 0) \end{cases} \tag{5.11}$$

$f_{Y,0}(\tilde{y}|y)$ is the PDF of a random variable Y_0 serving as an estimator of *Y*, which does not take into account the non-negativity of the measurand.

Second, the PDF $f_{Y,0}(\tilde{y}|y)$ is obtained by applying **Bayes Theorem** or the **Principle of Maximum (Information) Entropy**[7] (PME) (Jaynes 1982, 1989) in accordance with the GUM S1 which does not use any prior information or at least non-informative priors.

In other cases, the Bayes Theorem is applied in combination with a non-informative prior $f_{Y,0}(\tilde{y})$ or with an informative prior reflecting any other information available before a measurement is performed.

---

[6] In a normal evaluation of the measurement result y the non-negativity is not taken into account. Therefore, a result y can occasionally become negative.

[7] The Principle of Maximum Entropy (PME): $S = -\int f_{Y,0}\left(\tilde{y}|y\right) \cdot \ln\left(f_{Y,0}\left(\tilde{y}|y\right)\right)\mathrm{d}y = \max$

$$f_{Y,0}(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}) \cdot f_{Y,0}(y|\tilde{y}) \tag{5.12}$$

Given the same available information, the two approaches (PME and Bayes Theorem) are equivalent; see Weise et al. (2013, Appendix A.3) for a proof.

The **primary measurement result** or **primary estimate** y is the expectation of $f_{Y,0}(\tilde{y}|y)$, i.e. $y = \mathrm{E}(f_{Y,0}(\tilde{y}|y))$, and its squared associated **standard uncertainty** is the variance of $f_{Y,0}(\tilde{y}|y)$, i.e. $u^2(y) = \mathrm{Var}(f_{Y,0}(\tilde{y}|y))$.

ISO 11929-2:2019[8] is based on the GUM S1 approach. For a model of evaluation, $Y = G(X_i, i = 1,...,n)$ and a given set $\mathbf{a}$ of information, first the PDFs $f_{X_i}(\tilde{x}_i|\mathbf{a})$ of the input quantities have to be stablished. This is done by making use of the PME or the Bayes Theorem in combination with non-informative priors. The Markov Formula and Monte Carlo methods are used to propagate the distributions and to obtain the PDF $f_{Y,0}(\tilde{y}|\mathbf{a})$. The factorization according to Equation (5.5) is maintained. If additional prior information is available it can be taken into account using the methodology of ISO 11929- 2:2019.

The available information is summarized in ISO 11929-2:2019 by the symbol $\mathbf{a}$. It comprises all information about the measurement problem and the values of the input quantities and the measurand available when the evaluation of a measurement is performed. From this information $\mathbf{a}$ the PDFs $f_{X_i,0}(\tilde{x}_i|\mathbf{a}) = f_{X_i,0}(\tilde{x}_i|a_i)$ of the input quantities X_i have to be stablished (a_i being the subsets of information available and relevant for X_i).

The GUM S1 gives detailed advice how the PDFs have to be established. Special cases of obtaining PDFs $f_{X_i,0}(\tilde{x}_i|a_i)$ are dealt with in ISO 11929-1:2019 Appendix A for repeated counting measurements with random influences and in ISO 11929-2:2019 Appendix A for measurements with low count numbers.

After establishing the PDFs of all input quantities, a **joint PDF**, $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$, has to be formed which in case of independent input quantities is given as

$$f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a}) = \prod_1^m f_{X_i,0}(\tilde{x}_i|a_i) \tag{5.13}$$

A joint PDF has to be assigned to those X_i that are not independent and inserted in Equation (5.13) for the respective input quantities.

The **posterior** PDF $f_{Y,0}(\tilde{y}|\mathbf{a})$ is calculated from the joint PDF $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ using the model Equation $Y = G(X_i, i = 1,...,n)$ by the so-called **Markov-Formula**

[8] We start here with of ISO 11929-2:2019 because it provides the more general case described in the GUM S1. Historically, the GUM which provides the basis of ISO 11929-1:2019 was published before the GUM S1. However, the statistical basis of the GUM was only properly described in the GUM S1. Therefore, the GUM is presently under revision to make it more consistent with the statements in the GUM S1.

$$f_{Y,0}(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{+\infty} f_{\mathbf{X},0}(\xi|\mathbf{a}) \cdot \delta(\tilde{y} - G(\xi))\, \mathrm{d}\xi \tag{5.14}$$

The GUM S1 recommends the application of **Monte Carlo techniques** to solve Equation (5.14) and to derive $f_{Y,0}(\tilde{y}|\mathbf{a})$. Suitable numeric is described in detail elsewhere (Weise et al. 2013). Using the Monte Carlo approach, $i = 1,...,n_{\mathrm{M}}$ Monte Carlo trials are performed for propagating the distributions by drawing sets $x_{1,i},...,x_{n,i}$ from the PDFs $f_{X_i,0}(\tilde{x}_i|a_i)$. For each of these sets, one calculates $y_i = G(x_{1,i},...,x_{n,i})$. The vector $\mathbf{y}_{\mathrm{M}} = \{y_1,...,y_{n_{\mathrm{M}}}\}$, ordered ascendingly and afterwards assigning probabilities i / n_{M} to the y_i of $\mathbf{y}_{\mathrm{M}}$, is a discrete representation of the **distribution function of the posterior** $F_{Y,0}(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{\tilde{y}} f_{Y,0}(\eta|\mathbf{a})\, \mathrm{d}\eta$ of Y.

If the PDF $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ is explicitly available, the integral of Equation (5.14) may also have an explicit solution.

The **primary estimate** y of the measurand Y is calculated as the expectation of $f_{Y,0}(\tilde{y}|\mathbf{a})$

$$y = \mathrm{E}(f_{Y,0}(\tilde{y}|\mathbf{a})) = \int_{-\infty}^{+\infty} \eta \cdot f_{Y,0}(\eta|\mathbf{a})\, \mathrm{d}\eta \tag{5.15}$$

and the squared **standard uncertainty** associated with the primary estimate y is the variance of $f_{Y,0}(\tilde{y}|\mathbf{a})$

$$u^2(y) = \mathrm{Var}(f_{Y,0}(\tilde{y}|\mathbf{a})) = \int_{-\infty}^{+\infty} (\eta - y)^2 \cdot f_{Y,0}(\eta|\mathbf{a})\, \mathrm{d}\eta \tag{5.16}$$

The final result, i.e. the best estimate, is obtained by considering the non-negativity of the measurand by the Heaviside step function according to Equations (5.18) and (5.19)

$$f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y}) \tag{5.17}$$

This yields the **best estimate**

$$\hat{y} = \mathrm{E}(f_Y(\tilde{y}|\mathbf{a})) = \int_{0}^{+\infty} \eta \cdot f_Y(\eta|\mathbf{a})\, \mathrm{d}\eta \tag{5.18}$$

and its associated **standard uncertainty**

$$.u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}|\mathbf{a})) = \int_{0}^{+\infty} (\eta - \hat{y})^2 \cdot f_Y(\eta|\mathbf{a})\, \mathrm{d}\eta \tag{5.19}$$

The characteristic limits are derived as quantiles of the suitable PDFs as described in detail in ISO 11929-2:2019.

5.3 The GUM: an approximate and minimalistic concept

The Bayesian theory of measurement uncertainties (Weise and Wöger 1993), which provides a basis of the GUM approach, factorizes the desired PDF $f_Y(\tilde{y}|y)$

$$f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y}) \tag{5.20}$$

with

$$f_Y(\tilde{y}) = \mathrm{H}(\tilde{y}) = \begin{cases} \text{const} & (\tilde{y} \geq 0) \\ 0 & (\tilde{y} < 0) \end{cases} \tag{5.21}$$

and derives $f_{Y,0}(\tilde{y}|y)$ by PME (Jaynes 1982) or the Bayes' Theorem. Again, one assumes as the only prior information that the measurand Y is non-negative.

If only y and $u(y)$ are known, they are the best estimate and its associated standard uncertainty. Thus, they give for the application of the PME the constraints, $y = \mathrm{E}\left(f_{Y,0}(\tilde{y}|y)\right)$ and $u^2(y) = \mathrm{Var}\left(f_{Y,0}(\tilde{y}|y)\right)$. The PME leads with these constraints y and $u(y)$ to the searched PDF $f_{Y,0}(\tilde{y}|y)$ by means of variational methods and Lagrange multiplicators and yields the solution $f_{Y,0}(\tilde{y}|y) = C \cdot \exp\left(-(\tilde{y}-y)^2 / (2 \cdot u^2(y))\right)$ and thus

$$f_Y(\tilde{y}|y) = C \cdot f_Y(\tilde{y}) \cdot \exp\left(-(\tilde{y}-y)^2 / (2 \cdot u^2(y))\right) \tag{5.22}$$

The Gaussian distribution in Equation (5.22) is neither an approximation nor a probability distribution from repeated or counting measurements. It is based on the information about the measurand in terms of constraints on the PME. This is different from the frequentist point of view which allows only for likelihoods based on repeated measurements.

If by turning the argument one assumes that only a true value $\tilde{y}$ and its associated standard uncertainty $\tilde{u}(\tilde{y})$ are known, one obtains the constraints $\tilde{y} = \mathrm{E}\left(f_Y(y|\tilde{y})\right)$ and $\tilde{u}^2(\tilde{y}) = \mathrm{Var}\left(f_Y(y|\tilde{y})\right)$ which yield with the PME $S = -\int f_Y(y|\tilde{y}) \cdot \ln\left(f_Y(y|\tilde{y})\right) \mathrm{d}y = \max$ the solution

$$f_Y(y|\tilde{y}) = C \cdot \exp\left(-(\tilde{y}-y)^2 / (2 \cdot \tilde{u}^2(\tilde{y}))\right) \tag{5.23}$$

Again, this is neither an approximation nor a probability distribution from repeated or counting measurements.

The GUM and ISO 11929:2010 and ISO 11929-1:2019 are minimalistic for the purpose of general applicability and therefore assume that only y and $u(y)$ are known. This leads to the Gaussian PDF $f_Y(\tilde{y}|y)$ in Equation (5.22). The PDF describing the prior knowledge is also

minimalistic, namely it is only assumed that the measurand is non-negative. The knowledge $\tilde{y} \geq 0$ is then taken into account by a Heaviside function $\mathrm{H}(y)$ as PDF.

It must be emphasized that the user is free to take into account more information, if it is available. Then, one has to follow the GUM S1 approach and to use the tools provided by the PME, the Product Rule, and the Bayes Theorem for establishing, updating and propagating distributions.

ISO 11929-1:2019 is based on the GUM approach and only makes use of the factorization of Equation (5.20) and of expectations and variances of suitable PDFs $f_{Y,0}(\tilde{y}|y)$ without specifying the underlying PDFs. Thus, ISO 11929-1:2019 is a special case of the methodology of ISO 11929-2:2019. It is assumed that the information available **a** consists only of the primary measurement results x_i of X_i and the standard uncertainties $u(x_i)$ associated with the x_i.

For those input quantities X_i which are count rates the PDFs $f_{X_i,0}(\tilde{x}_i|x_i)$ are Gamma distributions. In ISO 11929-1:2019 the GUM approach with Gaussian PDFs is used and the variances of the PDFs of the count rates are taken into account.

The primary measurement result y and its associated standard uncertainty $u(y)$ are obtained according to Equations (5.22).

The final result is obtained by considering the non-negativity of the measurand by the Heaviside step function according to Equations (5.24) to (5.26)

$$f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot \exp\left(-(\tilde{y}-y)^2 / 2 \cdot u^2(y)\right) \cdot \mathrm{H}(\tilde{y}) \tag{5.24}$$

then yields the **best estimate**

$$\hat{y} = \mathrm{E}(f_Y(\tilde{y}|\mathbf{a})) = \int_{-\infty}^{+\infty} \eta \cdot f_Y(\eta|\mathbf{a})\,\mathrm{d}\eta \tag{5.25}$$

and its associated **standard uncertainty**

$$u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}|\mathbf{a})) = \int_{-\infty}^{+\infty} (\eta - y)^2 \cdot f_Y(\eta|\mathbf{a})\,\mathrm{d}\eta \tag{5.26}$$

The characteristic limits are calculated as quantiles of the suitable PDFs using the quantiles of the standard normal distribution.

In spite of its widespread and increasing use, the applicability of the GUM is limited. It is an approximation. The standard uncertainties are the result of a Taylor expansion of the model of evaluation truncated after the linear term. In principle, it should only be applied to models which are linear or which at least can be sufficiently linearized. If the GUM approximation is not sufficient for the model of evaluation the GUM offers a second order approximation which, however, is extremely unhandy and moreover has its own limitations. In this case, the approach of the GUM S1 helps to solve the problem in general. The GUM is actually a special case contained in the more general approach of the GUM S1 which widely extends the applicability of the GUM methodology.

5.4 Calculation of uncertainties in measurement according to the GUM

Given a model of evaluation $Y = G(X_i; i = 1,...,n)$ an estimate y of the output quantity Y is calculated by $y = G(x_i; i = 1,...,n)$. From the estimates x_i of the input quantities X_i the combined standard uncertainty $u(y)$ associated with y is calculated using the covariances $u(x_i, x_j)$

$$u^2(y) = \sum_{i,j=1}^{n} \frac{\partial G}{\partial X_i} \frac{\partial G}{\partial X_j} u(x_i, x_j) \ , \tag{5.27}$$

$$u^2(y) = \sum_{i=1}^{n} c_i^2 \cdot u^2(x_i) + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_i \cdot c_j \cdot u(x_i) \cdot u(x_j) \cdot r(x_i, x_j) \tag{5.28}$$

which in the case of independent input quantities X_i yields the well-known Equation

$$u^2(y) = \sum_{i=1}^{n} \left(\left. \frac{\partial G}{\partial X_i} \right|_{x_1,...,x_n} \right)^2 \cdot u^2(x_i) = \sum_{i=1}^{n} c_i^2 \cdot u^2(x_i) \tag{5.29}$$

with

$$c_i \equiv \partial G / \partial X_i \tag{5.30}$$

If the partial derivatives are not explicitly available, they can be numerically approximated sufficiently by differential quotients making use of the standard uncertainty $u(x_k)$ as an increment of the x_k.

$$\frac{\partial G_k}{\partial x_i} = \frac{1}{u(x_i)} \cdot \left(G_k\left(x_1,...,x_i + u(x_i)/2,...,x_n\right) - G_k\left(x_1,...,x_i - u(x_i)/2,...,x_n\right) \right) \tag{5.31}$$

In many instances, a simple calculus for independent quantities is helpful.

Sums of quantities

$$y = \sum_{i=1}^{n} (\pm x_i) \Rightarrow u^2(y) = \sum_{i=1}^{n} u^2(x_i) \tag{5.32}$$

Relative uncertainty

$$u_{\text{rel}}(x) = u(x) / x \tag{5.33}$$

Product of quantities with $m < n$

$$y = \prod_{i=1}^{m} x_i \Big/ \prod_{j=m+1}^{n} x_i \Rightarrow u_{\text{rel}}^2(y) = \sum_{i=1}^{m} u_{\text{rel}}^2(x_i) + \sum_{j=m+1}^{n} u_{\text{rel}}^2(x_j) = \sum_{i=1}^{n} u_{\text{rel}}^2(x_i) \tag{5.34}$$

Throughout this equation, modify the summation over *i* to start at 1 and end at *m* and the summation over *j* to start at *m*+1 and end at *n*. Moreover, replace *z* by *x* and introduce the additional constraint that *m*<*n*.

Empirical correlation

$$s(x_i, x_j) = \frac{1}{n-1} \cdot \sum_{k=1}^{n} (x_{i,k} - \bar{x}_i) \cdot (x_{j,k} - \bar{x}_j) \tag{5.35}$$

Correlation coefficient

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i) \cdot u(x_j)} \tag{5.36}$$

Explanations of and rules to handle covariances are given in ISO 11929-3:2019, Annex A.

5.5 Calculation of uncertainties for counting measurements according to the GUM

In order to clarify this procedure by a simple example, assume e.g. a counting measurement with a model of evaluation

$$y = (r_g - r_0) \cdot w = r_n \cdot w = (n_g / t_g - n_0 / t_0) \cdot w \tag{5.37}$$

and an associated standard uncertainty

$$\begin{aligned} u^2(y) &= (n_g / t_g - n_0 / t_0)^2 \cdot u^2(w) + w^2 \cdot (u^2(n_g) / t_g^2 + u^2(n_0) / t_0^2) \\ &= (n_g / t_g - n_0 / t_0)^2 \cdot u^2(w) + w^2 \cdot (n_g / t_g^2 + n_0 / t_0^2) \end{aligned} \tag{5.38}$$

Equation (5.37) is a general model in which a net signal, calculated by subtraction of a background or blank signal from a gross signal, is multiplied by a calibration factor to calculate the estimate of the measurand. Since no further assumption about the measurement is made it demonstrates the general applicability of the concept of ISO 11929:2010 also beyond nuclear radiation measurements. A more general form of the above model is given by:

$$y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} \text{ with } x_1 = r_g = n_g / t_g \text{ and } x_2 = r_0 = n_0 / t_0 \tag{5.39}$$

Equation (5.39) is a very common model of evaluation in nuclear radiation measurements. The particular is that $x_1 = r_g = n_g / t_g$ and $x_2 = r_0 = n_0 / t_0$ are the gross and the background count rates, respectively. x_3 is shielding factor, x_4 a general background term and $w = \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots}$ a calibration factor. With $x_3 = 1$, $u(x_3) = 0$, $x_4 = 0$, and $u(x_4) = 0$ the model of Equation (5.39) also contains the simplest model according to Equation (5.37).

Knowing that the counts n_g and n_0 are obtained from a Poisson process allows quantifying the uncertainty as $u^2(x_1) = n_g / t_0^2$ and $u^2(x_2) = n_0 / t_0^2$, respectively. Thus, one obtains the standard uncertainty

$$\begin{aligned} u^2(y) &= w^2 \cdot (u^2(x_1) + x_3^2 u^2(x_2) + x_2^2 u^2(x_3) + u^2(x_4)) + y^2 u_{\mathrm{rel}}^2(w) \\ &= w^2 \cdot (r_g / t_g + x_3^2 r_0 / t_0 + r_0^2 u^2(x_3) + u^2(x_4)) + y^2 u_{\mathrm{rel}}^2(w) \end{aligned} \tag{5.40}$$

with $u_{\mathrm{rel}}^2(w) = \sum_{i=5}^{m} \frac{u^2(x_i)}{x_i^2}$ being the relative uncertainty of the calibration factor w.

5.6 Calculation of uncertainties for counting measurements with random influences according to the GUM

5.6.1 Repeated measurements and random influences

Random influences due to, for instance, sample treatment and instruments cause measurement deviations, which can be different from sample to sample. In such cases, the counting results n_i of the counting measurements on several samples of a radioactive material to be examined, on several blanks of a radioactively labelled blank material, and on several reference samples of a standard reference material are therefore respectively averaged to obtain suitable estimates x_1 and x_2 of the input quantities X_1 and X_2 and their associated standard uncertainties $u(x_1)$ and $u(x_2)$, respectively; see Equation (5.39). Accordingly, X_1 shall be considered as the mean gross count rate and X_2 as the mean background count rate. Therefore, the measur- and Y shall also be taken as an averaged quantity, for instance as the mean net count rate or mean activity of the samples. All symbols belonging to the countings on the samples, blanks and reference samples are marked by the subscripts g, 0 and r, respectively. In each case, arithmetic averaging over m countings of the same kind carried out with the same preselected measurement duration, t (time preselection), is denoted by an overline. For m counting results $(i = 1, ..., m;\ m > 1)$ which are obtained in such a way and shall be averaged, the mean value $\overline{n}$ and its uncertainty $u^2(\overline{n})$ of the values n_i are given by

$$\overline{n} = \frac{1}{m} \sum_{i=1}^{m} n_i \qquad u^2(\overline{n}) = \frac{1}{m} \left(\overline{n} + \frac{m-1}{m-3} \overline{n} + \frac{1}{m-3} \sum_{i=1}^{m} (n_i - \overline{n})^2 \right) \tag{5.41}$$

Obviously, $m > 3$ is required. A derivation of equation (5.41) was given by Weise et al. (2013).

5.6.2 Procedure with unknown influences

In the case of unknown influences, the following expressions are valid for the mean gross count rate X_1 and the mean background count rate X_2

$$x_1 = \overline{n}_g / t_g \qquad x_2 = \overline{n}_0 / t_0 \tag{5.42}$$

$$u^2(x_1) = u^2(\overline{n}_g) / t_g^2 \qquad u^2(x_2) = u^2(\overline{n}_0) / t_0^2 \tag{5.43}$$

With the approach according to Equations (5.39) and (5.40) one obtains

$$y = \left(\frac{\overline{n}_g}{t_g} - \frac{\overline{n}_0}{t_0} x_3 - x_4 \right) \cdot w \tag{5.44}$$

$$u(y) = \sqrt{w^2 \cdot \left[u^2(\overline{n}_g) / t_g^2 + x_3^2 u^2(\overline{n}_0) / t_0^2 + (\overline{n}_0 / t_0)^2 \cdot u^2(x_3) + u^2(x_4) \right] + y^2 u_{rel}^2(w)} \tag{5.45}$$

5.6.3 Procedure with known influences

Another procedure, appropriate when small random influences are present, is based on the approach

$$u^2(\overline{n}) = \frac{\overline{n} + u^2}{m} = \frac{\overline{n} + \vartheta^2 \overline{n}^2}{m} \text{ or } \vartheta = u / \overline{n} \tag{5.46}$$

The symbol u denotes here the two terms with $m-3$ in Equation (5.41), i.e. $u^2 = \frac{m-1}{m-3}\overline{n} + \frac{1}{m-3}\sum_{i=1}^{m}(n_i - \overline{n})^2$.

