



# Choice of positive distribution law for nuclear data

Sébastien Lahaye

► **To cite this version:**

Sébastien Lahaye. Choice of positive distribution law for nuclear data. EPJ N - Nuclear Sciences & Technologies, EDP Sciences, 2018, 4, pp.38. 10.1051/epjn/2018047 . cea-02304354

**HAL Id: cea-02304354**

**<https://hal-cea.archives-ouvertes.fr/cea-02304354>**

Submitted on 3 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Choice of positive distribution law for nuclear data

Sébastien Lahaye\*

DEN – Service d'études des réacteurs et de mathématiques appliquées (SERMA), CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

Received: 6 November 2017 / Received in final form: 9 March 2018 / Accepted: 9 July 2018

**Abstract.** Nuclear data evaluation files in the ENDF6 format provide mean values and associated uncertainties for physical quantities relevant in nuclear physics. Uncertainties are denoted as  $\Delta$  in the format description, and are commonly understood as standard deviations. Uncertainties can be completed by covariance matrices. The evaluations do not provide any indication on the probability density function to be used when sampling. Three constraints must be observed: the mean value, the standard deviation and the positivity of the physical quantity. MENDEL code generally uses positively truncated Gaussian distribution laws for small relative standard deviations and a lognormal law for larger uncertainty levels ( $>50\%$ ). Indeed, the use of truncated Gaussian laws can modify the mean and standard deviation value. In this paper, we will make explicit the error in the mean value and the standard deviation when using different types of distribution laws. We also employ the principle of maximum entropy as a criterion to choose among the truncated Gaussian, the fitted Gaussian and the lognormal distribution. Remarkably, the difference in terms of entropy between the candidate distribution laws is a function of the relative standard deviation only. The obtained results provide therefore general guidance for the choice among these distributions.

## 1 Introduction

Nuclear data evaluation files in the ENDF6 format [1] provide mean values and associated uncertainties for physical quantities relevant in nuclear physics. These uncertainties are denoted as  $\Delta$  in the format description for most of the nuclear data parameter types, and are understood as standard deviations. Uncertainties can be completed (for microscopic cross sections, for example) by covariance matrices.

For uncertainty propagation based on random sampling, one needs to know the probability density function associated with each random variable. Current nuclear data evaluation files [2–4] do not provide any indication which probability density function to use, and users have to choose a distribution law. This is the case for all uncertain data propagated in fuel cycle code systems, such as independent fission yields, radioactive decay constants, radioactive decay energies, radioactive decay branching ratios and multigroup microscopic cross sections. For some data, in particular independent neutron fission yields and some microscopic cross sections, the relative standard deviation can be high (more than 50%), and Gaussian

laws naturally lead to negative occurrences, which is not acceptable for those physical quantities.

Three constraints must therefore be respected when sampling:

- positivity of the physical quantity;
- its mean value;
- and its standard deviation.

Often the positively truncated Gaussian law is used, which takes into account positivity but introduces a bias in the mean value and the standard deviation. Furthermore, it is not symmetric around the mean value.

MENDEL [5], is the new generation of the CEA code system for nuclear fuel cycle studies. Its depletion solver is provided to the transport code systems APOLLO3<sup>®</sup> [6] and TRIPOLI-4<sup>®</sup> [7]. MENDEL is the successor of DARWIN/PEPIN2 [8].

Uncertainty quantification in MENDEL is based on a propagation method using a Monte Carlo approach by correlated sampling [9], and sampling is done by the CEA uncertainty platform URANIE [10,11]. Nuclear data uncertainties are propagated to physical quantities of interest (such as decay heat or concentrations) [12].

Until now, the choice of distribution law in MENDEL is done in the following way:

- positively truncated Gaussian law if the relative standard deviation is less than 50%;

\* e-mail: [sebastien.lahaye@cea.fr](mailto:sebastien.lahaye@cea.fr)

– lognormal law if the relative standard deviation is more than 50%.

This choice is based on physical reasons, as truncated Gaussian distributions modify the mean value and the standard deviation values for large relative uncertainties. The switching point between the Gaussian and the lognormal distributions is a pragmatic choice. This paper aims to give a formal justification for this choice.

The structure of this paper is as follows. First, we will investigate the introduced bias in both the mean value and the standard deviation by a truncation of the Gaussian distribution. We will then describe how to modify the Gaussian law parameters in order to obtain after truncation the mean value and the standard deviation as specified in evaluation files.

In the second part, we employ the principle of maximum entropy [13,14] to choose between different distribution laws. We will show that the choice of the distribution law depends on the relative standard deviation.

We limit our study in this paper to the distribution laws themselves, without propagation in numerical code systems.

## 2 Maximum entropy method

For a continuous probability density function  $p(x)$  defined on  $I$ , we introduce the differential entropy defined as:

$$h(p) = -\int_I p(x) \ln p(x) dx \quad (1)$$

We define  $p(x) \ln p(x) = 0$  when  $p(x) = 0$  (due to  $\lim_{p \rightarrow 0^+} p \ln p = 0$ ).

This entropy function appears in statistical physics and thermodynamics where higher entropy is associated with states closer to equilibrium.

The maximum entropy principle [14] states that for a given set of constraints (for example known mean value and standard deviation), probability distribution with the largest entropy should be chosen.