The first term, $\overline{n}$, of Equation (5.46) corresponds to the numbers n_i of pulses according to the Poisson law in the absence of random influences. These influences are described by the second term, $\vartheta^2 \overline{n}^2$, assuming an empirical relative standard deviation ϑ valid for all samples and countings and caused by these influences. This influence parameter ϑ can be calculated from the data of counting measurements of the reference samples by combining Equation (5.43) with Equation (5.46)

$$\vartheta^2 = (m \cdot u^2(\overline{n}_r) - \overline{n}_r) / \overline{n}_r^2 \tag{5.47}$$

Instead of the data from counting measurements of the reference samples, those for other samples can be used which were previously examined, not explicitly for reference purposes but under conditions similar to those of the reference samples.

If $\vartheta^2 < 0$ results, the approach and the data are not compatible. The number m_r of the reference samples should then be enlarged or $\vartheta = 0$ be set. Moreover, $\vartheta < 0,2$ should be obtained.

NOTE The influence parameter ϑ is assumed to apply to both, the gross and the background measurements.

Instead of Equation (5.43), the expressions

$$u^2(x_1) = (\overline{n}_g + \vartheta^2 \overline{n}_g^2) / (m_g t_g^2) \quad u^2(x_2) = (\overline{n}_0 + \vartheta^2 \overline{n}_0^2) / (m_0 t_0^2) \tag{5.48}$$

now apply. The cases $m_g = 1$ and $m_0 = 1$ are permitted here. With $x_1 = \overline{n}_g / t_g$ and Equation (5.48), $u^2(x_1)$ is given as a function of x_1 by

$$u^2(x_1) = (x_1 / t_g + \vartheta^2 x_1^2) / m_g \tag{5.49}$$

With $x_3 = 1$ with $u(x_3) = 0$ and $x_4 = 0$ with $u(x_4) = 0$ it follows that

$$u(y) = \sqrt{w^2 \cdot \left[(\overline{n}_g + \vartheta^2 \overline{n}_g^2) / (m_g t_g^2) + (\overline{n}_0 + \vartheta^2 \overline{n}_0^2) / (m_0 t_0^2) \right] + y^2 u_{rel}^2(w)} \tag{5.50}$$

5.7 Calculation of uncertainties in measurements according to the GUM S1

After establishing the PDFs of all input quantities, a **joint PDF** $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ has to be formed which in the case of independent input quantities is given as

$$f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a}) = \prod_1^m f_{X_i,0}(\tilde{x}_i|a_i) \tag{5.51}$$

A joint PDF has to be assigned to those X_i that are not independent and inserted in Equation (5.51) for the respective input quantities.

The **posterior** PDF $f_{Y,0}(\tilde{y}|\mathbf{a})$ is calculated from the joint PDF $f_{X,0}(\tilde{\mathrm{x}}|\mathbf{a})$ using the model Equation $Y = G\left(X_i, i = 1,...,n\right)$ and the so-called **Markov Formula**

$$f_{Y,0}(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{+\infty} f_{\mathbf{X},0}(\xi|\mathbf{a}) \cdot \delta(\tilde{y} - G(\xi))\,\mathrm{d}\xi \tag{5.52}$$

The **primary result** is then calculated as $y = \mathrm{E}\left(f_{Y,0}(\tilde{y}|\mathbf{a})\right)$ and its associated standard uncertainty as $u^2(y) = \mathrm{Var}\left(f_{Y,0}(\tilde{y}|\mathbf{a})\right)$.

The **best estimate** is $\hat{y} = \mathrm{E}\left(f_Y(\tilde{y}|\mathbf{a},\Im)\right) = \mathrm{E}\left(f_Y(\tilde{y}|\mathbf{a}) \cdot f_Y(\tilde{y}|\Im)\right)$ and its associated standard uncertainty $u^2(\hat{y}) = \mathrm{Var}\left(f_Y(\tilde{y}|\mathbf{a},\Im)\right) = \mathrm{Var}\left(f_Y(\tilde{y}|\mathbf{a}) \cdot f_Y(\tilde{y}|\Im)\right)$.

6 Characteristic limits according to ISO 11929

6.1 The history of ISO 11929

The decision problem dealt with in ISO 11929 arises from the fact that in nuclear radiation measurements, the radiation of a sample of radioactive material has to be measured in the presence of a natural background radiation. It is the same problem as faced in trace element analysis, where the concentration of an element, a compound or a radionuclide has to be determined in the presence of an analytical blank. As a consequence of the natural background radiation, of analytical blanks and of measurement uncertainties the detection capability of a measurement is downward limited.

There are three questions to be answered:

- Is there a contribution from the sample among the counted events?
- How large is the smallest true value of the measurand which can be detected with high reliability?
- If a contribution of the sample has been observed, how large is the range of values of the measurand containing the true value with high probability?

The three questions are answered by the concept of the characteristic limits:

– the **decision threshold** decides the question whether there is a contribution from the sample among the counted events.
– the **detection limit** is the smallest true value of the measurand which can be detected with high reliability.
– the **limits of the coverage interval** define an interval which contains the true value of the measurand with a pre-selected probability.

The term "**characteristic limits"** is used in ISO 11929 to summarily denote the decision threshold, the detection limit and the limits of a coverage interval. ISO 11929 provides stipulations for the calculation of characteristic limits and values, i.e. the decision threshold, the detection limit, the limits of coverage intervals and the best estimate and its associated standard uncertainty.

The development of ISO 11929 started in the end of 1981 when a working group (AK SIGMA) of the German-Swiss Association for Radiation Protection (Deutsch-Schweizerischer Fachverband für Strahlenschutz e.V.) was established which should standardize the calculation of the characteristic limits. The work was based on concepts developed by Currie (1968), Nicholson (1963) and Altschuler und Pasternak (1963) who used statistical hypothesis testing of the equality of gross and background counting events. All these concepts made use of frequentist statistics.

It took until the year 1989 that the work of the AK SIGMA resulted in the publication of the first part of the German standard DIN 25482. In the year 1997 there were already 7 parts of DIN 25482 available, all using frequentist statistics and dealing exclusively with Type A uncertainties (Table 1).

Between 2000 and 2003 the Parts 10 – 13 of DIN 25482 were published which now were based on a Bayesian Theory of Uncertainty in Measurement (Weise and Wöger, 1992, 1993) and allowed to take into account all sources of uncertainty (Type A and Type B). A survey on the historical development of the standards based on Bayesian statistics is given in Table 2.

In parallel to the work of the AK SIGMA, the concepts of DIN 25482 were used to develop the first parts of ISO 11929 since the mid of the 1990ties. This work was done in the Working Group 17 (Radioactivity Measurements) of ISO/TC85 (Nuclear Energy)/SC2 (Radiation Protection). Between 2000 and 2005 eight parts of ISO 11929 were published, the Parts 1 – 6 and 8 being based on frequentist statistics and Part 7 on Bayesian statistics. To end this confusion and to get a consistent standard for the calculation of the characteristic limits, ISO 11929:2010 was developed which combined all topics of the earlier parts. ISO 11929:2010 was exclusively based on the GUM.

Table 1: Out-dated ISO 11929 and DIN 25482 standards based on frequentist statistics.

ISO 11929	**DIN 25482**	**Determination of detection limit and decision threshold for ionizing radiation measurements**
ISO 11929- 1:2000	DIN 25482-1:1989	Fundamentals and application to counting measurements without the influence of sample treatment
-	DIN 25482-1 Supplement:1992	Examples for the application

ISO 11929	DIN 25482	Determination of detection limit and decision threshold for ionizing radiation measurements
-	DIN 25482-2:1992	Counting spectrometric measurements without the influence of sample treatment
-	DIN 25482-2 Supplement:1992	Examples for the application
ISO 11929-2:2000	DIN 25482-6:1993	Fundamentals and applications to counting measurements with the influence of sample treatment
-	DIN 25482-6:1993 Supplement:1998	Examples for the application
-	DIN 25482-4:1996	Counting alpha-spectrometric measurements without the influence of sample treatment
ISO 11929-3:2000	DIN 25482-5:1993	Fundamentals and application to counting measurements by high resolution gamma spectrometry, without the influence of sample treatment
-	DIN 25482-5:1993 Supplement:1997	Examples for the application
ISO 11929-4:2001	DIN 25482-3:1993	Fundamentals and application to measurements by use of linear scale analogue ratemeters, without the influence of sample treatment
ISO 11929-5:2005	-	Fundamentals and Application to a Transient Measurement Mode
ISO 11929-6:2005	DIN 25482-7:1997	Fundamentals and application to measurements of aerosols and liquid effluents while running

With ISO 11929:2010 the earlier versions, Parts 1 – 8, of ISO 11929 were withdrawn. Except for the earlier Part 7, they had been based on frequentist statistics and could not take into account type B uncertainties. ISO 11929:2010 was also adopted as DIN ISO 11929:2011 and the earlier standard DIN 25482 Parts 1 – 8 and 10 – 13 were withdrawn. DIN 25482 Parts 1 – 8 had also been based on frequentist statistics.

ISO 11929 makes use of a Bayesian theory of measurement uncertainty according to Weise and Wöger (1992, 1993, 1999). The concept of Bayesian characteristic limits was introduced by Weise (1997) and later extended to spectrometric measurements (Weise and Michel 1995). The theoretical background of ISO 11929:2010 was described in detail by Weise et al. (2004, 2005, 2006). Since ISO 11929:2010 was dealing exclusively with measurement uncertainties as stipulated in the GUM (ISO 1993, JCGM 2008a). The publication of the GUM Supplement 1, (JCGM 2008b) required an extension of the methodology of ISO 11929:2010 (ISO 2010) and consequently a revision of ISO 11929:2010. An extension of ISO 11929 to the methodology of the GUM S1 was made by Weise et al. (2009) and material for the revision of ISO 11929:2010 was collected and published (Weise et al. 2013). A particular problem of a Bayesian analysis of ratemeter measurements was solved by Weise (2004).

Explanatory papers for ISO 11929 were published by Michel (2000, 2016, 2017, 2019) including an application of ISO 11929:2010 to conformity assessments (Michel 2017; SSK 2016). A description of the early historical development of ISO 11929 as well of its German predecessor DIN 25482 may be found elsewhere (Michel and Kirchhoff, 1999).

Table 2: ISO 11929, DIN 25482, and DIN ISO 11929 standards based on Bayesian statistics.

ISO 11929	**DIN 25482 resp. DIN ISO 11929**	**Determination of the detection limit and decision threshold for ionizing radiation measurements**
ISO 11929-7:2005	DIN 25482-10:2000	Fundamentals and general applications
-	DIN 25482-11	Measurements with albedo dosemeters
ISO 11929-8:2005	DIN 25482-12	Fundamentals and application to unfolding of spectrometric measurements without the influence of sample treatment
-	DIN 25482-13:2003	Fundamentals and Application to a Transient Measurement Mode
-	DIN 25482-13:2003 Supplement 1:2014	Supplement 1: Examples
ISO 11929:2010	DIN ISO 11929:2011	Determination of the characteristic limits (decision threshold, detection limit and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and application
-	DIN ISO 11929:2011 Supplement 1:2014	Supplement 1: Examples
ISO 11929 Parts 1 - 4	**DIN ISO 11929-1 to -4**	**Determination of the characteristic limits (decision threshold, detection limit and limits of the coverage interval) for measurements of ionizing radiation - Fundamentals and application**
ISO 11929-1:2019	DIN ISO 11929-1:2020	Elementary applications
ISO 11929-2:2019	DIN ISO 11929-2:2020	Advanced applications
ISO 11929-3:2019	DIN ISO 11929-3:2020	Applications to unfolding methods
ISO 11929-4:2020	DIN ISO 11929-4:2021	Guidelines to applications
EN ISO 11929-1:2021	DIN EN ISO 11929-1:2021	Elementary applications
EN ISO 11929-2:2021	DIN EN ISO 11929-2:2021	Advanced applications
EN ISO 11929-3:2021	DIN EN ISO 11929-3:2021	Applications to unfolding methods
EN ISO 11929-4	DIN EN ISO 11929-4	Guidelines to applications (in preparation)

Though the history of the application of ISO 11929:2010 can be considered a success and the basics of this standard series remained unchanged, there were quite some needs to revise ISO 11929:2010. The following items have been considered in the revision as necessary improvements:

- Fix for the problem of a not-existent detection limit by application of GUM S1.
- Extension to general applicability of ISO 11929 to GUM Suppl. 1.
- Correction of an error in the stipulations for random influences (Weise et al. 2013).
- Correction of the wording for proper use of Bayesian terminology, e.g. “decision making” instead of “hypothesis testing”, “coverage interval” instead of “confidence interval”.

- Clarification of stipulations for assessments and documentation in order to prevent misuse of stipulations.
- Extension of the standard to dosimetric measurements.
- Allowance for the shortest coverage interval as an alternative to the probabilistically symmetric one.
- Facilitation of the application of the standard by more explicit formulas, explanations and explicit examples.

By the revision, ISO 11929 was divided into 4 parts, of which parts 1 to 3 were published in 2019, the 4th part is presently under FDIS voting and hopefully, will be published in 2022. The first three parts of the standard were also published as a German DIN ISO standard in 2020.

- ISO 11929-1:2019 corresponds closely to the old ISO 11929:2010 and is exclusively based on the GUM.

- ISO 11929-2:2019 covers the application of the GUM S1 and allows calculating the characteristic limits using Monte Carlo methods.

- ISO 11929-3:2019 deals with unfolding methods according to the GUM.

- ISO 11929-4 is an explanatory paper giving a lot of numerical examples. The German supplement to DIN ISO 11929:2010 will remain valid. Part 4 will be complementary to the DIN ISO supplement to DIN ISO 11929 Beiblatt 1:2010 (DIN 2014).

In 2021, ISO 11929 became a European standard and the parts 1 to 3 were published as EN ISO and DIN EN ISO standards. Part 4 waits for some revision because of too many misprints and shall only afterwards become an European standard as well.

6.2 The scope of ISO 11929

ISO 11929 makes stipulations for the calculation of the primary result and its associated standard uncertainty, of the characteristic limits, the decision threshold, the detection limit, limits of coverage intervals, and the best estimate and its associated standard uncertainty.

The **primary result of a measurement** and its associated standard uncertainty are calculated as expectation and square root of the variance of the posterior PDF $f_{Y,0}(\tilde{y}|y)$, respectively. If the GUM methodology is used the primary result is obtained as point estimate by the Equation $y = G(x_i; i = 1,...,n)$ and the standard uncertainty is calculated as described in chapter 5.4.

The **decision threshold** y^* provides a decision rule whether or not the physical effect of interest has been observed, i.e. $\tilde{y} > 0$. It is defined as the $(1-\alpha)$-quantile of the likelihood $f_{Y,0}(y|\tilde{y})$:

$$P(y > y^* | \tilde{y} = 0) = \int_{y^*}^{\infty} f_{Y,0}(y|\tilde{y} = 0)\,\mathrm{d}y = \alpha \tag{6.1}$$

α is the preselected probability of a wrong decision in favor of the physical effect of interest if in reality it is absent. From the pure mathematical point of view, α is the probability of getting a measurement result which exceeds the decision threshold if there is no physical effect. The decision rule makes use of a quadratic loss function which is the only one satisfying the requirement for consistency (Weise et al. 2013). For all $\tilde{y} > y^*$ the probability of a correct decision in favor of the presence of the physical effect of interest exceeds that of a wrong decision against its presence.

The **detection limit** $y^\#$ is the smallest true value of the measurand that can be determined with a high reliability; i.e. with a preselected probability β of a wrong decision in favor of the absence of the physical effect of interest given the criterion of the decision threshold. It is defined as the β-quantile of the likelihood $f_{Y,0}(y|\tilde{y} = y^\#)$:

$$P(y < y^* \,|\tilde{y} = y^\#) = \int_{-\infty}^{y^*} f_{Y,0}(y|\tilde{y} = y^\#)\,\mathrm{d}y = \beta \tag{6.2}$$

For graphical presentations of the concept of decision threshold and detection limit see chapters 8 and 9.

Both, the decision threshold and the detection limit are calculated via the **predictive PDF** $f_{Y,0}(y|\tilde{y})$ describing the probability to obtain a primary measurement result y given an assumed true value $\tilde{y}$ of the measurand. If the GUM methodology is applied only the **uncertainty** $\tilde{u}(\tilde{y})$ **as a function of an assumed true value** $\tilde{y}$ of the measurand is needed. It is given as the square root of the variance of $f_{Y,0}(y|\tilde{y})$: $\tilde{u}(\tilde{y}) = \sqrt{\mathrm{Var}\left(f_{Y,0}(y|\tilde{y})\right)}$.

The limits of the **probabilistically symmetric coverage interval** or alternatively of the **shortest coverage interval** define intervals that contain the true value of the measurand with a preselected probability $(1-\gamma)$. The limits of the coverage intervals are calculated as suitable quantiles of the posterior $f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$.

The **best estimate** $\hat{y}$ and its associated **standard uncertainty** $u(\hat{y})$, in contrast to the primary result y and its associated standard uncertainty $u(y)$, take into account the non-negativity of the measurand and are calculated as the mean and the square root of the variance of $f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$.

Though ISO 11929 is entitled “Determination of the characteristic limits (decision threshold, detection limit and limits of the coverage interval) for measurements of ionizing radiation – Fundamentals and application” the applicability of this standard extends beyond measurements of ionizing radiation and can be practically applied to any measurement problem in which a background or blank quantity has to be subtracted from a gross measurement quantity.

It has to be emphasized that there are two conceptual parts in the procedures according to ISO 11929. The first one deals with the establishment of the model of evaluation and the calculation of the primary result and its standard uncertainty using the GUM or the GUM S1. The second part is dealing with the characteristic limits, i.e. the decision threshold, the detection limit, the limits of coverage intervals, and the best estimate and its associated standard uncer-

tainty. The limits of the coverage intervals and the best estimate and its associated standard uncertainty go beyond the strict application of the GUM and the GUM S1 since they take into account of the non-negativity of the measurand.

6.3 The methodology of ISO 11929-2:2019

The methodology of ISO 11929 starts from a model of evaluation

$$Y = G(X_i, i = 1,...,n) \tag{6.3}$$

connecting the input quantities X_i mathematically to the measurand Y. It makes use of the Bayesian theory of measurement uncertainties as dealt with in chapter 5, in particular with the factorization of the posterior $f_Y(\tilde{y}|y)$ according to

$$f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y}) \tag{6.4}$$

ISO 11929-2:2019 is based on the GUM S1 approach. For a model of evaluation, $Y = G(X_i, i = 1,...,n)$ and a given set $\mathbf{a}$ of information, first the PDFs $f_{X_i}(\tilde{x}_i|a_i)$ of the input quantities have to be stablished. This is done by making use of the PME or the Bayes Theorem in combination with non-informative priors. The Markov Formula and Monte Carlo methods are used to **propagate the distributions** and to obtain the PDF $f_{Y,0}(\tilde{y}|\mathbf{a})$. The factorization according to Equation (6.4) is maintained. If additional prior information is available it can be taken into account using the methodology of ISO 11929-2:2019.

The available information is summarized in ISO DIS 11929-2:2019 by the symbol $\mathbf{a}$. It comprises all information about the measurement problem and the values of the input quantities and the measurand available when the evaluation of a measurement is performed. From this information $\mathbf{a}$ the PDFs $f_{X_i,0}(\tilde{x}_i|a_i)$ of the input quantities X_i have to be stablished (a_i being the subsets of information available and relevant for X_i).

The GUM S1 gives detailed advice how the PDFs have to be established. Special cases of obtaining PDFs $f_{X_i,0}(\tilde{x}_i|a_i)$ are dealt with in ISO DIS 11929-1:2019 Appendix A for repeated counting measurements with random influences and in ISO DIS 11929-2:2019 Appendix A for measurements with low count numbers.

After establishing the PDFs of all input quantities, a **joint PDF** $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ has to be formed which in case of independent input quantities is given as

$$f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a}) = \prod_1^m f_{X_i,0}(\tilde{x}_i|a_i) \tag{6.5}$$

A joint PDF has to be assigned to those X_i that are not independent and inserted in Equation (6.5) for the respective input quantities.

The **posterior** PDF $f_{Y,0}(\tilde{y}|\mathbf{a})$ is calculated from the joint PDF $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ using the model Equation $Y = G\left(X_i, i = 1,...,n\right)$ by the so-called **Markov Formula**

$$f_{Y,0}(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{+\infty} f_{\mathbf{X},0}(\boldsymbol{\xi}|\mathbf{a}) \cdot \delta(\tilde{y} - G(\boldsymbol{\xi}))\, \mathrm{d}\boldsymbol{\xi} \ . \tag{6.6}$$

The GUM S1 recommends the application of **Monte Carlo techniques** to solve Equation (6.6) and to derive $f_{Y,0}(\tilde{y}|\mathbf{a})$. Suitable numeric is described in detail elsewhere (Weise et al. 2013). Using the Monte Carlo approach, $i = 1,...,n_{\mathrm{M}}$ Monte Carlo trials are performed for propagating the distributions by drawing sets $x_{1,i},...,x_{n,i}$ from the PDFs $f_{X_i,0}(\tilde{x}_i|a_i)$. For each of these sets, one calculates $y_i = G(x_{1,i},...,x_{n,i})$. The vector $\mathbf{y}_{\mathrm{M}} = \left\{y_1,...,y_{n_{\mathrm{M}}}\right\}$, ordered ascendingly and afterwards assigning probabilities i / n_{M} to the y_i of $\mathbf{y}_{\mathrm{M}}$, is a discrete representation of the **distribution function of the posterior** $F_{Y,0}(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{\tilde{y}} f_{Y,0}(\eta|\mathbf{a})\, \mathrm{d}\eta$ of Y.

If the PDF $f_{\mathbf{X},0}(\tilde{\mathbf{x}}|\mathbf{a})$ is explicitly available, the integral of Equation (6.6) may also have an explicit solution.

The **decision threshold** y^* is the (1-α)-quantile of the predictive probability density function $f_{Y,0}(y|\tilde{y} = 0)$. Assuming a quadratic loss function, the decision threshold provides a decision rule for the decision of whether or not there is a contribution from the sample in a measurement result. Consequently, one decides to conclude that a contribution of the sample is recognized if $y > y^*$. The definition of the decision threshold is therefore $P(y > y^*|\tilde{y} = 0) = \alpha$. The probability of a wrong decision in favor of the physical effect equals a predefined value α. With the decision rule provided by the decision threshold, it holds that the probability for decisions in favor of the physical effect exceeds 50 % for all true values $\tilde{y} > y^*$; see chapter 8.1.

The **detection limit** $y^\#$ is the smallest true value of the measurand which can be recognized with a high probability. Given a predefined probability β of a wrong decision based on the decision threshold in favor of there being no effect of the sample, the detection limit is defined as the β-quantile of the predictive probability density function $f_{Y,0}(y|\tilde{y} = y^\#)$. The definition of the detection limit is therefore $P(y \le y^*|\tilde{y} = y^\#) = \beta$.

The **limits of the probabilistically symmetric coverage interval** $\left[y^\lhd, y^\rhd\right]$ are the γ/2- and (1-γ/2)-quantiles of the posterior PDF $f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$ (Equation 6.4). The **limits of the shortest coverage interval** $\left[y^<, y^>\right]$ also make use of the posterior PDF $f_Y(\tilde{y}|y)$ according to Equation 6.6. See chapter 10.

The **best estimate and its associated standard uncertainty** are calculated as the mean and the square root of the variance of $f_Y(\tilde{y}|y) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot f_Y(\tilde{y}) = C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$. See chapter 11.

6.4 The methodology of ISO 11929-1:2019

ISO 11929-1:2019 is based on the GUM approach and only makes use of the factorization of Equation (6.4) and of expectations and variances of suitable PDFs $f_{X_i,0}(\tilde{x}_i|x_i)$ and $f_{Y,0}(\tilde{y}|y)$ without specifying the underlying PDFs. Thus, ISO 11929-1:2019 is a special case of the methodology of ISO 11929-2:2019 and therefore is presented after the more general ISO 11929-2. It is assumed that the available information **a** consists only of the primary measurement results x_i of X_i and the standard uncertainties $u(x_i)$ associated with the x_i and that the model of evaluation is sufficiently linear to allow for the application of the GUM. This has the consequence that the PDFs are Gaussian distributions so that explicit formulas for the characteristic limits can be given; see chapters 7 to 11 for the formulas.

6.5 UncertRadio – a software package for the application of ISO 11929

There is software available for calculation of characteristic values according to ISO 11929 for radioactivity measurements: **UncertRadio**. It was developed by G. Kanisch (Kanisch 2015a, 2015b). The software „UncertRadio“ enables the automated calculation of characteristic values of activity measurements according to ISO 11929 on computers with Windows operation System. These include the activity concentration or specific activity and its combined standard measurement uncertainty, an uncertainty budget and values of the decision threshold and the detection limit. The uncertainties of the single output values are calculated using numerical uncertainty propagation according to ISO GUM.

UncertRadio can be used for a variety of applications from Alpha, Beta and Gamma measurements, but also from dosimetry. It has the capability to derive the characteristic values for up to three radionuclides simultaneously, whose output quantity values, e.g. activity values, are dependent from each other due to the measurement. Therefore, it is especially suited for modern liquid scintillation measurement procedures of e.g. strontium isotopes.

There are two main analytical approaches used within the software:

- procedures without linear unfolding: numerically applied propagation of uncertainty values of input quantities ,
- procedures with linear unfolding: linear least squares procedures (e.g. for decay or build-up curves);

Alternatively, an evaluation by Monte Carlo simulation may be used within both approaches. This represents the method of propagating whole distributions and is in advantage in the case of significant deviations from the normal distribution; see ISO GUM Supplements 1 and 2.

Applying UncertRadio means that the user should be able to formulate the formulas used for the calculation of values of the output quantity. A special advantage is, however, that no partial derivatives are to be supplied. For a better understanding of the course of the software and

of user-supplied formulas and functions an extensive bundle of application examples as project files is included into the installation.

Many of the application examples are from the working group "AK-SIGMA" of the “German-Suisse Association for Radiation Protection”, from Measuring Instructions of the German Coordinating Offices and from the literature. These examples including those of a German National Supplement 1 of DIN ISO 11929 (2014) contributed to the validation of UncertRadio.

The author is grateful for the feedback and new requirements of the user-community, especially from federal coordination centres and the laboratories of the German Federal States working according to AVV-IMIS on monitoring of environmental radioactivity, which significantly improved the quality and usability of UncertRadio.

The actual version can be downloaded for free from https://www.thuenen.de/en/fi/fields-of-activity/marine-environment/coordination-centre-of-radioactivity/uncertradio/.

The download consists of one executable file, which combines all required components of the Software (help-files, short installation guide, and collection of validated example projects). Additionally, a brief instruction to use the software may be downloaded. The private or commercial use of the software is free of charge.

The program’s author is Günter Kanisch. Contact person for questions and suggestions is Dr. Marc-Oliver Aust from the “Federal co-ordinating office for fish and fishery products, crustaceans, molluscs and marine algae” in the Thünen-Institute of Fisheries Ecology.

Other software packages, including EXCEL™ in combination with e.g. Chrystal Ball™, allow for calculating the characteristic limits and are commercially available.

7 Uncertainty associated with an assumed true value of the measurand

7.1 General aspects

The uncertainty $\tilde{u}(\tilde{y})$ associated with an assumed true value $\tilde{y}$ of the measurand Y is calculated via the predictive PDF $f_{Y,0}(y|\tilde{y})$ by $\tilde{u}(\tilde{y}) = \sqrt{\mathrm{Var}\left(f_{Y,0}(y|\tilde{y})\right)}$. $f_{Y,0}(y|\tilde{y})$ or $\tilde{u}(\tilde{y})$ are needed for the calculation of the decision threshold (chapter 8) and the detection limit (chapter 9) on the basis of the GUM S1 or the GUM, respectively.

7.2 Predictive PDF associated with an assumed true value of the measurand

If GUM S1 is used, the following holds. For the provision and numerical calculation of the decision threshold in chapter 8 and of the detection limit in chapter 9, the standard uncertainty of the measurand is needed for assumed true values of the measurand. In ISO 11929-1, in which the GUM methodology is used, the standard uncertainty as a function of assumed true values of the measurand could be explicitly given. This is not possible if the GUM S1 is used, but standard uncertainties for individual assumed true values of the measurand can be determined in a way similar to $u(y)$ on the basis of the GUM S1.

In the approach according to GUM S1 the standard uncertainty as a function of assumed true values of the measurand can be derived as approximation by repeated numerical calculations for assumed true values $\tilde{y}$ of the measurand resulting in probability distribution $f_Y(y|\mathbf{a}')$ with a modified set of information $\mathbf{a}'$ which takes into account the remaining information of $\mathbf{a}$. To this end, the value of the gross input quantity x_1 has to be modified; see equation (5.39). The standard uncertainty for an assumed true value of the measurand is derived from the variance of the probability distribution $f_Y(y|\mathbf{a}')$.

For the calculation of the decision threshold and the detection limit, iterative methods have to be used. For each iteration step, a new probability distribution $f_{X_1}(\tilde{x}_1|\mathbf{a}', x_1)$ has to be established with a modified value x_1 of the gross quantity X_1. Then, $i = 1,...,n_M$ new Monte Carlo trials are performed by drawing sets $x_{1,i},...,x_{n,i}$ from the probability distributions $f_{X_1}(\tilde{x}_1|\mathbf{a}', x_1), f_{X_i}(\tilde{x}_i|\mathbf{a}'), i = 2,...,n)$. For each of these sets one again calculates $y_i = G(x_{1,i},...,x_{n,i})$. The new vector $\boldsymbol{y}_M(\boldsymbol{a}') = \{y_1,...,y_{n_M}\}$ ordered ascendingly and afterwards assigning probabilities i/n_M to the y_i of $\mathbf{y}_M(\mathbf{a}')$ is a discrete representation of the distribution function $F_Y(\tilde{y}|\mathbf{a}', x_1) = \int_{-\infty}^{\tilde{y}} f_Y(\eta|\mathbf{a}', x_1)\,\mathrm{d}\eta$ of Y.

7.3 Uncertainty as a function of an assumed true value of the measurand

For the provision and numerical calculation of the decision threshold in 8.2 and of the detection limit in 8.3, the standard uncertainty of the measurand is needed as a function $\tilde{u}(\tilde{y})$ of an assumed true value $\tilde{y} > 0$ of the measurand. This function shall be determined in a way similar to $u(y)$ within the framework of the evaluation of the measurements by application of the GUM or GUM Supplement 1. In most cases, $\tilde{u}(\tilde{y})$ shall be formed as a positive square root of a variance function $\tilde{u}^2(\tilde{y})$ calculated first. This function shall be defined, unique and continuous for all $\tilde{y} > 0$, and shall not assume negative values.

The function $\tilde{u}^2(\tilde{y})$ is schematically shown in Fig. 17. Several cases are shown in this figure. In principle, this function can be determined experimentally from multiple measurements of a series of reference materials of different levels of the physical effect of interest. It is the characteristics of a reference material that its value is assumed to represent the true value of the measurand. Such multiple measurement results are indicated by the circles in the left panel of Fig. 17. A particular point is $\tilde{u}^2(\tilde{y} = 0)$. For any model $y = (r_g - r_0) \cdot w$ a measurement of the background r_0 and the evaluation of its associated uncertainty $u(r_0)$ are needed. If the true value of the measurand $\tilde{y}$ is zero one expects $r_g = r_0$ and consequently one obtains $\tilde{u}^2(0) = 2 \cdot u^2(r_0) \cdot w$. Thus, $\tilde{u}^2(0)$ is always available.

One may ask what the minimum requirements are to calculate $\tilde{u}^2(\tilde{y})$ and to obtain the decision threshold and the detection limit. For the successful invention of a measurement method it is required that a measurement of the background and at least one successful measurement

of a sample with $y \neq 0$ is performed and that the associated uncertainties $u^2(0)$ and $u(y)$ are calculated. Then, the function $\tilde{u}^2(\tilde{y})$ can be approximated by interpolation using the Equation (7.1).

$$\tilde{u}^2(\tilde{y}) = \tilde{u}^2(0) \cdot (1 - \tilde{y}/y) + u^2(y) \cdot \tilde{y}/y \tag{7.1}$$

This case is shown in the right panel of Fig. 17. If more measurements are available, higher order interpolations can be used as described in detail in ISO 11929-1:2019 Appendix A.2.

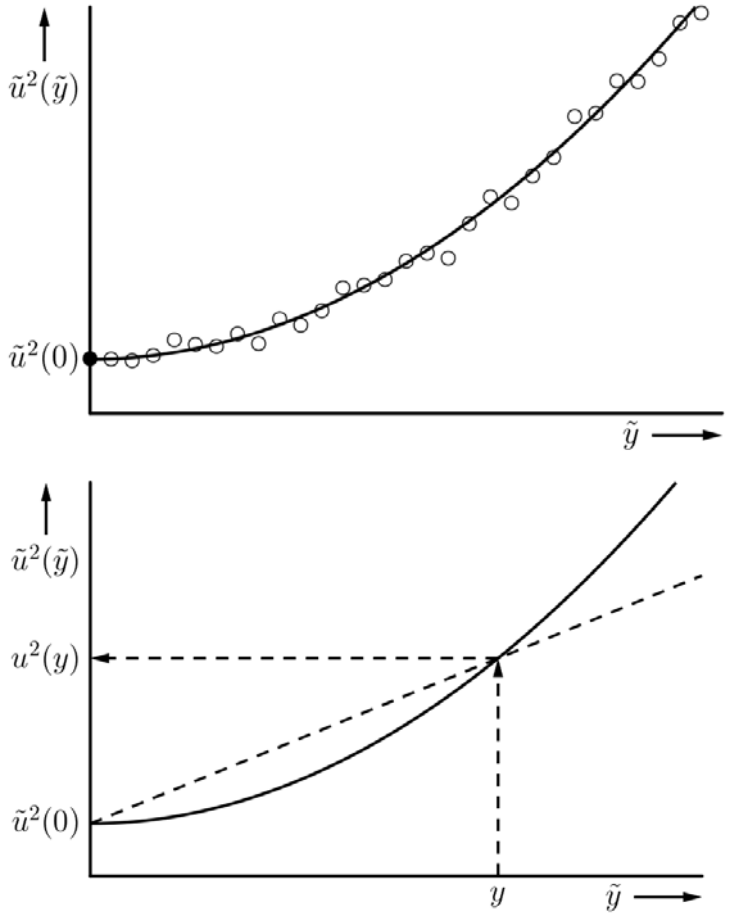


Fig. 17: Simple picture of $\tilde{u}^2(\tilde{y})$ (upper panel) and of the interpolation equation (7.1) (lower panel).

In most cases of models of the general form $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots}$ with $x_1 = r_g = n_g / t_g$ and $x_2 = r_0 = n_0 / t_0$, $\tilde{u}(\tilde{y})$ can be explicitly specified, provided that $u(x_1)$ is given as a function of x_1. Such cases are dealt with below.

7.4 Counting measurements

ISO 11929-1:2019 covers the general model of evaluation

$$y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} \text{ with } x_1 = r_g = n_g / t_g \text{ and } x_2 = r_0 = n_0 / t_0 \tag{7.2}$$

In the case of the model, the standard uncertainty, $u(x_1)$, of the gross count rate $X_1 = R_g$, is given as a function of the estimate, $x_1 = r_g$, either $\sqrt{x_1 / t_g}$ or $x_1 / \sqrt{n_g}$ applies if the measurement duration, t_g (time preselection), or, respectively, the number, n_g, of recorded pulses (preselection of counts) is specified.

In order to calculate $\tilde{u}^2(\tilde{y})$, the value y shall be formally replaced by $\tilde{y}$ in Equation (7.2). This allows the elimination of x_1 in the general case and, in particular, of n_g with time preselection and of t_g with preselection of counts. This yields in the case according to Equation (7.3)

$$x_1 = \tilde{y}/w + x_2 x_3 + x_4 \tag{7.3}$$

with time preselection

$$n_g = t_g \cdot (\tilde{y}/w + r_0 x_3 + x_4) \,. \tag{7.4}$$

Then, with $u^2(x_1) = x_1 / t_g = n_g / t_g^2$ and by substituting n_g according to Equation (7.4) and with $u^2(x_2) = r_0 / t_0$, Equation (7.2) leads in the case of **time preselection** to

$$\tilde{u}(\tilde{y}) = \sqrt{w^2 \cdot \left[(\tilde{y}/w + r_0 x_3 + x_4)/t_g + x_3^2 r_0 / t_0 + r_0^2 u^2(x_3) + u^2(x_4) \right] + \tilde{y}^2 u_{\text{rel}}^2(w)} \tag{7.5}$$

For the simple model $y = (x_1 - x_2) \cdot w$ one obtains in the case of **time preselection**

$$\tilde{u}^2(\tilde{y}) = \tilde{y}^2 \cdot u_{\text{rel}}^2(w) + w^2 \cdot \left[\frac{\tilde{y}}{t_g \cdot w} + \frac{n_0}{t_0} \cdot \left(\frac{1}{t_g} + \frac{1}{t_0} \right) \right]$$

With preselection of counts

$$t_g = \frac{n_g}{\tilde{y}/w + r_0 x_3 + x_4} \tag{7.6}$$

is analogously obtained. Then, with $u^2(x_1) = x_1 / t_g = n_g / t_g^2$ and by substituting t_g according to Equation (7.6) and with $u^2(x_2) = r_0^2 / n_0$, Equation (7.2) leads in the case of **preselection of counts** to

$$\tilde{u}(\tilde{y}) = \sqrt{w^2 \cdot \left[(\tilde{y}/w + r_0 x_3 + x_4)^2 / n_g + x_3^2 r_0^2 / n_0 + r_0^2 u^2(x_3) + u^2(x_4) \right] + \tilde{y}^2 u_{\text{rel}}^2(w)} \,. \tag{7.7.}$$

If the standard uncertainty cannot be explicitly given as a function of an assumed true value of the measurand, an approximation by interpolation described in chapter 7.3 has to be applied.

Preselection of counts is frequently chosen in order to economize the available time for measurements. If in addition a maximum measurement time $t_{\max}$ is set, it must be ensured that $t_{\max} \geq n_g / (n_0 \cdot x_3 / t_0 + x_4)$. Otherwise, the evaluation changes the chosen model of evaluation and the decision threshold and the detection limit are wrongly calculated.

If no explicit specification of $\tilde{u}(\tilde{y})$ is available, it is often sufficient to use the following approximations for the function $\tilde{u}(\tilde{y})$, in particular, if the standard uncertainty $u(x_1)$ is not known as a function of x_1. A prerequisite is that measurement results y_j and associated standard uncertainties $u(y_j)$ calculated from previous measurements of the same kind, are already available $(j = 0,1,2,...)$. The measurements shall be carried out on different samples with differing activities, but in other respects as far as possible under similar conditions. One of the measurements can be a background effect measurement or a blank measurement with $\tilde{y} = 0$ and, for instance, $j = 0$. Then, $y_0 = 0$ shall be set and $\tilde{u}(0) = u(y_0)$. The measurement currently carried out can be taken as a further measurement with $j = 1$.

7.5 Counting measurements with random influences

7.5.1 Procedure with unknown influences

Assume that $u^2(x_1)$ is not given as a function of x_1. Therefore, $\tilde{u}^2(\tilde{y})$ shall be determined as an approximation, for instance, according to Equation (7.1), where the current result y can be used as y_1. For this purpose and for the calculation of $\tilde{u}^2(0)$, i.e. for $\tilde{y} = 0$, the missing $\overline{n}_g^2 / t_g^2$ shall be replaced by $\overline{n}_0^2 / t_0^2$, since both these values are then variance estimates of the same distribution of count rate values, independent of t_g, t_0, m_g and m_0.

It is often sufficient to use the approximations of Equations (7.1) for the function $\tilde{u}(\tilde{y})$, in particular, if the standard uncertainty $u(x_1)$ is not known as a function of x_1. A prerequisite is that measurement results y_j and associated standard uncertainties $u(y_j)$ calculated according to chapter 5.6.2 from previous measurements of the same kind, are already available $(j = 0,1,2,...)$. The measurements shall be carried out on different samples with differing activities, but in other respects as far as possible under similar conditions. One of the measurements can be a background effect measurement or a blank measurement with $\tilde{y} = 0$ and, for instance $j = 0$. Then, $y_0 = 0$ shall be set and $\tilde{u}(0) = u(y_0)$. The measurement currently carried out can be taken as a further measurement with $j = 1$.

For an assumed true value of the measurand $\tilde{y} = 0$ one expects $\frac{\overline{n}_g}{t_g} = \frac{\overline{n}_0}{t_0}$ and obtains for the model according to Equation (7.2) with Equation (5.45)

$$\tilde{u}^2(0) = w^2 \cdot \left[u^2(\overline{n}_0) / t_g^2 + x_3^2 \cdot u^2(\overline{n}_0) / t_0^2 + (\overline{n}_0 / t_0)^2 \cdot u^2(x_3) + u^2(x_4) \right] \tag{7.8}$$

7.5.2 Procedure with known influences

In order to calculate $\tilde{u}(\tilde{y})$, the result y is replaced by $\tilde{y}$ and Equation (7.2) is solved for $x_1 = \overline{n}_g / t_g$. This yields $x_1 = \tilde{y}/w + \overline{n}_0 / t_0$. The estimate x_1 determined in this way in the current case, shall be substituted in Equation (5.48) and $u^2(x_1)$ from Equation (5.48). This finally leads to $\tilde{u}(\tilde{y})$

$$\tilde{u}(\tilde{y}) = \sqrt{\tilde{y}^2 \left[\frac{\vartheta^2}{m_g} + u_{rel}^2(w) \right] + \frac{\tilde{y}w}{m_g} \left(\frac{2\overline{n}_0 \vartheta^2}{t_0} + \frac{1}{t_g} \right) + w^2 \left(\frac{\overline{n}_0}{m_g t_0 t_g} + \frac{\overline{n}_0^2 \vartheta^2}{m_g t_0^2} + \frac{\overline{n}_0 + \vartheta^2 \overline{n}_0^2}{m_0 t_0^2} \right)} \tag{7.9}$$

7.5.3 Black box measurements

Applying the GUM in ISO 11929-1:2019, one assumes in order to obtain the decision threshold that $\overline{y}_g = \overline{y}_b$ and $s_g = s_b$ will hold for a true value $\tilde{y} = 0$ of *Y*. This yields

$$\tilde{u}^2(\tilde{y} = 0) = \left(\frac{n_g - 1}{n_g - 3} \cdot \frac{1}{n_g} + \frac{n_b - 1}{n_b - 3} \cdot \frac{1}{n_b} \right) \cdot s_b^2 \tag{7.10}$$

8 The decision threshold

8.1 General aspects and the definition of the decision threshold

The decision threshold provides a decision rule with given probability criteria whether or not a contribution of the sample has been observed taking into account the background measurement and the uncertainty associated with it.

The decision threshold (Fig. 18) is defined by the basic decision criterion for the presence or absence of a contribution of the sample by

$$P(y > y^* \,|\, \tilde{y} = 0) = \int_{y^*}^{\infty} f_{Y,0}(y \,|\, \tilde{y} = 0)\, \mathrm{d}y = \alpha \tag{8.1}$$

α is a preselected probability for the wrong decision to accept the existence of a contribution from the sample if in reality there is none. If a measured value y exceeds the decision threshold y^* one decides that a contribution from the sample has been observed.

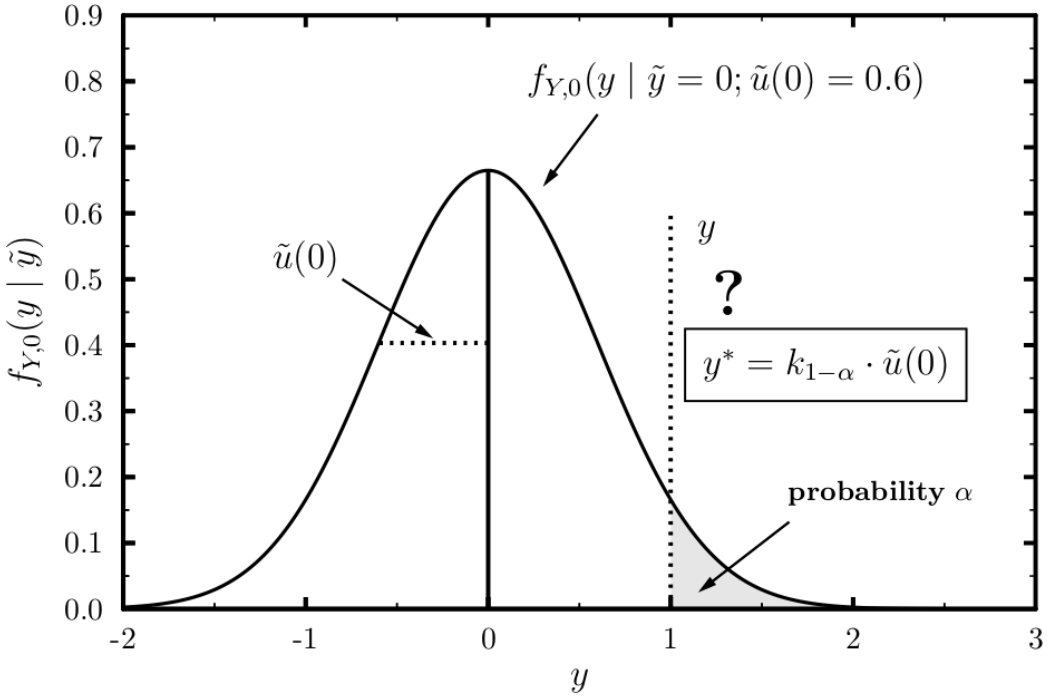


Fig. 18: The decision threshold.

At this point it is somewhat obscure how to derive such a criterion or how to fix a decision threshold in ISO 11929:2010. The reason for both is that decisions are a matter of decision theory, e.g. (Berger, 1985). According to decision theory, decisions need a loss function. For a given set of information the optimal action to decide between two options a_0 and a_1 will be to accept option a_0 if (and only if) the expected posterior loss of accepting is smaller than the expected posterior loss of rejecting the option a_0 and accepting a_1. Also the consequences of loss or gain can only be quantified by probabilities. But, the critical probability α for acceptance and rejection has to be chosen by humans weighting the importance of loss or gain.

In metrology, only second-order loss functions, e.g. used to define uncertainty or in least squares fits including an uncertainty treatment, can meet – at least in a linear model approximation in the data range of interest – the requirement of consistency (Weise and Wöger 1999, 2000). This choice goes deeply into the basic metrological requirements. Any uncertainty theory should meet at least the following six metrological requirements: (1) generality, (2) consistency (information conservation), (3) unique statistical basis, (4) possibility of a critical comparison of different measurement results of the same measurand, (5) protection against large possible measurement deviations, and (6) simplicity, transparency, practicability, and small computing effort.

Weise and Wöger (1999, 2000) have shown that only quadratic functions such as in the least squares method satisfy this requirement for consistency. The PME and Bayes Theorem are also consistent when connected with the same preconditions.

The decision threshold has to be calculated on the basis of the background or blank measurement alone which give the information about $f_{Y,0}(y|\tilde{y}=0)$. For the decision threshold this is sufficient and no further information is needed. For the calculation of the detection limit some more information is required. For the calculation of $f_{Y,0}(y|\tilde{y})$ the expected uncertainties $\tilde{u}(\tilde{y})$ for assumed true values $\tilde{y}$ of the measurand Y have to be known. The required information

regards the standard uncertainty as a function of the true value $\tilde{y}$ of the measurand Y. $\tilde{u}(\tilde{y}=0)$ can be obtained from $f_{Y,0}(y|\tilde{y}=0)$ resulting from a background or blank measurement. For another assumed true value $\tilde{y}$ it is possible to experimentally determine the value of this function, e.g. by measurements of reference materials. $\tilde{u}(\tilde{y})$ can also be approximated by interpolation using the information about $\tilde{u}(\tilde{y}=0)$ and a measurement result y and its associated standard uncertainty $u(y)$. In many cases of counting radiation measurements $\tilde{u}(\tilde{y})$ can be explicitly calculated (see below).

The decision threshold is not based on a frequentist hypothesis test testing the null hypothesis $H_0:\tilde{y}=0$ against the alternative hypothesis $H_1:\tilde{y}>0$. It has to be mentioned that in frequentist statistics rejecting the null hypothesis does not imply accepting the alternative hypothesis. This is only done with a decision theory using Bayesian probabilities and an appropriate loss function. Any decision bears the risk of being wrong. With the decision rule given by the decision threshold and with a preselected probability of $\alpha=0,05$, the probability for a wrong decision in favor of an effect of the sample, if in reality there is none effect, is $P(\tilde{y}=0)\le 0,05$. For any true value larger than the decision threshold the probability of making the right decision (i.e. the decision in favor of an effect of the sample) is $P(\tilde{y}>0)>0,5$.

A better estimation of the probabilities involved in the decision rule provided by the decision threshold is given by the Bayes factor; see e.g. Goodman (1999). The **Bayes factor** is a likelihood ratio for two competing hypotheses, called models M_1 and M_2, given data D. It is defined as

$$K=\frac{P(D|M_1)}{P(D|M_2)}=\frac{\dfrac{P(M_1|D)\cdot P(D)}{P(M_1)}}{\dfrac{P(M_2|D)\cdot P(D)}{P(M_2)}}=\frac{P(M_1|D)}{P(M_2|D)}\cdot\frac{P(M_2)}{P(M_1)} \tag{8.2}$$

Where $P(M_1|D)$ and $P(M_2|D)$ are the likelihoods of the two models given the data and $P(M_1)$ and $P(M_2)$ are the respective priors.

H. Jeffreys (1939) proposed the following scheme for judgements about the evidence provided by the Bayes factor. The reason of the scheme is as follows. If $P(D|M_1)>P(D|M_2)$ and hence $K>1$ the data is more likely to arise in the case that model 1 applies than in the case that model 2 applies.

K	Evidence for M_1
< 1	Negative
1 – 3,16	Barely worth mentioning
3,16 – 10	Substantial
10 to 31,6	Strong
31,6 – 100	Very strong
>100	Decisive

Assuming ignorance the probabilities $P(M_1) = P(M_2) = 0,5$ should hold. Remember he probability assessment for throwing a coin. With a probability of $\alpha = 0,05$ and $M_1 \triangleq (\tilde{y} = 0)$ and $M_2 \triangleq (\tilde{y} > 0)$ one obtains for the Bayes factor

$$K = \frac{P(\tilde{y} = 0 \,|\, y = y^*)}{P(\tilde{y} > 0 \,|\, y = y^*)} \cdot \frac{P(\tilde{y} > 0)}{P(\tilde{y} = 0)} = \frac{\mathrm{N}(0)}{\mathrm{N}(0,05)} \cdot \frac{0,5}{0,5} = \frac{0,398}{0,1} = 3,98 \tag{8.3}$$

which gives substantial evidence for the model $M_2 \triangleq (\tilde{y} > 0)$.

Whereas the approach put forward here is correct in principle, there remains a *caveat*. It is not suited to test a point model ($\tilde{y} = 0$) against an interval model ($\tilde{y} > 0$). With the exception of very rare instances mentioned below, perfect zero does not exist in practice. It is impossible to claim that the true value is exactly zero rather than being very small but distinct from zero. It is also virtually impossible to decide by an experiment whether the true value is perfectly zero or, say, 1E-9. Consequently, without prior knowledge it will have to be taken as equally likely that the true value is 0, 1E-9, 2E-9, 3E-9 and so on. Therefore, in the absence of prior knowledge the probability of $\tilde{y} > 0$ will always be far greater than the probability of $\tilde{y}$ being exactly zero.

8.2 The decision threshold with uncertainties according to the GUM

The GUM and ISO 11929-1 assume that only y and its associated standard uncertainty $u(y)$ are available and that the measurand is non-negative. With this the respective PDFs $f_{Y,0}(y \,|\, \tilde{y} = 0)$ and $f_{Y,0}(y \,|\, \tilde{y})$ are Gaussians and the following formula can be given for the decision threshold

$$y^* = k_{1-\alpha} \cdot \tilde{u}(0) \tag{8.4}$$

with $k_{1-\alpha}$ being the $(1-\alpha)$-quantile of the standard normal distribution. The Equation (8.4) in combination with the formulas given in chapter 7.4 and 7.5 lead to the explicit formulas for the decision threshold.

For a counting measurement with the general model of evaluation $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \cdots}{x_5 x_7 \cdots} = (x_1 - x_2 x_3 - x_4) \cdot w$ and for **preselection of time** the decision threshold according to Equation (7.5) is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(\tilde{y} = 0) = k_{1-\alpha} \cdot w \sqrt{(r_0 x_3 + x_4)/t_\mathrm{g} + x_3^2 r_0 / t_0 + r_0^2 u^2(x_3) + u^2(x_4)} \tag{8.5}$$

For the simple and frequently used counting measurement with the model of evaluation $y = (x_1 - x_2) \cdot w$ and **preselection of time** the decision threshold is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(\tilde{y} = 0) = k_{1-\alpha} \cdot w \cdot \sqrt{r_0 \cdot \left(\frac{1}{t_\mathrm{g}} + \frac{1}{t_0} \right)} \tag{8.6}$$

For a counting measurement with the model of evaluation according to Formula (5.39) $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} = (x_1 - x_2 x_3 - x_4) \cdot w$ and **preselection of counts** the decision threshold according to Equation (7.7) is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(\tilde{y} = 0) = k_{1-\alpha} \cdot w \sqrt{(r_0 x_3 + x_4)^2 / n_\mathrm{g} + x_3^2 r_0^2 / n_0 + r_0^2 u^2(x_3) + u^2(x_4)} \tag{8.7}$$

For **counting measurements with unknown random influences** with the model of evaluation according to $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} = (x_1 - x_2 x_3 - x_4) \cdot w$ with $x_1 = \overline{n}_\mathrm{g} / t_\mathrm{g}$; $x_2 = \overline{n}_0 / t_0$ the decision threshold according to Equation (7.8) is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(0) = k_{1-\alpha} \cdot w \cdot \sqrt{u^2(\overline{n}_0) / t_\mathrm{g}^2 + x_3^2 \cdot u^2(\overline{n}_0) / t_0^2 + (\overline{n}_0 / t_0)^2 \cdot u^2(x_3) + u^2(x_4)} \tag{8.8}$$

For **counting measurements with known random influences** with the model of evaluation $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} = (x_1 - x_2 x_3 - x_4) \cdot w$ the decision threshold according to Equation (7.9) is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(\tilde{y} = 0) = k_{1-\alpha} \sqrt{w^2 \left(\frac{\overline{n}_0}{m_\mathrm{g} t_0 t_\mathrm{g}} + \frac{\overline{n}_0^2 \vartheta^2}{m_\mathrm{g} t_0^2} + \frac{\overline{n}_0 + \vartheta^2 \overline{n}_0^2}{m_0 t_0^2} \right)} \tag{8.9}$$

For **black box measurements** with the model of evaluation $y = (\overline{y}_\mathrm{g} - \overline{y}_\mathrm{b}) \cdot \mathrm{w}$ with $\overline{y}_\mathrm{g} = \frac{1}{n_\mathrm{g}} \sum_{i=1}^{n_\mathrm{g}} y_{\mathrm{g},i}$, $\overline{y}_\mathrm{b} = \frac{1}{n_\mathrm{b}} \sum_{i=1}^{n_\mathrm{b}} y_{\mathrm{b},i}$, and $s_\mathrm{b} = \left(\frac{1}{n_\mathrm{b} - 1} \sum_{i=1}^{n_\mathrm{b}} (y_{\mathrm{b},i} - \overline{y}_\mathrm{b})^2 \right)^{1/2}$ the decision threshold according to Equation (7.10) is given by

$$y^* = k_{1-\alpha} \cdot \tilde{u}(\tilde{y} = 0) = k_{1-\alpha} \cdot w \cdot \sqrt{\frac{n_\mathrm{g} - 1}{n_\mathrm{g} - 3} \cdot \frac{1}{n_\mathrm{g}} + \frac{n_\mathrm{b} - 1}{n_\mathrm{b} - 3} \cdot \frac{1}{n_\mathrm{b}}} \cdot s_\mathrm{b} \tag{8.10}$$

The decision threshold does not depend on the uncertainty of the calibration factor. This is because – assuming a properly working instrument – a measurement can indicate that there is a contribution from the sample without knowing its value; i.e. independent from the calibration.

8.3 The decision threshold with uncertainties according to the GUM S1

The decision threshold y^* of the non-negative measurand quantifying the physical effect of interest, is the value of the estimator Y which allows the conclusion that the physical effect is present, if the primary measurement result y exceeds the decision threshold y^*. If the result y is below the decision threshold y^* the result cannot be attributed to the physical effect, nevertheless it cannot be concluded that it is absent. If the physical effect is really absent, the prob-

ability of taking the wrong decision, that the effect is present, is equal to the specified probability, α (probability of the wrong decision that it is present if it actually is absent).

A determined primary measurement result y for the non-negative measurand suggests that the true value of the measurand is likely to differ from zero ($\tilde{y} > 0$) only if it is larger than the decision threshold

$$P(y > y^* \,|\, \tilde{y} = 0) = \int_{y^*}^{\infty} f_{Y,0}(y' \,|\, \mathbf{a}', \tilde{y} = 0)\, \mathrm{d}y' = \alpha \tag{8.11}$$

The integral according to Equation (8.11) has to be evaluated using Monte Carlo techniques. See Weise et al. (2009) for details of the suitable numeric.

Using the Monte Carlo approach, the decision threshold y^* is calculated by iteration in form of a root-finding problem which can be solved by bisection methods, *regula falsi* or interpolation. For each iteration step, a new probability distribution $f_{X_1,0}(\tilde{x}_1 \,|\, \mathbf{a}', x_1)$ has to be established with a modified value x_1 of the gross quantity X_1. Then, $i = 1, ..., n_\mathrm{M}$ new Monte Carlo trials are performed by drawing sets $x_{2,i}, ..., x_{n,i}$ from the probability distributions $f_{Y,0}(y' \,|\, \mathbf{a}', \tilde{y} = 0)$. For each of these sets, one again calculates $y_i = G(x_{1,i}, ..., x_{n,i})$. The new vector $\mathbf{y}_\mathrm{M}(\mathbf{a}') = \{y_1, ..., y_{n_\mathrm{M}}\}$, ordered ascendingly and afterwards assigning cumulative probabilities i / n_M to the y_i of $\vec{\mathbf{y}}_\mathrm{M}(\mathbf{a}')$, is a discrete representation of the distribution function $F_Y(\tilde{y} \,|\, \mathbf{a}', x_1) = \int_{-\infty}^{\tilde{y}} f_Y(\eta \,|\, \mathbf{a}', x_1)\, \mathrm{d}\eta$ of Y. By iteration one searches for the function $F_Y(\tilde{y} \,|\, \mathbf{a}', x_1) = \int_{-\infty}^{\tilde{y}} f_Y(\eta \,|\, \mathbf{a}', x_1)\, \mathrm{d}\eta$ with $\mathrm{E}\left(f_Y(\tilde{y} \,|\, \mathbf{a}', x_1)\right) = 0$. This yields a particular value x_1 for which it is known that on average $\tilde{y}$ equals 0 or is at least very close to 0. From the respective $\mathbf{y}_\mathrm{M}(\mathbf{a}') = \{y_1, ..., y_{n_\mathrm{M}}\}$ of the probability distribution, one calculates the probability $P(y^* < y_k) = k / n_\mathrm{M}$ by searching the largest index k with $y_k > y^*$. The (1-α)-quantile of this vector $\mathbf{y}_\mathrm{M}(\mathbf{a}') = \{y_1, ..., y_{n_\mathrm{M}}\}$ is the decision threshold y^*.

If the probability distribution $f_{Y,0}(y' \,|\, \mathbf{a}', \tilde{y} = 0)$ is explicitly available, the integral according to Equation (8.11) can be evaluated by any suitable means.

8.4 Assessment of a measurement

The primary measurement result y has to be compared with the decision threshold y^*. If the primary measurement result y exceeds the decision threshold y^*, it is decided that the physical effect provided by the measurand is present, i.e. that a contribution from the sample has been recognized.

If the result y is below the decision threshold y^*, it is decided that the result cannot be attributed to the physical effect. Nevertheless, it cannot be concluded that it is absent. If the

physical effect is really absent, the probability of taking the wrong decision, that the effect is present, is equal to the specified probability α.

9 The detection limit

9.1 General aspects and the definition of the detection limit

The detection limit (Fig. 19) has to be set sufficiently high above the decision threshold to avoid an unduly high probability for the wrong decision that there is no contribution from the sample if in reality there is one. Given the criterion of the decision threshold, the detection limit is defined by

$$P(y < y^* \,|\, \tilde{y} = y^\#) = \int_{-\infty}^{y^*} f_{Y,0}(y \,|\, \tilde{y} = y^\#)\, \mathrm{d}y = \beta \tag{9.1}$$

β is the probability of the wrong decision that the effect of the sample is absent if in reality there is one.

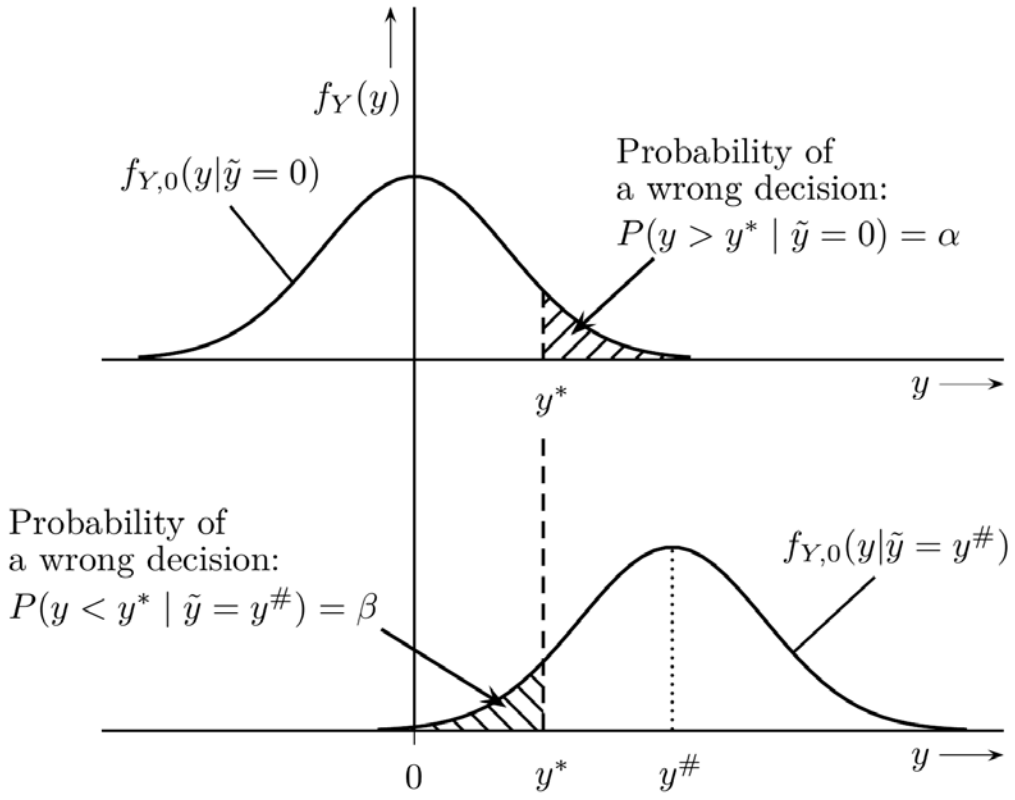


Fig. 19: Decision threshold y^* and detection limit $y^\#$ and the probabilities for wrong decisions.

9.2 The detection limit with uncertainties according to the GUM

If the GUM is used for the evaluation of uncertainties, the detection limit is given by

$$y^{\#} = y^{*} + k_{1-\beta} \cdot \tilde{u}(y^{\#}) = k_{1-\alpha} \cdot \tilde{u}(0) + k_{1-\beta} \cdot \tilde{u}(y^{\#}) \tag{9.2}$$

with $k_{1-\alpha}$ and $k_{1-\beta}$ being the $(1-\alpha)$- and $(1-\beta)$-quantiles of the standard normal distribution, respectively.

The detection limit $y^{\#}$ is the smallest true value of the measurand, for which, by applying the decision rule according to Chapter 8.2, the probability of the wrong decision that the physical effect is absent does not exceed the specified probability β. In order to find out whether a measurement procedure is suitable for the measurement purpose, the detection limit $y^{\#}$ is compared with the specified guideline value y_{r} of the measurand.

For the general model of evaluation $y = (x_1 - x_2 x_3 - x_4) \cdot \dfrac{x_6 x_8 \ldots}{x_5 x_7 \ldots} = (x_1 - x_2 x_3 - x_4) \cdot w$ the detection limit $y^{\#}$ is obtained in the case of **time preselection** as the smallest solution of Equation (9.3)

$$y^{\#} = y^{*} + k_{1-\beta} \cdot \tilde{u}(y^{\#}) = y^{*} + k_{1-\beta} \cdot \sqrt{w^2 \cdot \left((y^{\#} / w + x_3 \cdot n_0 / t_0 + x_4) / t_g + x_3^2 \cdot n_0 / t_0^2 + u^2(x_3) \cdot n_0^2 / t_0^2 + u^2(x_4) \right) + y^{\#^2} \cdot u_{\mathrm{rel}}^2(w)} \tag{9.3}$$

Equation (9.3) has a solution, which is the detection limit $y^{\#}$, if with preselection of time the following condition is satisfied

$$k_{1-\beta} \cdot u_{\mathrm{rel}}(w) < 1 \tag{9.4}$$

For the simple and frequently used counting measurement with the model of evaluation $y = (x_1 - x_2) \cdot w$ and **preselection of time** the detection limit is implicitly given by

$$y^{\#} = y^{*} + k_{1-\beta} \cdot \tilde{u}(y^{\#}) = y^{*} + k_{1-\beta} \cdot \sqrt{y^{\#2} \cdot u_{\mathrm{rel}}^2(w) + w^2 \cdot \left[\frac{y^{\#}}{t_{\mathrm{g}} \cdot w} + \frac{n_0}{t_0} \cdot \left(\frac{1}{t_{\mathrm{g}}} + \frac{1}{t_0} \right) \right]} \tag{9.5}$$

For $k_{1-\alpha} = k_{1-\beta} = k$ one obtains the explicit Equation for the detection limit

$$y^{\#} = \frac{2 \cdot y^{*} + (k^2 \cdot w) / t_{\mathrm{g}}}{1\text{-}k^2 \cdot u_{\mathrm{rel}}^2(w)} \tag{9.6}$$

If $1\text{-}k^2 \cdot u_{\mathrm{rel}}^2(w) \le 0$, i.e. if $k \cdot u_{\mathrm{rel}}(w) \ge 1$, the Equations (9.5) and (9.6) for the detection limit have no solution. This is a consequence of the fact that for large relative uncertainties of the calibration factor w the GUM approximation is no longer sufficient. In this case one has to proceed according to the GUM S1 and ISO 11929-2:2019.

The detection limit $y^{\#}$ is obtained in the case of **preselection of counts** as the smallest solution of Equations (9.7)

$$y^{\#} = y^{*} + k_{1-\beta} \cdot \tilde{u}(y^{\#}) = y^{*} + k_{1-\beta} \cdot \sqrt{w^2 \cdot \left[(y^{\#} / w + n_0 \cdot x_3 / t_0 + x_4)^2 \big/ n_{\mathrm{g}} + x_3^2 \cdot n_0 \big/ t_0^2 + n_0^2 \cdot u^2(x_3) / t_0^2 + u^2(x_4) \right] + y^{\#2} \cdot u_{\mathrm{rel}}^2(w)} \tag{9.7}$$

Equation (9.7) has a solution, which is the detection limit $y^{\#}$, if, with preselection of counts, the following condition is satisfied

$$k_{1-\beta} \cdot \sqrt{\frac{1}{n_{\mathrm{g}}} + u_{\mathrm{rel}}^2(w)} < 1 \tag{9.8}$$

A non-existence of the detection limit points to the fact that there are too large relative uncertainties of the quantities X_5 to X_m, summarily expressed by $u_{\mathrm{rel}}(w)$, and that the GUM approximation is not sufficient. In such a case the solution is to obtain the necessary information to allow for an evaluation of the uncertainty on the basis of distributions according to GUM Supplement 1 and to proceed according to ISO 11929-2:2019.

For **counting measurements with unknown random influences** with the model of evaluation $y = (x_1 - x_2 x_3 - x_4) \cdot \dfrac{x_6 x_8 \cdots}{x_5 x_7 \cdots} = (x_1 - x_2 x_3 - x_4) \cdot w$ with $x_1 = \overline{n}_{\mathrm{g}} / t_{\mathrm{g}}$; $x_2 = \overline{n}_0 / t_0$ the detection limit $y^{\#}$ can only be obtained by applying the interpolation formula according to Equation (7.1).

$$\tilde{u}^2(\tilde{y}) = \tilde{u}^2(0) \cdot (1 - \tilde{y}/y_1) + u^2(y_1) \cdot \tilde{y}/y_1 \tag{9.9}$$

If the results y_0, y_1 and y_2 as well as the associated standard uncertainties $u(y_0)$, $u(y_1)$, and $u(y_2)$ from three measurements are available, the following bilinear interpolation can be used

$$\tilde{u}^2(\tilde{y}) = u^2(y_0) \cdot \frac{(\tilde{y} - y_1)(\tilde{y} - y_2)}{(y_0 - y_1)(y_0 - y_2)} + u^2(y_1) \cdot \frac{(\tilde{y} - y_0)(\tilde{y} - y_2)}{(y_1 - y_0)(y_1 - y_2)} + u^2(y_2) \cdot \frac{(\tilde{y} - y_0)(\tilde{y} - y_1)}{(y_2 - y_0)(y_2 - y_1)} \tag{9.10}$$

If results from many similar measurements are given, the parabolic shape of the function $\tilde{u}^2(\tilde{y})$ can also be determined by an adjustment calculation.

The linear interpolation according to Formula (7.1) leads to the approximation

$$y^{\#} = a + \sqrt{a^2 + (k_{1-\beta}^2 - k_{1-\alpha}^2)\,\tilde{u}^2(0)} \tag{9.11}$$

with

$$a = k_{1-\alpha}\,\tilde{u}(0) + \frac{1}{2}\left\{ \left(k_{1-\beta}^2 / y_1 \right) \left[(u^2(y_1) - \tilde{u}^2(0) \right] \right\} \tag{9.12}$$

For **counting measurements with known random influences** with the model of evaluation $y = (x_1 - x_2 x_3 - x_4) \cdot \frac{x_6 x_8 \ldots}{x_5 x_7 \ldots} = (x_1 - x_2 x_3 - x_4) \cdot w$ the detection limit is given with the help of equation (7.9) implicitly by

$$y^{\#} = y^{*} + k_{1-\beta} \sqrt{y^{\#2} \left[\frac{\vartheta^2}{m_\mathrm{g}} + u_\mathrm{rel}^2(w) \right] + \frac{y^{\#} w}{m_\mathrm{g}} \left(\frac{2\overline{n}_0 \vartheta^2}{t_0} + \frac{1}{t_\mathrm{g}} \right) + w^2 \left(\frac{\overline{n}_0}{m_\mathrm{g} t_0 t_\mathrm{g}} + \frac{\overline{n}_0^2 \vartheta^2}{m_\mathrm{g} t_0^2} + \frac{\overline{n}_0 + \vartheta^2 \overline{n}_0^2}{m_0 t_0^2} \right)} \quad (9.13)$$

For **black box measurements** with the model of evaluation $y = (\overline{y}_\mathrm{g} - \overline{y}_\mathrm{b}) \cdot \mathrm{w}$ with $\overline{y}_\mathrm{g} = \frac{1}{n_\mathrm{g}} \sum_{i=1}^{n_\mathrm{g}} y_{\mathrm{g},i}$, $\overline{y}_\mathrm{b} = \frac{1}{n_\mathrm{b}} \sum_{i=1}^{n_\mathrm{b}} y_{\mathrm{b},i}$, and $s_\mathrm{b} = \left(\frac{1}{n_\mathrm{b} - 1} \sum_{i=1}^{n_\mathrm{b}} (y_{\mathrm{b},i} - \overline{y}_\mathrm{b})^2 \right)^{1/2}$ the detection limit $y^{\#}$ can only be obtained by applying the interpolation formula according to Equation (7.1) as described above for counting measurements with unknown random influences.

9.3 The detection limit with uncertainties according to the GUM S1

The detection limit $y^{\#}$ is the smallest true value of the measurand, for which, by applying the decision rule according to Chapter 8.2, the probability of the wrong decision, that the physical effect is absent if it is present, does not exceed the specified probability β.

In order to find out whether a measurement procedure is suitable for the measurement purpose, the detection limit $y^{\#}$ is compared with the specified guideline value y_r of the measurand. The detection limit $y^{\#}$ is the smallest true value of the measurand which can be detected with the measurement procedure used. The detection limit $y^{\#}$ is obtained as the smallest solution of Equation (9.14)

$$P(y < y^{*} \,|\, \tilde{y} = y^{\#}) = \int_{-\infty}^{y^{*}} f_{Y,0}(y' \,|\, \mathbf{a}', \tilde{y} = y^{\#}) \, \mathrm{d}y' = \beta \quad (9.14)$$

Using the Monte Carlo approach, the decision threshold y^{*} is calculated by iteration in form of a root-finding problem which can be solved by bisection methods, *regula falsi* or interpolation. For each iteration step, a new probability distributions $f_{X_1}(\tilde{x}_1 \,|\, \mathbf{a}', x_1)$ has to be established with a modified value x_1 of the gross quantity X_1. Then, $i = 1, \ldots, n_\mathrm{M}$ new Monte Carlo trials are performed by drawing sets $x_{1,i}, \ldots, x_{n,i}$ from the probability distributions $f_{X_1}(\tilde{x}_1 \,|\, \mathbf{a}', x_1), f_{X_i}(\tilde{x}_i \,|\, \mathbf{a}'), i = 2, \ldots, n$. For each of these sets, one again calculates $y_i = G(x_{1,i}, \ldots, x_{n,i})$. The new vector $\mathbf{y}_\mathrm{M}(\mathbf{a}') = \{ y_1, \ldots, y_{n_\mathrm{M}} \}$ ordered ascendingly and to whose elements y_i afterwards cumulative probabilities i / n_M are assigned is a discrete representation of the distribution function $F_Y(\tilde{y} \,|\, \mathbf{a}', x_1) = \int_{-\infty}^{\tilde{y}} f_Y(\eta \,|\, \mathbf{a}', x_1) \, \mathrm{d}\eta$ of Y. From this new $\mathbf{y}_\mathrm{M}(\mathbf{a}') = \{ y_1, \ldots, y_{n_\mathrm{M}} \}$, one calculates the probability $P(y_i < y^{*})$. This procedure is repeated

until the condition $P(y_i < y^*) - \beta = 0$ is met and one obtains a vector $\mathbf{y}_{\mathrm{M}}(\mathbf{a}') = \left\{ y_1, ..., y_{n_{\mathrm{M}}} \right\}$ which fulfils this condition. The arithmetic mean of this vector $\mathbf{y}_{\mathrm{M}}(\mathbf{a}') = \left\{ y_1, ..., y_{n_{\mathrm{M}}} \right\}$ is the detection limit $y^{\#}$.

If the probability distribution $f_{Y,0}(y' | \tilde{y})$ is explicitly available, the integral according to Equation (9.14) can be evaluated by any suitable mathematical means.

9.4 Assessment of a measurement procedure

The decision on whether or not a measurement procedure to be applied sufficiently satisfies the requirements regarding the detection of the physical effect quantified by the measurand is made by comparing the detection limit $y^{\#}$ with the specified guideline value y_{r}. If $y^{\#} > y_{\mathrm{r}}$, the measurement procedure is not suitable for the intended measurement purpose with respect to the requirements.

To improve the situation in the case of $y^{\#} > y_{\mathrm{r}}$, it can often be sufficient to choose longer measurement durations or to preselect more counts of the measurement procedure. This reduces the detection limit.

Note that, occasionally, it can happen that a primary measurement result is larger than the decision threshold, i.e. $y > y^*$, and thus an effect of the sample is recognized, but the detection limit is larger than the specified guideline, i.e. $y^{\#} > y_{\mathrm{r}}$. This is, for instance, the case if due to particular circumstances the background counting rate is too high and at the same time the contribution from the sample is also high. If the primary measurement result y and its associated standard uncertainty $u(y)$ conform to the measurement objective the result can be accepted although formally the criterion of the detection limit in comparison with the guideline value is not fulfilled.

If Equations (9.3) or (9.5) have no solution $y^{\#}$ the GUM approximation, which provides a basis for Part 1 of ISO 11929:2019, is not sufficient to evaluate the measurement uncertainties. In this case the GUM S1 has to be used for the evaluation of the uncertainties and the characteristic limits can be obtained as stipulated in ISO 11929-2:2019.

It is important to note that frequently the detection limit is used in the wrong way, i.e. that the detection limit is compared with the primary measurement result and used to answer the question whether or not a contribution from the sample has been recognized. This is fundamentally wrong. The decision threshold has to be compared with the primary measurement result to decide this question. The detection limit is exclusively meant to characterize the detectability or sensitivity of a measurement procedure.

10 The limits of coverage intervals

10.1 General aspects and the definition of coverage intervals

Coverage intervals are frequently used in metrology. The limits of a coverage interval $y^{\triangleleft}, y^{\triangleright}$ define an interval $[y^{\triangleleft}, y^{\triangleright}]$ which contains the true value $\tilde{y}$ of the measurand Y with a pre-selected probability $(1 - \gamma)$.

$$P(y^{\triangleleft} < \tilde{y} < y^{\triangleright} \,|\, y, \mathfrak{I}) = \int_{y^{\triangleleft}}^{y^{\triangleright}} f_Y(\tilde{y} \,|\, y, \mathfrak{I})\, \mathrm{d}\tilde{y} = 1 - \gamma \tag{10.1}$$

Frequently, the probabilistically symmetric coverage interval is used with the additional conditions

$$\begin{aligned} P(\tilde{y} < y^{\triangleleft} \,|\, y, \mathfrak{I}) &= \int_{-\infty}^{y^{\triangleleft}} f_Y(\tilde{y} \,|\, y, \mathfrak{I})\, \mathrm{d}\tilde{y} = \gamma / 2 \\ P(\tilde{y} > y^{\triangleright} \,|\, y, \mathfrak{I}) &= \int_{y^{\triangleright}}^{\infty} f_Y(\tilde{y} \,|\, y, \mathfrak{I})\, \mathrm{d}\tilde{y} = \gamma / 2 \end{aligned} \tag{10.2}$$

Note that the definition of a coverage interval is not unique and that different types of coverage interval can be defined for the same probability $(1 - \gamma)$. For example, the shortest coverage interval $[y^{<}, y^{>}]$ can be used alternatively

$$P(y^{<} < \tilde{y} < y^{>} \,|\, y) = \int_{y^{<}}^{y^{>}} f_Y(\tilde{y} \,|\, y, \mathfrak{I})\, \mathrm{d}\tilde{y} = 1 - \gamma \text{ with } y^{>} - y^{<} = \min \tag{10.3}$$

See Weise et al. (2013) for detailed formulas in the case of the GUM approximation. For the general case of multiple output quantities see the GUM S2.

The term coverage interval which occasionally also is called credible interval has to be distinguished from the term confidence interval used in frequentist statistics since the meanings of the two terms are different. The coverage interval contains the true value $\tilde{y}$ of the measurand Y with a pre-selected probability $(1 - \gamma)$. The confidence interval is an interval in which the result of a new experiment is expected with the probability $(1 - \gamma)$ given a true value $\tilde{y}$ of the measurand Y. These two definitions make quite some difference.

10.2 Coverage intervals according to the GUM

With a primary measurement result y of the measurand and the standard uncertainty $u(y)$ associated with y, the lower limit of the probabilistically symmetric coverage interval $y^{\triangleleft}$ and the upper limit of the probabilistically symmetric coverage interval $y^{\triangleright}$ are calculated by

$$y^{\triangleleft} = y - k_p \cdot u(y) \text{ with } p = \omega \cdot (1 - \gamma/2) \tag{10.4}$$

$$y^{\triangleright} = y + k_q \cdot u(y) \text{ with } q = 1 - \omega \cdot \gamma / 2 \tag{10.5}$$

where

$$\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/u(y)} \exp(-\frac{v^2}{2}) \, \mathrm{d}v = \Phi\left[y / u(y)\right] \tag{10.6}$$

For the distribution function $\Phi(t)$ of the standard normal distribution and for its inversion, $k_p = t$ for $\Phi(t) = p$ see ISO 11929-1:2019 (ISO 2019a).

In general, the limits of the **probabilistically symmetric coverage interval** are located neither symmetrical to y, nor to the best estimate $\hat{y}$ but the probabilities of the measurand being smaller than $y^{\triangleleft}$ or larger than $y^{\triangleright}$ both equal $\gamma/2$. The relations $0 < y^{\triangleleft} < y^{\triangleright}$ apply.

$\omega = 1$ may be set if $y > 4u(y)$. In this case, the following approximations symmetrical to y apply

$$y^{\triangleleft} = y - k_{1-\gamma/2} \cdot u(y) \text{ and } y^{\triangleright} = y + k_{1-\gamma/2} \cdot u(y) \tag{10.7}$$

and the result may be expressed as $y \pm k_{1-\gamma/2} \cdot u(y)$. See Fig. 20 for an example.

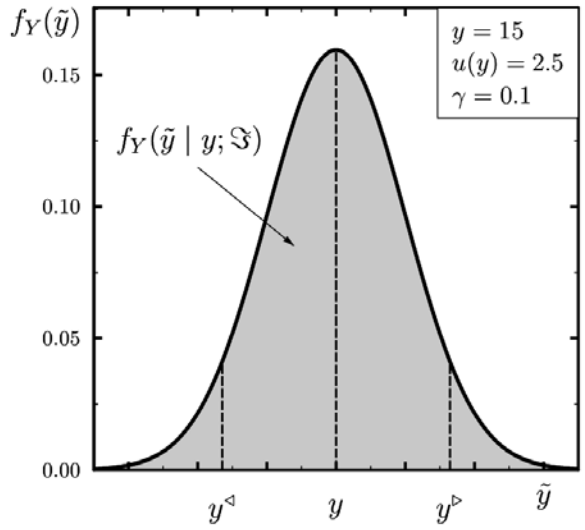


Fig. 20: Example of a probabilistically symmetric coverage interval. The shaded area represents the truncated Gaussian distribution.

As described in detail by Weise et al (2013), the lower limit of the **shortest coverage interval** $y^{<}$ and the upper limit of the shortest coverage interval $y^{>}$ are calculated from a primary measurement result y of the measurand and the standard uncertainty $u(y)$ associated with y either by

$$y^{<}, y^{>} = y \pm k_p \cdot u(y) \quad p = (1 + \omega \cdot (1 - \gamma)) / 2 \tag{10.8}$$

or if $y^{<} < 0$

$$y^{<} = 0 \quad y^{>} = y + k_q \cdot u(y) \quad q = 1 - \omega \cdot \gamma \tag{10.9}$$

with ω given by Equation (10.6). The relations $0 \le y^{<} < y^{>}$ apply and the approximation of Equation (10.7) is valid.

Fig. 21 gives examples of the probabilistically symmetric and the shortest coverage interval for a case with a large uncertainty of the primary measurement result.

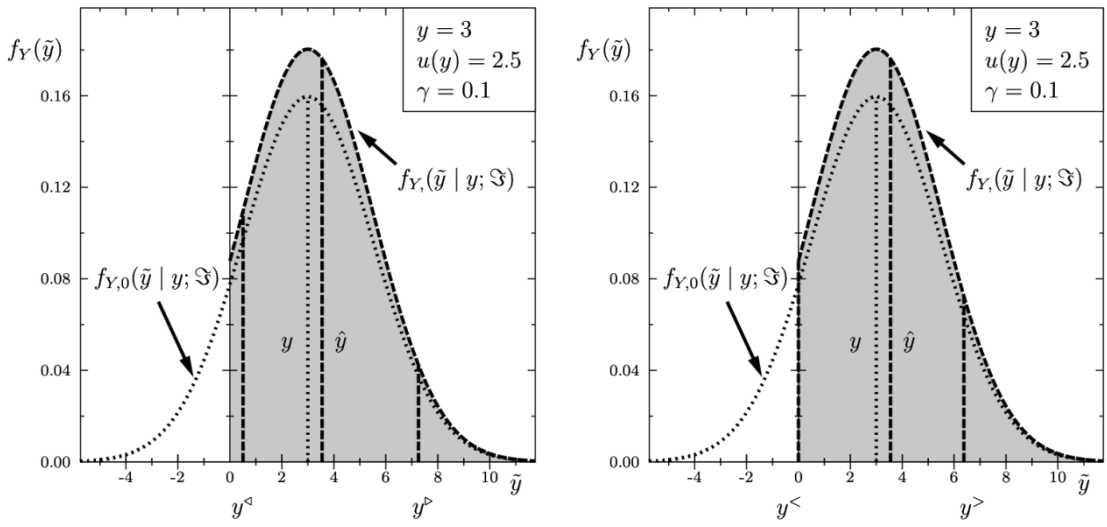


Fig. 21: Example of a probabilistically symmetric coverage interval (left) and a shortest coverage interval (right). The shaded areas represent the truncated Gaussian distributions.

10.3 Coverage intervals according to the GUM S1

The probabilistically symmetric coverage interval (Fig. 21) includes for a result y of a measurement, which exceeds the decision threshold y^*, the true value of the measurand with a probability $1-\gamma$. It is enclosed by the lower and upper limit of the symmetric coverage interval, respectively $y^{\triangleleft}$ and $y^{\triangleright}$, derived as $(1-\gamma/2)$-quantiles of the probability distribution $f_Y(\tilde{y}|\mathbf{a})$ of the true value given the experimental result and the prior knowledge that the measurand is non-negative. They are calculated as the upper and lower $\gamma/2$ quantile of the posterior probability distribution which takes into account that the measurand is non-negative (Equation (10.10))

$$\int_{-\infty}^{y^{\triangleleft}} f_Y(\eta|\mathbf{a}, Y \ge 0)\,\mathrm{d}\eta = \gamma/2 \tag{10.10}$$

$$\int_{y^{\triangleright}}^{\infty} f_Y(\eta|\mathbf{a}, Y \ge 0)\,\mathrm{d}\eta = \gamma/2 \tag{10.11}$$

Using the Monte Carlo approach, the limits of the probabilistically symmetric coverage interval $y^{\triangleleft}$ and $y^{\triangleright}$ are the $q_{1-\gamma/2}$ quantiles of the probability distribution $f_Y(\tilde{y}|\mathbf{a})$ represented by the vector $\mathbf{y}_{\mathrm{M}} = \{y_1, ..., y_{n_{\mathrm{M}}}\}$ taking into account that the measurand is non-negative. These quantiles can be conveniently calculated from the vector $\mathbf{y}_{\mathrm{M}} = \{y_1, ..., y_{n_{\mathrm{M}}}\}$ by searching the values $y_{k_1} = y^{\triangleleft}$ and $y_{k_2} = y^{\triangleright}$ with the conditions $y_i \geq 0$, $k_1 / n_{\mathrm{M},1} = \gamma / 2$ and $k_2 / n_{\mathrm{M},1} = 1 - \gamma / 2$; $n_{\mathrm{M},1}$ is the number of $y_i \geq 0$.

If the probability distribution $f_Y(\tilde{y}|\mathbf{a})$ is explicitly available the quantiles according to Equation (10.2) can be evaluated by any suitable means.

11 The best estimate and its associated standard uncertainty

11.1 Definition of the best estimate and its associated standard uncertainty

In the example of Fig. (21) the following holds: If the relative standard uncertainty associated with the measurement result exceeds 25 %, a substantial part of the PDF $f_{Y,0}(\tilde{y}|y)$ will extend to negative values $\tilde{y}$ which is physically not meaningful since the quantity Y is non-negative. The primary measurement result y, and the expectation of $f_{Y,0}(\tilde{y}|y)$, may be negative. Taking, however, into account the prior knowledge that the quantity Y is non-negative, as expressed by the Heaviside function $\mathrm{H}(\tilde{y})$, the best estimate and its associated standard uncertainty can be calculated by

$$\begin{aligned} \hat{y} &= \mathrm{E}(f_Y(\tilde{y}|y, Y \geq 0)) = \int_{-\infty}^{+\infty} \tilde{y} \cdot f_Y(\tilde{y}|y, Y \geq 0)\, \mathrm{d}\tilde{y} \\ &= \int_{-\infty}^{+\infty} \tilde{y} \cdot C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})\, \mathrm{d}\tilde{y} \end{aligned} \tag{11.1}$$

$$\begin{aligned} u^2(\hat{y}) &= \mathrm{Var}(f_Y(\tilde{y}|y, Y \geq 0)) = \int_{-\infty}^{+\infty} (\tilde{y} - \hat{y})^2 \cdot f_Y(\tilde{y}|y, Y \geq 0)\, \mathrm{d}\tilde{y} \\ &= \int_{-\infty}^{+\infty} (\tilde{y} - \hat{y})^2 \cdot C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})\, \mathrm{d}\tilde{y} \end{aligned} \tag{11.2}$$

The PDF $C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$ with the normalisation constant C is the posterior PDF of the complete evaluation of the measurement.

11.1 The best estimates according to the GUM

If the primary measurement result y exceeds the decision threshold y^* also the best estimate $\hat{y}$ and its associated uncertainty $u(\hat{y})$ can be calculated. In contrast to the primary result y

and its associated uncertainty $u(y)$ the best estimate $\hat{y}$ and its associated uncertainty $u(\hat{y})$ take into account that the measurand is non-negative.

If the GUM is used for the evaluation of the uncertainties, the best estimate and its associated uncertainty are calculated by

$$\hat{y} = y + \frac{u(y)\cdot \exp\left\{-y^2 \Big/ \left[2u^2(y)\right]\right\}}{\omega\sqrt{2\pi}} \quad \text{and} \quad u(\hat{y}) = \sqrt{u^2(y) - (\hat{y} - y)\,\hat{y}} \tag{11.3}$$

For $y \geq 4u(y)$, the approximations $\hat{y} = y$ and $u(\hat{y}) = u(y)$ are sufficient and a separate calculation of the best estimate $\hat{y}$ and its associated uncertainty $u(\hat{y})$ is not necessary.

The relations $0 \leq y^{<} < y^{\lhd} < y \leq \hat{y} < y^{>} \leq y^{\rhd}$ as well as $u(\hat{y}) < u(y)$ and $u(\hat{y}) < \hat{y}$ apply. Moreover, for $y \geq 4u(y)$, the approximations

$$\hat{y} = y \;;\; u(\hat{y}) = u(y) \tag{11.4}$$

are sufficient.

If the best estimate $\hat{y}$ and its associated standard uncertainty $u(\hat{y})$ are calculated, the recording of the primary measurement result y and its associated standard uncertainty $u(y)$ may be omitted.

If the decision rule defined by the decision threshold is not used and if $y < y^*$, the best estimate $\hat{y}$ and its standard uncertainty $u(\hat{y})$ can be calculated anyway.

11.2 The best estimates according to the GUM S1

The determined primary measurement result y of the measurand shall be compared with the decision threshold y^*. If $y > y^*$, the physical effect quantified by the measurand is recognized as present. Otherwise, it is decided to conclude that the effect is absent.

If $y \geq y^*$ the best estimate $\hat{y}$ of the measurand and its associated standard uncertainty $u(\hat{y})$ are given by

$$\hat{y} = \mathrm{E}(f_Y(\tilde{y}\,|\,\mathbf{a}, Y \geq 0)) = \int_{-\infty}^{+\infty} \eta \cdot f_Y(\eta\,|\,\mathbf{a}, Y \geq 0)\,\mathrm{d}\eta \tag{11.5}$$

$$u^2(\hat{y}) = \mathrm{Var}(f_Y(\tilde{y}\,|\,\mathbf{a}, Y \geq 0) = \int_{-\infty}^{+\infty} (\eta - \hat{y})^2 \cdot f_Y(\eta\,|\,\mathbf{a}, Y \geq 0)\,\mathrm{d}\eta\,. \tag{11.6}$$

If the uncertainties were evaluated according to the ISO/IEC Guide 98-3-1, the best estimate and its associated standard uncertainty have to be determined from Monte Carlo calculations and to be calculated by explicitly solving the integrals of Equations (11.5) and (11.6).

If the probability distribution $C \cdot f_{Y,0}(\tilde{y}|y) \cdot \mathrm{H}(\tilde{y})$ is explicitly available the integrals according to Equations (11.5) and (11.6) can be evaluated by any suitable means.

The best estimate and its associated standard uncertainty are independent of the type of coverage interval used.

If the best estimate $\hat{y}$ and its standard uncertainty $u(\hat{y})$ are calculated, the recording of the primary measurement result y and its standard uncertainty $u(y)$ can be omitted.

Using the Monte Carlo approach for the calculation of the best estimate and its associated uncertainty as well as for the limits of the coverage intervals, the original set $\mathbf{y}_\mathrm{M} = \{y_1, ..., y_{n_\mathrm{M}}\}$ representing the distribution function $F_Y(\tilde{y}|\mathbf{a}) = \int_{-\infty}^{\tilde{y}} f_Y(\eta|\mathbf{a})\,\mathrm{d}\eta$ according to section 6.3, is used considering the condition that the measurand is non-negative. The best estimate and its associated standard uncertainty are then calculated as the arithmetic mean and the standard deviation of the $y_i \geq 0$.

12 Conformity with requirements

12.1 General aspects

This chapter extends a previous paper on "Measuring, Estimating, and Deciding under Uncertainty" (Michel 2016) with respect to the problem of conformity assessments based on an actual recommendation (SSK 2016) of the German Commission on Radiological Protection (SSK; www.ssk.de); see also Michel (2017, 2019).

The simplest decision problem is to decide the question whether or not two measurement results conform to each other. As described by Weise and Wöger (1994) the decision can be made by the criterion for conformity

$$|y_1 - y_2| \leq \beta \cdot u(y_1 - y_2) = \beta \cdot \sqrt{u^2(y_1) + u^2(y_2)} \tag{12.1}$$

However, these authors correctly just state that the factor β which is representing a probability of conformity might be in the range 1 to 3, with $\beta = \sqrt{2}$ being the most favorable from a Bayesian viewpoint. The decision about the choice of β remains a task for the user.

The JCGM also made recommendations for assessments of conformity with requirements taking into account measurement uncertainties (JCGM 2012b). The German Commission on Radiological Protection (SSK) has used the JCGM 106 (JCGM 2012b) recommendation to give explicit advice how to assess conformity with requirements in ionizing radiation measurements for the purpose of radiological protection (SSK 2016).

12.2 Tolerance and acceptance intervals

Requirements for technical testing result from safety regulations according to the state of science and technology and from legal limits. Such requirements can be defined for various parameters (measurands) in the form of one- or two-sided limited tolerance intervals $[T_L, T_U]$ (Fig. 22).

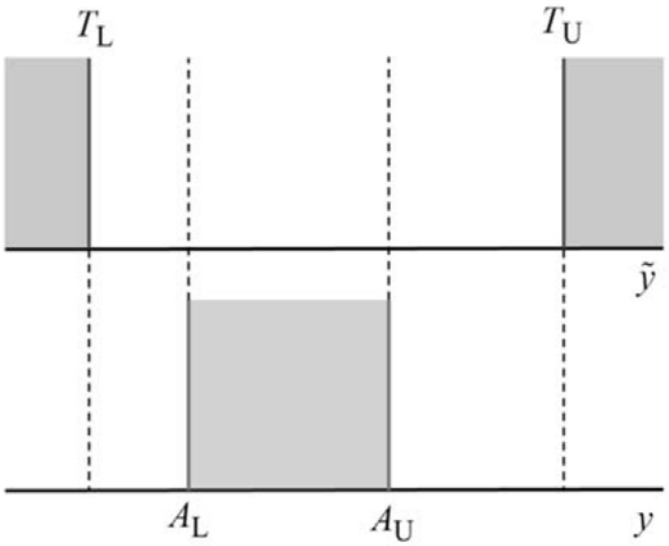


Fig. 22: Two-sided tolerance interval $[T_L, T_U]$ for the true value of the measurand $\tilde{y}$ and acceptance interval $[A_L, A_U]$ for measurement results y.

For ease of practical application also an acceptance interval $[A_L, A_U]$ may be defined in a way that one decides in favor of conformity with the requirement set by $[T_L, T_U]$ independent of the measurement uncertainty if the measured value lies within the acceptance interval $[A_L, A_U]$. Consequently, the acceptance interval lies inside the tolerance interval and is shorter than the former and will be located asymmetrically relative to the tolerance interval (Fig. 22). Setting the limits of an acceptance interval needs the knowledge of the standard uncertainty associated with a measurement result as a function of the measurement result. Fig. 22 gives a graphical representation of the concepts of the tolerance and acceptance intervals taking a two-sided tolerance interval as an example.

A proof of conformity is then based on measurements, the results of which have uncertainties associated with them and which have to be taken into account in the assessment. The uncertainties referred to are understood as standard uncertainties $u(y)$ according to the GUM or GUM S1 associated with the measurement result y. A decision in favor of conformity with the requirement is made if the true value of the measurand lies within the tolerance interval with a preselected probability. With the concept of measurement uncertainties according to the GUM and GUM S1 decision rules can be derived by means of the limits of the probabilistically symmetric coverage interval of the PDF $f_Y(\tilde{y}|y, \Im)$.

A requirement is fulfilled, i.e. the measurement result conforms to the requirement,

1. if – in case of a one-sided downward tolerance interval $[T_L,\infty]$ – the lower limit $y^{\triangleleft}$ of the probabilistically symmetric coverage interval $\left[y^{\triangleleft}, y^{\triangleright}\right]$ for the coverage probability 90 % is larger than the limit of the tolerance interval T_L ;

2. if – in case of a one-sided upward limited tolerance interval $[0,T_U]$ – the upper limit $y^{\triangleright}$ of the probabilistically symmetric coverage interval $\left[y^{\triangleleft}, y^{\triangleright}\right]$ for the coverage probability 90 % is smaller than the limit of the tolerance interval T_U ;

3. if – in case of a two-sided tolerance interval $[T_L,T_U]$ – the probabilistically symmetric coverage interval $\left[y^{\triangleleft}, y^{\triangleright}\right]$ for the coverage probability 95 % lies within the tolerance interval $[T_L,T_U]$.

If the GUM is used for the quantification of measurement uncertainties, the limits of the coverage interval can be calculated by means of the standard measurement uncertainties and quantiles of the standard normal distribution as stipulated in ISO 11929-1:2019 and described in chapter 5.3. If the GUM S1 is applied, numerical procedures are needed for the calculation of the limits of the coverage interval.

If requirements are stipulated by definition of a one- or two-sided tolerance interval, it is demanded that the true value of the measurand lies in the tolerance interval. Since the true value of the measurand is unknown and unknowable only probability statements can be made about it taking into account the measurement uncertainty.

Since only probability statements can be made about the true value of a measurand, there is the possibility to wrongly decide on the conformity with requirements. While the scientific judgement about a measurement objective can only give boundary conditions about an acceptable or tolerable probability for a wrong decision, the setting of this probability can only be performed by the regulator. It is the result of societal agreement.

The German Commission on Radiological Protection (SSK), which is consulting the German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety, recommended setting the probability for a correct decision in favor of conformity with requirements to 95 %. Consequently, the probability for a wrong decision in favor of conformity with a requirement is 5 %. The latter value of 5 % is frequently cited in international standards. Given the limited possibility to calculate reliably very small percentiles of an estimated PDF, the setting of 5 % for a wrong decision appears to be meaningful and justified.

12.3 Conformity with a tolerance interval

If the requirement is specified by a one-sided upward limited tolerance interval $[0,T_U]$ with an upper limit T_U below which the true value of the measurand shall lie with a high probability, one shall decide in favor of conformity if the upper limit of the probabilistically symmetric coverage interval for the coverage probability of 90 % is smaller than the limit T_U

$$P(\tilde{y} > T_U \,|\, y, u(y)) \leq 0,05 \tag{12.2}$$

If the GUM is used for quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition for conformity reads

$$y + 1,65 \cdot u(y) \leq T_U \tag{12.3}$$

For larger relative measurement uncertainties the limits of the coverage interval shall be calculated according to ISO 11929-2:2019 as described in Equation (12.4).

If the requirement is specified by a one-sided downward limited tolerance interval $[T_L, \infty]$ with a lower limit T_L above which the true value of the measurand shall lie with a high probability, one shall decide in favor of conformity if the lower limit of the probabilistically symmetric coverage interval for the coverage probability of 90 % is smaller than the limit T_L

$$P(\tilde{y} < T_L \,|\, y, u(y)) \leq 0,05 \tag{12.4}$$

If the GUM is used for quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition for conformity reads

$$y - 1,65 \cdot u(y) \geq T_L \tag{12.5}$$

For larger relative measurement uncertainties the limits of the coverage interval shall be calculated according to ISO 11929-2:2019 as described in Equation (12.4).

As an explanation of the rules given by Equations (12.2) to (12.5), Fig. 23 shows the extreme cases by which it is ensured that the coverage probability of 90 % fixes the probability of a correct decision in favor of conformity with the requirement at 95 % and limits that for a wrong decision at 5 %.

The use of a coverage probability of only 90 % in the case of a probabilistically symmetric coverage interval results from the fact that just one side of the distribution has the chance that the true value of the measurand lies above respectively below the limits of the tolerance interval T_L respectively T_U. The other side of the distribution covers also outside of the coverage interval true values which conform to the requirement.

So, the factor 1,65 in Equation (12.5) is the one-sided quantile of the standard normal distribution for a probability of 5 %. It is sufficient here since the tolerance interval is just one-sided. This is different for a two-sided tolerance interval, where the two-sided quantile of the standard normal distribution for a probability of 5 % has to be used. The latter quantile two sided quantile is 1,96 (Equation 12.7). See figure 9 in chapter 4.2 for an explanation of the different quantiles.

If requirements are defined in the regulations via a tolerance interval no further requirements regarding permissible magnitude of the measurement uncertainties are necessary.

If the requirement is specified by a two-sided tolerance interval $[T_L, T_U]$ in which the true value of the measurand shall lie with a high probability, one shall decide in favor of conformity if the probabilistically symmetric coverage interval for the coverage probability of 95 % lies in the tolerance interval

$$P(\tilde{y} < T_{\mathrm{L}} \vee \tilde{y} > T_{\mathrm{U}} \,|\, y, u(y)) \leq 0{,}05 \tag{12.6}$$

If the GUM is used for quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition for conformity reads

$$y - 1{,}96 \cdot u(y) \geq T_{\mathrm{L}} \text{ and } y + 1{,}96 \cdot u(y) \leq T_{\mathrm{U}} \tag{12.7}$$

For larger relative measurement uncertainties the limits of the coverage interval shall be calculated according to ISO 11929-2:2019 as described in Equation (12.6).

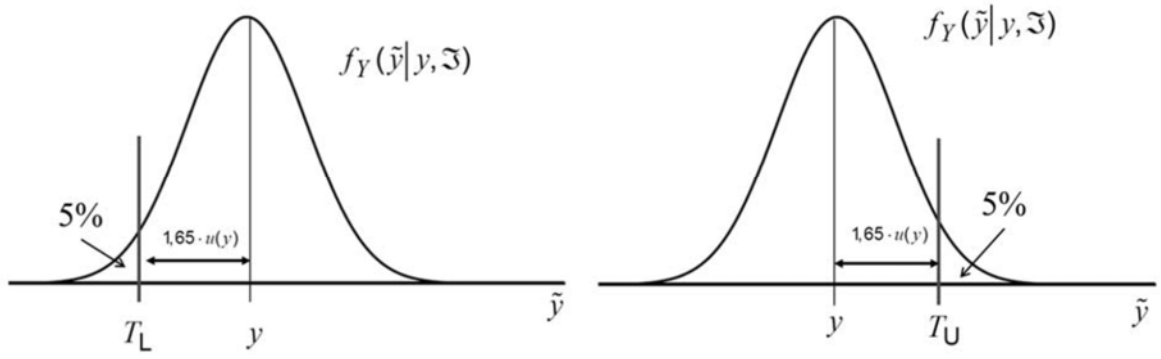


Fig. 23: Extreme cases of one-sided tolerance intervals. For one-sided tolerance intervals a coverage probability of 90 % for the probabilistically symmetric coverage interval is sufficient in order to limit the probability of wrong decisions in favor of conformity with a requirement to 5 %.

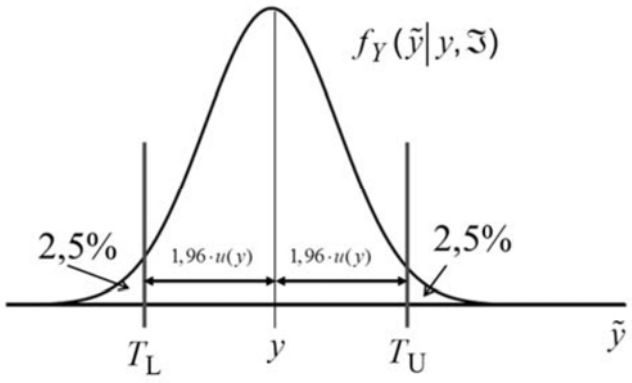


Fig. 24: Extreme case of a two-sided tolerance interval which makes it obligatory to apply a coverage probability of 95 %. In this extreme case, the uncertainty is so large that the probabilistically symmetric coverage interval for the coverage probability of 95 % completely fills the tolerance interval.

The setting of a coverage probability of 95 % is conservative and limits for any measurement relative uncertainty the probability of a wrong decision in favor of conformity with the requirement to 5 %. This is exemplified for a limiting case in Fig. 24, which shows the case of maximum possible standard uncertainty and the therefrom resulting demand for a coverage

probability of 95 %. A measurement procedure with a larger measurement uncertainty is not suitable to proof the conformity with the requirement.

The rules stated in chapter 12.3 are depicted in Fig. 25.

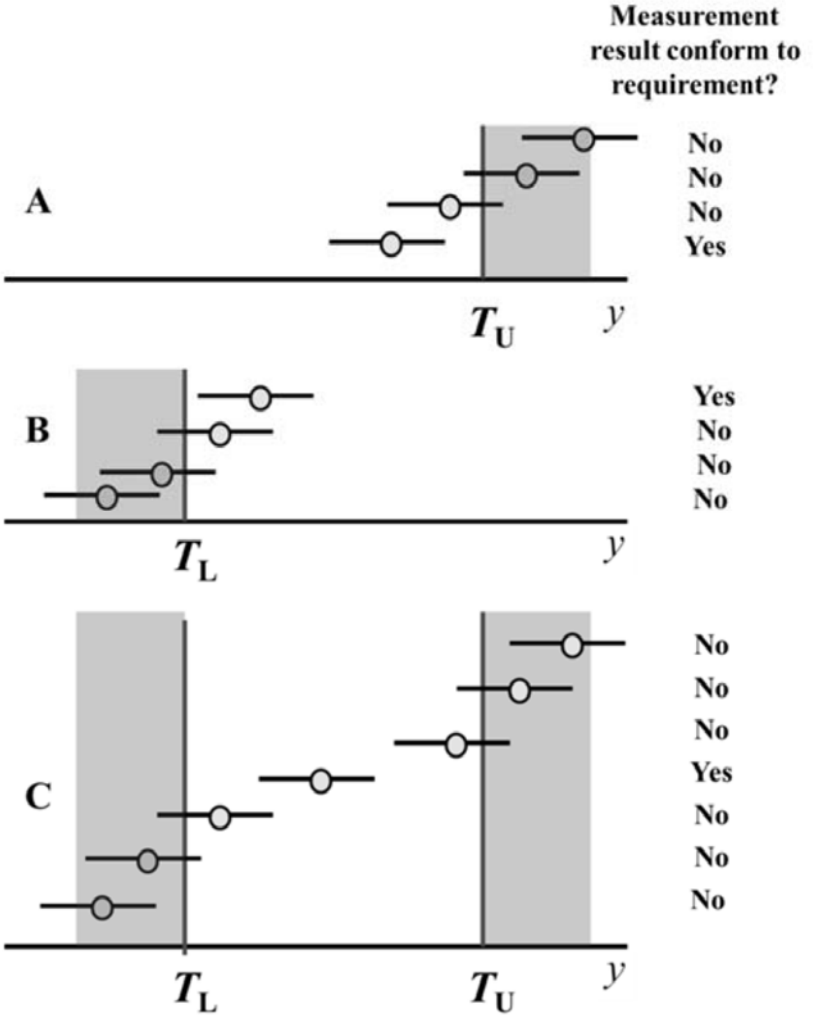


Fig. 25: Assessment of conformity with a requirement for a one-sided upward limited tolerance interval (A), for a one-sided downward limited tolerance interval (B), for a two-sided tolerance interval (C).

12.4 Acceptance interval inside a tolerance interval

The specification of an acceptance interval for a one-sided or two-sided tolerance interval requires the knowledge of the measurement uncertainty as a function of the measurement result and a sufficient reproducibility of the measurements and the measurement uncertainties.

For a one-sided downward limited tolerance interval $[T_L, +\infty]$ only the lower limit of the acceptance interval is needed. The acceptance interval $[A_L, \infty]$ is then given by the implicit Equation

$$P(\tilde{y} < T_L \,|\, y = A_L, u(y = A_L)) = 0,05 \tag{12.8}$$

If the GUM is used for quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition for the lower limit T_L of the acceptance interval reads (see Fig. 26)

$$A_L = T_L + 1,65 \cdot u(y = A_L) \tag{12.9}$$

For a one-sided upward limited tolerance interval $[0, T_U]$ only the upper limit A_U of the acceptance interval is needed. The acceptance interval $[0, A_U]$ is then given by the implicit Equations

$$P(\tilde{y} > T_U \,|\, y = K_U, u(y = K_U)) = 0,05 \tag{12.10}$$

If the GUM is used for the quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition for the upper limit of the acceptance interval reads (see Fig. 26)

$$A_U = T_U - 1,65 \cdot u(y = A_U) \tag{12.11}$$

For larger relative measurement uncertainties the limits of the coverage interval shall be calculated according to ISO 11929-2:2019 (ISO 2019).

For a two-sided tolerance interval $[T_L, T_U]$ both the lower and the upper limit of the acceptance interval are needed. The acceptance interval $[A_L, A_U]$ is then given by the implicit Equations

$$P(\tilde{y} < T_L \,|\, y = A_L, u(y = A_L)) = 0,025 \text{ and } P(\tilde{y} > T_U \,|\, y = A_U, u(y = A_U)) = 0,025 \tag{12.12}$$

If the GUM is used for quantifying the measurement uncertainty and if the relative measurement uncertainty is smaller than 25 % the condition reads (Fig. 27)

$$A_L = T_L + 1,96 \cdot u(y = A_L) \text{ and } A_U = T_U - 1,96 \cdot u(y = A_U) \tag{12.13}$$

For larger relative measurement uncertainties the limits of the coverage interval shall be calculated according to ISO 11929-2:2019.

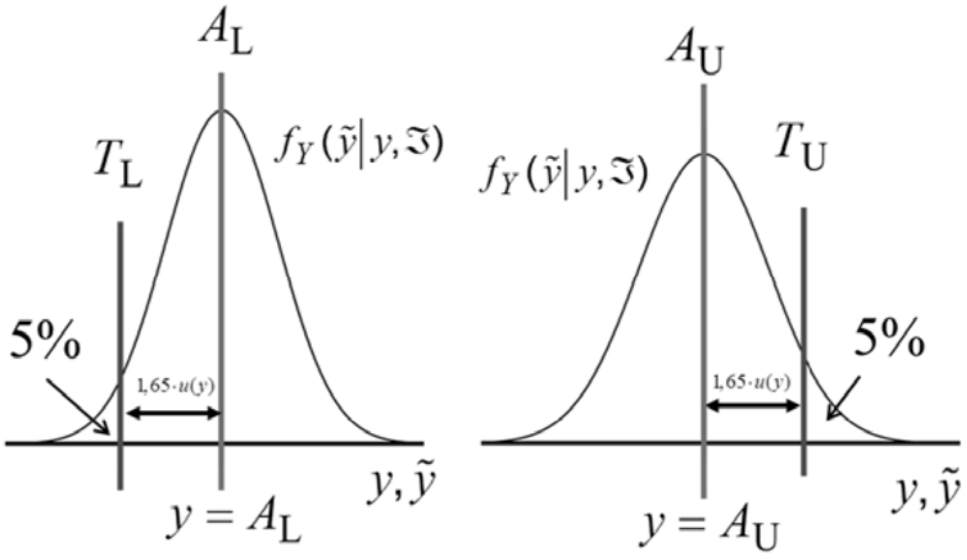


Fig. 26: Schematic diagram to explain the stipulations regarding limits of a acceptance interval for a one-sided downward limited (left) respectively upward limited (right) tolerance interval.

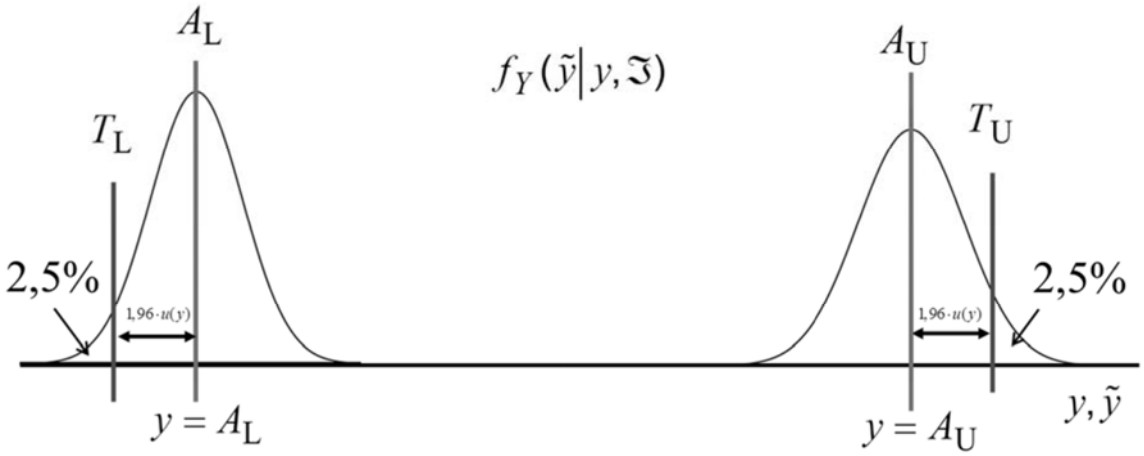


Fig. 27: Schematic diagram to explain the specifications for an acceptance interval in the case of a two-sided tolerance interval.

13 Epilogue

With the GUM standard series, there is an internationally accepted methodology to deal with uncertainty in measurements. With a Bayesian theory of measurement uncertainty, there exists a solid statistical foundation of the GUM standard series which has the splendor to reveal the normal methodology of human learning and improving of knowledge. The standard ISO 11929 extends the methodology of the GUM to decision making under uncertainty. Characteristic limits according to ISO 11929 can be calculated for practically every measurement problem. However, as metrology and the GUM are further developing, also ISO 11929 will have to be revised in the future to keep track with the development of metrology.

14 References

Altschuler, B, and B. Pasternack (1963) Statistical measures of the lower limit of detection of a radioactivity, Health Physics 9, 293 - 298.

Bayes, Th. (1763) Essay towards solving a problem in the doctrine of chances, The Philosophical Transactions 53,370–418; reprinted by G.A. Barnard; Thomas Bayes, 1958, Biometrika 45(3/4),293–315.

Berger, J. O. (1985) Statistical decision theory and Bayesian analysis. Springer, New York.

Berger, J.O., Bernardo, J.M., Sun D. (2009) The formal definition of reference priors, The Annals of Statistics Vol. 37, No. 2, 905–938 DOI: 10.1214/07-AOS587

Bernardo J. (1979) Reference prior distributions for Bayesian inference. J. R. Stat. Soc. B 41, 113 – 147.

Bernardo, J. M., A. F. M. Smith (1994) Bayesian Theory, Wiley, Chichester.

Bernardo J.M. (2003) Bayesian Statistics, in: Probability and Statistics (R. Viertl, ed.) Encyclopaedia of Life Support Systems (EOLSS). Oxford, UK: UNESCO.

Calmet, D. (2014) ISO Standards on test methods for radioactivity monitoring of food and the environment, Radiat. Emerg. Med., 3 (1), 7–20.

Calmet, D. et al. (2016); International Standards on Radioactivity Measurement for Radiological Protection: Status and Perspectives, Proc. IRPA14 Cape Town 2016, Radiation Protection Dosimetry (2016), pp. 1–8. doi:10.1093/rpd/ncw342.

CODATA (2017) D.B. Newell, F. Cabiati, J. Fischer, K. Fujii, S.G. Karshenboim, H.S. Margolis, E. de Mirandés, P.J. Mohr, F. Nez, K. Pachucki, T.J. Quinn, B.N. Taylor, M. Wang, B.M. Wood, and Z. Zhang, The CODATA 2017 values of h, e, k, and NA for the revision of the SI. Metrologia 55 (2018) L13–L16, https://iopscience.iop.org/article/10.1088/1681-7575/aa950a/pdf.

Cox, M., Harris, P., Nam, G., Thomas, D. (2006) The use of Monte Carlo method for uncertainty calculation , with an application to the measurement of neutron ambient dose equivalent rate. Radiation Protection Dosimetry 121, No. 1, 12–23 doi:10.1093/rpd/ncl098

Cox, R. T. (1946), Probability, frequency, and reasonable expectation, Am. J. Phys. 14, 1–13.

Currie, LA (1968) Limits for Qualitative Detection and Quantitative Determination: Application to Radiochemistry, Anal. Chem. 40, 586-593.

D'Agostini (2003) Bayesian Reasoning in Data Analysis: A critical introduction. https://www.roma1.infn.it/~dagos/WSPC/

DIN (1996) DIN 1319 Teil 3. Auswertung von Messungen einer einzelnen Messgröße, Messunsicherheit. Beuth Verlag, Berlin.

DIN (1999) DIN 1319 Teil 4. Behandlung von Unsicherheiten bei der Auswertung von Messungen. Beuth Verlag, Berlin.

DIN (2014) DIN ISO 11929 Beiblatt 1:2014-08 Bestimmung der charakteristischen Grenzen (Erkennungsgrenze, Nachweisgrenze und Grenzen des Vertrauensbereichs) bei Messungen ionisierender Strahlung – Grundlagen und Anwendungen; Beiblatt 1: Beispiele, Beuth Verlag, Berlin.

Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin (2014) Bayesian Data Analysis, Third Edition, CRC Press, Taylor & Francis Group, Boca Raton, http://bacbuc.hd.free.fr/WebDAV/data/Bouquins/Gelman %20- %20Bayesian %20Data %20Analysis.pdf.

Gregory, P.C. (2005) Bayesian Logical Data Analysis for the Physical Sciences, Cambridge University Press.

ISO (1993) ISO Guide to the Expression of Uncertainty in Measurement, corrected reprint, ISO, Geneva, corrected reprint 1995.

ISO (2010) ISO 11929:2010. Determination of characteristic limits (decision threshold, detection limit, and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and applications. ISO Geneva; DIN ISO 11929 Bestimmung der charakteristischen Grenzen (Erkennungsgrenze, Nachweisgrenze und Grenzen des Vertrauensbereichs) bei Messungen ionisierender Strahlung – Grundlagen und Anwendungen, DIN 2011.

ISO (2019a) ISO 11929-1:2019. Determination of characteristic limits (decision threshold, detection limit, and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and applications. Part 1: elementary applications, ISO Geneva; DIN ISO 11929-1:2020.

ISO (2019b) ISO 11929-2:2019. Determination of characteristic limits (decision threshold, detection limit, and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and applications. Part 2: Advanced applications, ISO Geneva; DIN ISO 11929-2:2020.

ISO (2019c) ISO 11929-3:2019. Determination of characteristic limits (decision threshold, detection limit, and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and applications. Part 3: Application to unfolding methods, ISO Geneva; DIN ISO 11929-3:2020.

ISO (2020) ISO 11929-4:2020. Determination of characteristic limits (decision threshold, detection limit, and limits of the confidence interval) for measurements of ionizing radiation - Fundamentals and applications. Part 4: Guideline to the application, ISO Geneva; DIN ISO 11929-4: 2021.

Jaynes, E. T. (1968), Prior probabilities, IEEE Trans. Systems Science and Cybernetics, SSC-4, 227–241. Reprinted in Tummala, V. M. Rao and Henshaw, R. C., eds., Concepts and Applications of Modern Decision Models, Michigan State University Business Studies Series (1976); and in Jaynes (1983).

Jaynes, E. T. (1983), in Rosenkrantz, R. D., ed., Papers on Probability, Statistics, and Statistical Physics, D. Reidel Publishing Co., Dordrecht, Holland; second paperbound edition, Kluwer Academic Publishers (1989). Reprints of 13 papers dated 1957–1980.

Jaynes, E.T. (2003) Probability theory – the logic of science, edited by G. Larry Bretthorst, Cambridge University Press, Cambridge/UK. https://bayes.wustl.edu/etj/prob/book.pdf

JCGM (2008a) ISO/IEC Guide 98-3:2008, Guide to the Expression of Uncertainty in Measurement (GUM). JCGM 100, http://www.bipm.org/en/publications/guides/gum.html.

JCGM (2008b) ISO/IEC Guide 98-3-1:2008, Guide to the expression of uncertainty in measurement (GUM) - Supplement 1: Propagation of distributions using a Monte Carlo method, JCGM 101, http://www.bipm.org/en/publications/guides/gum.html.

JCGM (2009) ISO/IEC Guide 98-1:2009, Uncertainty of measurement — An introduction to the "Guide to the expression of uncertainty in measurement, JCGM 104:2009

JCGM (2011) ISO/IEC Guide 98-3-2:2008, Guide to the expression of uncertainty in measurement (GUM) - Supplement 2: Extension to any number of output quantities, JCGM 102, http://www.bipm.org/en/publications/guides/gum.html.

JCGM (2012a) ISO/IEC Guide 99, International vocabulary of metrology – Basic and general concepts and associated terms (VIM), JCGM 200:2012; http://www.bipm.org/en/publications/guides/gum.html. Deutsche Version: DIN (2012) Internationales Wörterbuch der Metrologie. Grundlegende und allgemeine Begriffe und zugeordnete Benennungen (VIM). Deutsch-Englische Fassung ISO/IEC-Leitfaden 99:2007, korrigierte Fassung 2012, DIN Deutsches Institut für Normung e.V., 4. Auflage 2012

JCGM (2012b) ISO/IEC Guide 98-4:2012, Evaluation of measurement data – The role of measurement uncertainty in conformity assessment. Report JCGM 106:2012. http://www.bipm.org/en/publications/guides/gum.html.

JCGM (2019) ISO/IEC Draft Guide 98-3-SP3 - JCGM 103 CD Guide to the expression of uncertainty in measurement - Developing and using measurement models [constitutes draft ISO/IEC Guide 98-3-SP3: Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) — Supplement 3: Developing and using measurement models].

Jeffrey R. C. (1983), The Logic of Decision, 2nd edn, University of Chicago Press.

Jeffreys H. (1939), Theory of Probability, Clarendon Press, Oxford; later editions, 1948, 1961, 1967, 1988.

Jeffreys H. (1946) An invariant form for the prior probability in estimation problems, Proc. Roy. Soc. A 186, 453 – 461.

Jeffreys H. (1961), Theory of Probability, 3rd ed., London Oxford University Press

Kanisch, G (2015a): Generalized evaluation of environmental radioactivity measurements with UncertRadio Part I: Methods without linear unfolding http://dx.doi.org/10.1016/j.apradiso.2015.12.003.

Kanisch, G. (2015b) Generalized evaluation of environmental radioactivity measurements with UncertRadio Part II: Methods with linear unfolding, http://dx.doi.org/10.1016/j.apradiso.2015.12.046.

Kirkpatrick J.M., R. Venkataraman, B.M. Young (2012) Minimum detectable activity, systematic uncertainties, and the ISO 11929 standard, J. Radioanal. Nucl. Chem. DOI 10.1007/s10967-012-2083-5

Kolmogoroff, A. (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung; in: Ergebnisser der Mathematik und ihrer Grenzgebiete, herausgegeben von der Schriftleitung des „Zentralblatt für Mathematik“, 2. Band, Verlag Julius Springer, Berlin.

Laplace, P. S. (1812), Théorie Analytique des Probabilités, 2 vols., Courcier Imprimeur, Paris; 3rd edition with supplements, 1820. Reprinted in Laplace (1878–1912), Vol. 7. Reprints of this rare but very important work and the following one are available from: Editions Culture et Civilisation, 115 Ave. Gabriel Lebron, 1160 Brussels, Belgium.

Lee, P. M.; “Bayesian Statistics: An Introduction” (1989), Oxford University Press, New York.

Maxwell, J. C. (1850) in Harman, P. M. (ed.), 1990. The Scientific Letters and Papers of James Clerk Maxwell, Vol. 1, 1846 – 1862,197.

Mende, O. (1990) Untersuchung stochastischer Verfahren zur Festlegung von Nachweis- und Erkennungsgrenzen bei zählenden Kernstrahlungsmessungen, Diplomarbeit, Universität Hannover, 1990.

Michel R. (2000) Quality Assurance of Nuclear Analytical Techniques Based on Bayesian Characteristic Limits, Proceedings MTAA-10, J. Radioanal. Chem. 245:137 – 144.

Michel R. (2016) Measuring, Estimating, and Deciding under Uncertainty, Applied Radiation and Isotopes, pp. 6-11, DOI: 10.1016/j.apradiso.2015.12.013

Michel R. (2017) Uncertainty, detectability, and conformity in measurements of ionizing radiation, Japan Journal Health Phys. 52 (3) 179 – 191, DOI: 10.5453/jhps.52.179.

Michel R. (2019) Messunsicherheiten und Konformität mit Anforderungen. 9. VDI-Fachtagung Messunsicherheit 2019 - Messunsicherheit praxisgerecht bestimmen, Erfurt, 13. und 14. November 2019, VDI-Berichte 2365, pp. 241 – 256.

Michel R., Kirchhoff, K. (1999) Nachweis-, Erkennungs- und Vertrauensgrenzen bei Kernstrahlungsmessungen. Publikationsreihe Fortschritte im Strahlenschutz, Fachverbandes für Strahlenschutz, FS-99-108-AKSIGMA, ISSN 1013-4506, TÜV-Rheinland Verlag, Köln, https://www.fs-ev.org/fileadmin/user_upload/04_Arbeitsgruppen/06_Nachweisgrenzen/02_Dokumente/Arbeitsmaterialien/Michel_Kirchhoff_NWG_Buch_1999.pdf

Nicholson, W.L. (1963) Fixed time estimation of counting rates with background corrections, Hanford Laboratories, Richland.

PTB (2017) PTB-Infoblatt – Das neue Internationale Einheitensystem (SI) Stand 11/2017, Anhang: Einheiten und „ihre“ Konstanten.

Robert, C. P. (2001) The Bayesian Choice (2nd ed.), Springer, New York.

SSK (2016) Methodik zur Berücksichtigung von Messunsicherheiten bei messtechnischen Prüfungen im Geltungsbereich der Röntgenverordnung und der Strahlenschutzverord-

nung, Empfehlung der Strahlenschutzkommission. Verabschiedet in der 283. Sitzung der SSK am 15./16. September 2016. Veröffentlicht im BAnz AT 24.02.2017 B2.

Ullrich T., Xu, Z. (2007) Treatment of errors in efficiency calculations, http://arxiv.org/abs/physics/0701199v1.

Weise K. (1997) Bayesian-statistical detection limit, decision threshold, and confidence interval in nuclear radiation measurement, Kerntechnik 63:214–224.

Weise K. (2004) The Bayesian count rate probability distribution in measurement of ionizing radiation by use of a ratemeter, PTB Ra-44.

Weise K., Michel R. (1995) Erkennungsgrenze, Nachweisgrenze und Vertrauensbereich in der allgemeinen Kernstrahlungsspektrometrie, Kerntechnik 60 No. 4, 189 - 196

Weise K., Wöger W. (1992) Eine Bayessche Theorie der Messunsicherheit, PTB-Bericht N-11.

Weise K. , Wöger, W. (1993) A Bayesian theory of measurement uncertainty. Meas. Sci. Technol. 4,1-11.

Weise K., Wöger, W. (1994) Comparison of two measurement results using the Bayesian theory of measurement uncertainty. Meas. Sci. Technol. 5:879–882.

Weise K. ,Wöger, W. (1999) Messunsicherheiten und Messdatenauswertung, Wiley-VCH, Weinheim.

Weise K. et al. (2006) Bayesian decision threshold, detection limit and confidence limits in ionizing radiation measurements, Radiation Protection Dosimetry 121(1), 52–63.

Weise K. et al. (2009) Monte Carlo determination of the characteristic limits in measurement of ionising radiation: Fundamentals and numerics. Radiation Protection Dosimetry 135(3), 169–196; doi:10.1093/rpd/ncp105.

Weise K. et al. (2004) Nachweisgrenzen und Erkennungsgrenzen bei Kernstrahlungsmessungen: Spezielle Anwendungen, ISSN 1013-4506, ISBN 3-8249-0904-9, FS-04-127-AKSIGMA, TÜV Verlag Rheinland (2004) https://www.fs-ev.org/fileadmin/user_upload/04_Arbeitsgruppen/06_Nachweisgrenzen/02_Dokumente/Arbeitsmaterialien/Weise_et_al_NWG_Vorschlag_fuer_eine_Norm_FS-04-127-AKSIGMA.pdf

Weise K. et al. (2005) Determination of the detection limit and decision threshold for ionizing-radiation measurements: fundamentals and particular applications, ISSN 1013-4506, ISBN 3-8249-0945-6, FS-05-129-AKSIGMA, TÜV Verlag Rheinland, Köln (2005) https://www.fs-ev.org/fileadmin/user_upload/04_Arbeitsgruppen/06_Nachweisgrenzen/02_Dokumente/Arbeitsmaterialien/Weise_et_al_Characteristic_limits_-_proposal_for_a_standard_FS-05-129-AKSIGMA.pdf

Weise K. et al. (2013) Characteristics values in measurement of ionizing radiation – material for a critical discussion on fundamentals and alternatives, FS-2013-167-AKSIGMA, http://www.fs-ev.org/fileadmin/user_upload/04_Arbeitsgruppen/06_Nachweisgrenzen/02_Dokumente/

Arbeitsmaterialien/Weise_et_al_Characteristic_Values_-_material_for_a_critical_discussion_FS-2013-167-AKSIGMA.pdf, Verlag TÜV Rheinland.

Yang R., Berger J.O. (1996) A catalogue of noninformative priors, http://www.stats.org.uk/priors/noninformative/YangBerger1998.pdf .

15 Appendix: PDFs and their application depending on the available information.

PDF $f(x)$	Formulas	Comments
Normal or Gaussian $\mathrm{N}(x\|\mu,\sigma^2)$	PDF $\mathrm{N}(x\|\mu,\sigma^2)=f_X(x\|\mu,\sigma)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{(x-\mu)^2}{2\cdot\sigma^2}\right)\quad x,\mu\in\Re;\sigma^2>0$ Expectation $\mathrm{E}\left(\mathrm{N}(x\|\mu,\sigma^2)\right)=\mu$ Variance $\mathrm{Var}\left(\mathrm{N}(x\|\mu,\sigma^2)\right)=\sigma^2$ Standard normal distribution PDF $f_{\mathrm{N}}(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ Distribution function $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}\,\mathrm{d}t$ Error function $\mathrm{erf}(x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{x}e^{-t^2}\,\mathrm{d}t$ Distribution function for a normal distribution with parameters μ and σ: $F_{\mathrm{N}}(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{x-\mu}{\sigma\cdot\sqrt{2}}\right)\right]$	The limiting theorem of statistics says: any sum of arbitrarily distributed random quantities goes for a normal distribution. Thus, the normal distribution is the limiting PDF for a sum of arbitrarily distributed quantities. In metrology $f_Y(\tilde{y}\|y,u(y))=\frac{1}{\sqrt{2\pi}\cdot u(\mathrm{y})}\exp\left(-\frac{(\tilde{y}-y)^2}{2u^2(y)}\right)$ is according to the PME the exact solution if only y and $u(y)$ are known.
Exponential $\mathrm{Exp}(x\|\lambda)$	PDF $\mathrm{Exp}(x\|\lambda)=f_x(x\|\lambda)=\lambda\cdot e^{-\lambda\cdot x}\quad x\geq 0,\text{else } 0\,;\lambda>0$ Distribution function $F_{\mathrm{Exp}}(x\|\lambda)=1-e^{-\lambda\cdot x}$ Expectation $\mathrm{E}(\mathrm{Exp}(x\|\lambda))=1/\lambda$ Variance $\mathrm{Var}(\mathrm{Exp}(x\|\lambda))=1/\lambda^2$	The exponential distribution is the probability distribution of the time between events in a Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate as when counting a long-lived radionuclide. The exponential distribution is the maximum entropy distribution of a random variable $X\geq 0$ with a given and fixed expectation $\mathrm{E}(X)$.

PDF $f(x)$	Formulas	Comments
Binomial $\mathrm{Bi}(k\|p,n)$ discrete	PDF $\mathrm{Bi}(k\|p,n) = f_K(k\|p,n) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \quad k \in \{0,1,2,...\}, \text{else} = 0;\ p > 0$ Distribution function $F_{\mathrm{Bi}}(k\|p,n) = \sum_{i=1}^{\lfloor k \rfloor} \binom{n}{i} p^i \cdot (1-p)^{n-i}$ $\lfloor k \rfloor$ denotes the rounding function Expectation Variance $\mathrm{Var}(\mathrm{Bi}(k\|p,n)) = n \cdot p \cdot (1-p)$	The binomial distribution describes the number of successes having a probability p in n independent Bernoulli experiments which only have the outcomes success or failure. For the radioactive decay this means counting a radionuclide with short half-life for a counting time t. N is the number of atoms at the beginning of the experiment. $\mathrm{Bi}(n\|\lambda, N, t_\mathrm{m}) = \binom{N}{n} (\lambda \cdot t_\mathrm{m})^n (1-\lambda \cdot t_\mathrm{m})^{N-n}$
Poisson $\mathrm{Po}(n\|\lambda)$ discrete	PDF $\mathrm{Po}(n\|\lambda) = f_N(n\|\lambda) = \lambda^n \cdot e^{-\lambda} / n! \quad \lambda \in \Re > 0; n = 0,1,2,...$ Distribution function $F_{\mathrm{Po}}(n\|\lambda) = p$ Expectation = Variance $\mathrm{E}(\mathrm{Po}(n\|\lambda)) = \mathrm{Var}(\mathrm{Po}(n\|\lambda)) = \lambda$	When counting a radionuclide with a long half-life for a counting time t the number of counts obey to a Poisson distribution with the expectation $\mathrm{E}(n) = \lambda \cdot N \cdot t_\mathrm{m}$ and the variance $\mathrm{Var}(n) = \lambda \cdot t_\mathrm{m} \cdot (1 - \lambda \cdot t_\mathrm{m}) \cdot N$ The parameter λ describes the expected number of events observed. p is the probability to find at most n events, if λ is expected on the average.
Negative binomial $\mathrm{NB}(n\|r,p)$ discrete	PDF $\mathrm{NB}(n\|r,p) = f_N(n\|r,p) = \binom{n-1}{r-1} \cdot p^r \cdot (1-p)^{n-r}; n = 0,1,2,3...$ $r > 0$ number of successful trials; $p \in [0,1]$ probability for success in a single trial Expectation $\mathrm{E}(\mathrm{NB}(n\|r,p)) = \frac{r \cdot (1-p)}{p}$	Counting a radionuclide with random influences which are characterized by ϑ. $1/r = \vartheta^2$ In such cases, the observations are overdispersed with respect to a Poisson distribution, for which the mean is equal to the variance. Hence a Poisson distribution is not an appropriate model. This results in a continuous mixture of Poisson distributions where the parameter λ is variable. This

PDF $f(x)$	Formulas	Comments
	Variance $\mathrm{Var}(\mathrm{NB}(n\|r,p)) = \frac{r\cdot(1-p)}{p^2} = \mathrm{E}(\mathrm{NB}(n\|r,p)) + \frac{\mathrm{E}(\mathrm{NB}(n\|r,p))^2}{r}$ coefficient of variation $\frac{\sigma}{\mu} = \frac{\sqrt{\mathrm{Var}(\mathrm{NB}(n\|r,p))}}{\mathrm{E}(\mathrm{NB}(n\|r,p))} = \sqrt{\frac{1}{\mathrm{E}(\mathrm{NB}(n\|r,p))} + \frac{1}{r}} = \sqrt{\frac{1}{\mathrm{E}(\mathrm{NB}(n\|r,p))} + \vartheta^2}$	results in a mixing distribution of the Poisson rate with gamma distributed parameter λ. For counting measurements with random influences ϑ^2 is the contribution of the random influences to the relative uncertainty.
Gamma $\mathrm{Ga}(x\|\alpha,\beta)$	PDF for a parametrisation with shape parameter α and a scale parameter β $\mathrm{Ga}(x\|\alpha,\beta) = f_X(x\|\alpha,\beta) = \frac{\beta^\alpha \cdot x^{\alpha-1}\cdot e^{-\beta\cdot x}}{\Gamma(\alpha)};\quad x>0;\alpha,\beta>0$ Expectation $\mathrm{E}\left(\mathrm{Ga}(x\|\alpha,\beta)\right) = \alpha/\beta$ Variance $\mathrm{Var}\left(\mathrm{Ga}(x\|\alpha,\beta)\right) = \alpha/\beta^2$	Rate with *n* counted events during the time *t* with a non-informative prior $f_R(\tilde{r}) = C/\tilde{r}$ for a stationary Poisson process $\mathrm{Ga}(\tilde{r}\|n,1/t) = t\cdot(\tilde{r}\cdot t)^{n-1}\cdot e^{-\tilde{r}\cdot t}/(n-1)!;\quad (\tilde{r}\geq 0)$ with a shape parameter $\alpha = n$ and a scale parameter $\beta = t$. $\mathrm{E}(\mathrm{Ga}(\tilde{r}\|n,t)) = r = n/t$ $\mathrm{Var}(\mathrm{Ga}(\tilde{r}\|n,t)) = r/t = n/t^2$
Beta $\mathrm{B}(x\|\alpha,\beta,x_\mathrm{L},x_\mathrm{U})$	PDF $\mathrm{B}(x\|\alpha,\beta,x_\mathrm{L},x_\mathrm{U}) = f_X\left(x\|x_i,i=1,\ldots,n\right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\cdot x^{\alpha-1}\cdot(1-x)^{\beta-1}$ $\alpha,\beta\in\Re>0\quad x\in[0,1]$ Expectation $\mathrm{E}(\mathrm{B}(x\|\alpha,\beta,x_\mathrm{L},x_\mathrm{U})) = \frac{\alpha}{\alpha+\beta}$ Variance $\mathrm{Var}(\mathrm{B}(x\|\alpha,\beta,x_\mathrm{L},x_\mathrm{U})) = \frac{\alpha\cdot\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ The parameters of the Beta distribution are calculated as	Series of indications with lower and upper limits $x_\mathrm{L}, x_\mathrm{U}$ with the parameters $\alpha = \overline{x}^2\cdot\left(\frac{1-\overline{x}}{s_x^2} - \frac{1}{\overline{x}}\right)$ and $\beta = \alpha\cdot\left(\frac{1}{\overline{x}} - 1\right)$ Applied in ISO 11929-4:2020 to multiple measurements of efficiencies. $\overline{\varepsilon} = \frac{1}{n-1}\sum_{i=1}^{n}\varepsilon_i$ and $s_\varepsilon^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - \overline{\varepsilon})^2$

PDF $f(x)$	Formulas	Comments
	$\alpha = \left[\mathrm{E}(\mathrm{B}(\alpha,\beta,x_\mathrm{L},x_\mathrm{U}))\right]^2 \times$ $\times\left(\frac{1-\mathrm{E}(\mathrm{B}(x\vert\alpha,\beta,x_\mathrm{L},x_\mathrm{U}))}{\mathrm{Var}(\mathrm{B}(x\vert\alpha,\beta,x_\mathrm{L},x_\mathrm{U}))}-\frac{1}{\mathrm{E}(\mathrm{B}(x\vert\alpha,\beta,x_\mathrm{L},x_\mathrm{U}))}\right)$ $\beta = \alpha\cdot\left(\frac{1}{\mathrm{E}(\mathrm{B}(x\vert\alpha,\beta,x_\mathrm{L},x_\mathrm{U}))}-1\right)$	$\alpha = \bar{\varepsilon}^2\left(\frac{1-\bar{\varepsilon}}{s_\varepsilon^2}-\frac{1}{\bar{\varepsilon}}\right)$ and $\beta = \alpha\cdot\left(\frac{1}{\bar{\varepsilon}}-1\right)$ $f\left(\tilde{\varepsilon}\vert(\varepsilon_i, i=1,...,n), \varepsilon_\mathrm{L}\le\tilde{\varepsilon}\le\varepsilon_\mathrm{U}\right) =$ $\mathrm{B}(\tilde{\varepsilon}\vert\alpha,\beta,\varepsilon_\mathrm{L},\varepsilon_\mathrm{U}) = \frac{1}{\varepsilon_\mathrm{U}-\varepsilon_\mathrm{L}}\cdot\mathrm{B}\left(\tilde{\varepsilon}\vert\frac{\tilde{\varepsilon}-\varepsilon_\mathrm{L}}{\varepsilon_\mathrm{U}-\varepsilon_\mathrm{L}},\alpha,\beta\right)$
Scaled and shifted *t*-distribution $t_{n-1}(x\vert\bar{x},s^2(n)/n)$	PDF $t_{n-1}(x\vert\bar{x},s^2(n)/n) = \frac{\Gamma(n/2)}{\sqrt{(n-1)\cdot\pi}\cdot\Gamma((n-1)/2)}\cdot\frac{1}{s/\sqrt{n}}\times\left(1+\frac{1}{n-1}\cdot\left(\frac{x-\bar{x}}{s/\sqrt{n}}\right)^2\right)^{-n/2}$ $\bar{x}\in\Re;\ s^2(n)>0$ Expectation $\mathrm{E}(t_{n-1}(x\vert\bar{x},s^2(n)/n)) = \bar{x}$ Variance $\mathrm{Var}(t_{n-1}(x\vert\bar{x},s^2(n)/n)) = \frac{n-1}{n-3}\cdot\frac{s^2}{n}$ $\bar{x} = \frac{1}{n}\cdot\sum_{i=1}^{n}x_i$ and $s^2 = \frac{1}{n-1}\cdot\sum_{i=1}^{n}(x_i-\bar{x})^2$	The scaled and shifted *t*-distribution $t_{n-1}(x\vert\bar{x},s^2(n)/n)$ with $\nu = n-1$ degrees of freedom and $t = \frac{x-\bar{x}}{s/\sqrt{n}}$ finds its application for repeated measurements. $f_\mathrm{t}(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\cdot\pi}\cdot\Gamma(\nu/2)}\cdot\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$ It is applied to series of indications $x_1,...,x_n$ sampled independently from a quantity having a Gaussian distribution, with unknown expectation and unknown variance.
Rectangular $\mathrm{R}(x\vert x_1,x_\mathrm{u})$	PDF $\mathrm{R}(x\vert x_1,x_\mathrm{u}) = f_X(x\vert x_1,x_\mathrm{u}) = \begin{cases}1/(x_\mathrm{u}-x_1), & x_1\le x_\mathrm{u}\\ 0 & \text{otherwise}\end{cases}$ Expectation $\mathrm{E}\left(\mathrm{R}(x\vert x_1,x_\mathrm{u})\right) = (x_\mathrm{u}+x_1)/2$ Variance $\mathrm{Var}\left(\mathrm{R}(x\vert x_1,x_\mathrm{u})\right) = (x_\mathrm{u}-x_1)^2/12$	The rectangular distribution applies to the case of a random variable for which only the lower and upper limits are known. The quantity is supposed to behave randomly between these limits. The GUM S1 gives also formulas if the limits are inexactly descried.

PDF $f(x)$	Formulas	Comments
Trapezoidal $\mathrm{Trap}(x\|a,b,\beta)$	See explicit formulas and figure in chapter 6.4.4 of GUM S1 $\mathrm{Trap}(x\|a,b,\beta)$ with $a=a_1$, $b=b_1+b_2$ and $\beta=\|(b_1-a_1)-(b_2-a_2)\|/(b-a)$	Sum of two quantities assigned rectangular distributions with lower and upper limits a_1, b_1 and a_2, b_2, respectively.
Triangular $\mathrm{Tri}(x\|x_{1,\mathrm{l}},x_{1,\mathrm{u}},x_{2,\mathrm{l}},x_{2,\mathrm{u}})$	See explicit formulas in chapter 6.4.5 of GUM S1 $\mathrm{Tri}(x\|x_{1,\mathrm{l}},x_{1,\mathrm{u}},x_{2,\mathrm{l}},x_{2,\mathrm{u}})$	Sum of two quantities assigned rectangular distributions with lower and upper limits $x_{1,\mathrm{l}}$ $x_{2,\mathrm{l}}$ and $x_{1,\mathrm{u}}$ $x_{2,\mathrm{u}}$ and the same semi-width $(x_{1,\mathrm{u}}-x_{1,\mathrm{l}}=x_{2,\mathrm{u}}-x_{2,\mathrm{l}})$.
Log-normal $\mathrm{LN}(x\|\overline{\ln(x_i)},\mathrm{s}^2(\ln(x_i)))$	PDF $\mathrm{LN}(x\|\overline{\ln(x_i)},\mathrm{s}^2(\ln(x_i)))=\mathrm{LN}_X(x\|\mu,\sigma)=\frac{1}{x}\frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{(\ln(x)-\mu)^2}{2\cdot\sigma^2}\right)$ Distribution function $F_{\mathrm{LN}}(x\|\mu,\sigma)=\mathbf{\Phi}\left(\frac{(\ln(x)-\mu)}{\sigma}\right)$ Expectation $\mathrm{E}\left(\mathrm{LN}(x\|\mu,\sigma)\right)=\exp\left(\mu+\sigma^2/2\right)$ Median $\mathrm{M}\left(\mathrm{LN}(x\|\mu,\sigma)\right)=\exp(\mu)$ Variance $\mathrm{Var}\left(\mathrm{LN}(x\|\mu,\sigma)\right)=\left[\exp(\sigma^2)-1\right]\cdot\exp\left(2\mu+\sigma^2\right)$	The multiplicative limiting theorem of statistics: Any product of arbitrarily distributed random quantities goes for a logarithmic normal distribution. The log-normal distribution with μ and σ being the logarithms of the geometric mean and of the geometric standard deviation, respectively. It is the limiting PDF for a product of arbitrarily distributed quantities. $\overline{\ln(x_i)}=\mu=\frac{1}{n}\cdot\sum_{i=1}^{n}\ln(x_i)$ and $s^2(\ln(x_i))=\sigma^2=\frac{1}{n-1}\cdot\sum_{i=1}^{n}(\ln(x_i)-\overline{\ln(x_i)})^2$

16 Glossary

The terms and definitions are according to the VIM 3rd ed. (JCGM 2008).

quantity

property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference (VIM 3rd ed. 2008, 1.1)

measurement unit, unit of measurement, unit

real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number (VIM 3rd ed. 2008, 1.9)

International System of Units, SI

system of units, based on the International System of Quantities, their names and symbols, including a series of prefixes and their names and symbols, together with rules for their use, adopted by the General Conference on Weights and Measures (CGPM) (VIM 3rd ed. 2008, 1.16)

quantity value, value of a quantity, value

number and reference together expressing magnitude of a quantity (VIM 3rd ed. 2008, 1.19)

measurement

process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity (VIM 3rd ed. 2008, 2.1)

metrology

science of measurement and its application (VIM 3rd ed. 2008, 2.2)

measurand

quantity intended to be measured (VIM 3rd ed. 2008, 2.3)

measurement method, method of measurement

generic description of a logical organization of operations used in a measurement (VIM 3rd ed. 2008, 2.5)

measurement result, result of measurement

set of quantity values being attributed to a measurand together with any other available relevant information (VIM 3rd ed. 2008, 2.9)

measured quantity value, value of a measured quantity, measured value

quantity value representing a measurement result (VIM 3rd ed. 2008, 2.10)

true quantity value, true value of a quantity, true value

quantity value consistent with the definition of a quantity (VIM 3rd ed. 2008, 2.11)

measurement uncertainty, uncertainty of measurement, uncertainty

non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used (VIM 3rd ed. 2008, 2.26)

Type A evaluation of measurement uncertainty, Type A evaluation

evaluation of a component of measurement uncertainty by a statistical analysis of measured quantity values obtained under defined measurement conditions (VIM 3rd ed. 2008, 2.28)

Type B evaluation of measurement uncertainty, Type B evaluation

evaluation of a component of measurement uncertainty determined by means other than a Type A evaluation of measurement uncertainty (VIM 3rd ed. 2008, 2.29)

Examples: Evaluations based on information

- associated with authoritative published quantity values,
- associated with the quantity value of a certified reference material,
- obtained from a calibration certificate,
- about drift,
- obtained from the accuracy class of a verified measuring instrument,
- obtained from limits deduced through personal experience.

standard measurement uncertainty, standard uncertainty of measurement, standard uncertainty

measurement uncertainty expressed as a standard deviation (VIM 3rd ed. 2008, 2.30)

combined standard measurement uncertainty, combined standard uncertainty

standard measurement uncertainty that is obtained using the individual standard measurement uncertainties associated with the input quantities in a measurement model (VIM 3rd ed. 2008, 2.31)

relative standard measurement uncertainty

standard measurement uncertainty divided by the absolute value of the measured quantity value (VIM 3rd ed. 2008, 2.32)

uncertainty budget

statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination (VIM 3rd ed. 2008, 2.33)

expanded measurement uncertainty, expanded uncertainty

product of a combined standard measurement uncertainty and a factor larger than the number one (VIM 3rd ed. 2008, 2.35)

coverage interval

interval containing the set of true quantity values of a measurand with a stated probability, based on the information available (VIM 3rd ed. 2008, 2.36)

coverage probability

probability that the set of true quantity values of a measurand is contained within a specified coverage interval (VIM 3rd ed. 2008, 2.37)

coverage factor

number larger than one by which a combined standard measurement uncertainty is multiplied to obtain an expanded measurement uncertainty (VIM 3rd ed. 2008, 2.38)

metrological traceability

property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty (VIM 3rd ed. 2008, 2.41)

measurement model, model of measurement, model

mathematical relation among all quantities known to be involved in a measurement (VIM 3rd ed. 2008, 2.48)

input quantity in a measurement model, input quantity

quantity that must be measured, or a quantity, the value of which can be otherwise obtained, in order to calculate a measured quantity value of a measurand (VIM 3rd ed. 2008, 2.50)

output quantity in a measurement model, output quantity

quantity, the measured value of which is calculated using the values of input quantities in a measurement model (VIM 3rd ed. 2008, 2.51)

influence quantity

quantity that, in a direct measurement, does not affect the quantity that is actually measured, but affects the relation between the indication and the measurement result (VIM 3rd ed. 2008, 2.52)

correction

compensation for an estimated systematic effect (VIM 3rd ed. 2008, 2.53)

indication

quantity value provided by a measuring instrument or a measuring system (VIM 3rd ed. 2008, 4.1)

blank indication, background indication

indication obtained from a phenomenon, body, or substance similar to the one under investigation, but for which a quantity of interest is supposed not to be present, or is not contributing to the indication (VIM 3rd ed. 2008, 4.2

detection limit, limit of detection

measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence (VIM 3rd ed. 2008, 4.18)

instrumental measurement uncertainty

component of measurement uncertainty arising from a measuring instrument or measuring system in use (VIM 3rd ed. 2008, 4.24)

maximum permissible measurement error, maximum permissible error, limit of error

extreme value of measurement error, with respect to a known reference quantity value, permitted by specifications or regulations for a given measurement, measuring instrument, or measuring system (VIM 3rd ed. 2008, 4.26)

probability distribution, distribution

probability measure induced by a random variable (JCGM 106 2012: 3.1.1)

distribution function

function giving, for every value ξ, the probability that the random variable X be less than or equal to ξ: (JCGM 106 2012: 3.1.2)

probability density function, PDF

derivative, when it exists, of the distribution function (JCGM 106 2012: 3.1.3)

conformity assessment

activity to determine whether specified requirements relating to a product, process, system, person or body are fulfilled (JCGM 106 2012: 3.3.1)

specified requirement

need or expectation that is stated (JCGM 106 2012: 3.3.3)

tolerance limit, specification limit

specified upper or lower bound of permissible values of a property (JCGM 106 2012: 3.3.4)

tolerance interval

interval of permissible values of a property (JCGM 106 2012: 3.3.5)

conformance probability

probability that an item fulfils a specified requirement (JCGM 106 2012: 3.3.7)

acceptance limit

specified upper or lower bound of permissible measured quantity values (JCGM 106 2012: 3.3.8)

acceptance interval

interval of permissible measured quantity values (JCGM 106 2012: 3.3.9)

rejection interval

interval of non-permissible measured quantity values (JCGM 106 2012: 3.3.10)

decision rule

documented rule that describes how measurement uncertainty will be accounted for with regard to accepting or rejecting an item, given a specified requirement and the result of a measurement (JCGM 106 2012: 3.3.12)

estimate

a guess of what the size, value, amount, cost, etc. of something might be

estimator

a function of the data that is used to estimate the value of an unknown parameter on the basis of a statistical model.