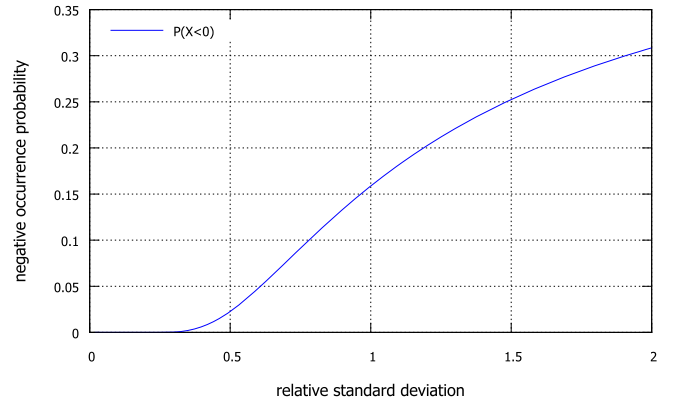
For given constraints, the law with the largest entropy is the one that contains the least amount of information about the physical quantity. For example, the maximum entropy principle will lead to the following choices:

- a uniform law if the constraints are minimal and maximal values;
- and a Gaussian law if the constraints are a given mean value and a standard deviation.

## 3 Candidate distribution laws

Candidate distribution laws must respect the three following criteria:

- **positivity of the realizations:**  $P(X < 0) = 0$  (*id est*  $I = R_+$ ), as nuclear data are positive;
- they must **match the mean value**  $\mu$  (within a given tolerance) as specified in the evaluation file;
- they must **match the standard deviation**  $\sigma$  (within a given tolerance) as specified in the evaluation file.



**Fig. 1.** Probability of negative occurrences for non truncated Gaussian distributions.

### 3.1 Gaussian distribution

Let  $G(\mu, \sigma)$  be the non-truncated Gaussian distribution with mean value  $\mu$  and standard deviation  $\sigma$ . Its probability distribution function reads:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \quad (2)$$

#### 3.1.1 Entropy

The Gaussian distribution maximizes the differential entropy among all distribution laws for given mean value and standard deviation. The Gaussian distribution entropy is given by:

$$h(g) = \ln\left(\frac{\sigma}{\mu}\sqrt{2\pi e}\right) + \ln\mu. \quad (3)$$

#### 3.1.2 Positivity

A Gaussian distribution yields negative values with the probability:

$$P(X < 0) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right). \quad (4)$$

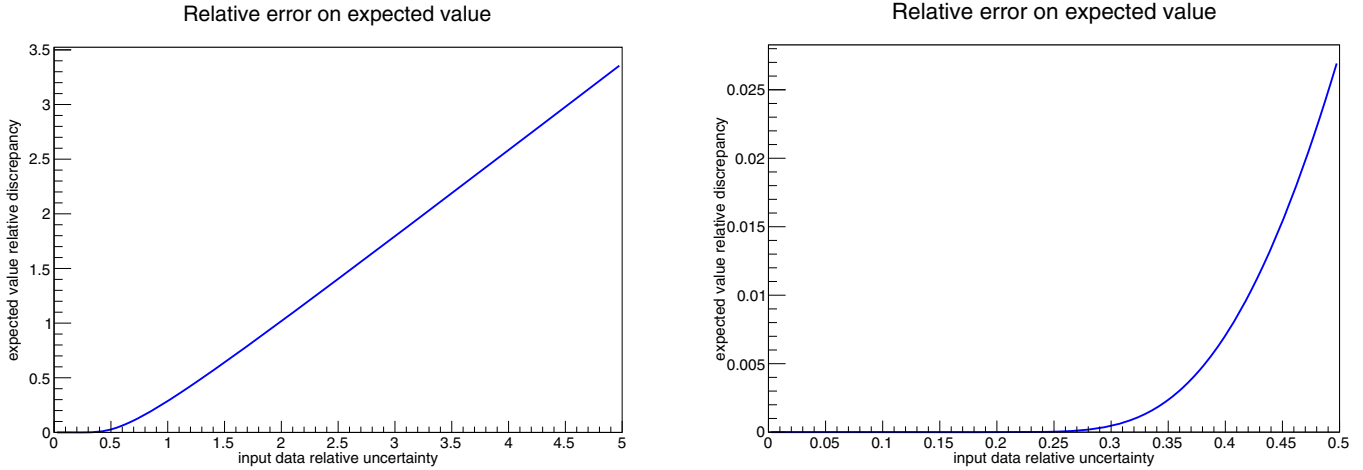
This negative occurrence probability is given as a function of the relative standard deviation in [Figure 1](#).

Due to this negative occurrence probability, which is non-negligible when  $\frac{\sigma}{\mu}$  is large, it is necessary to use other distribution laws to enforce the positivity constraint.

### 3.2 Positively truncated Gaussian distribution

#### 3.2.1 Distribution law

We define a Gaussian distribution with mean value  $\mu$  and standard deviation  $\sigma$ , and then set the probability density to zero for negative values. The resulting distribution – after normalization – is called a *positively truncated Gaussian distribution*.



**Fig. 2.** Relative error on mean value.

Draws can be realized by sampling from the original Gaussian distribution and rejecting all negative values. Its probability density function reads:

$$p(x) = \begin{cases} \frac{1}{\beta\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) & , x \geq 0 \\ 0 & , x < 0. \end{cases} \quad (5)$$

The constant  $\beta$  is defined so that  $\int_{\mathbb{R}} p(x) dx = 1$ , which means:

$$\beta = \frac{1}{2} \left( 1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \right).$$

If we substitute  $\delta = \frac{\sigma}{\mu}$ , we get the form:

$$\beta = \frac{1}{2} \left( 1 + \operatorname{erf}\left(\frac{1}{\delta\sqrt{2}}\right) \right). \quad (6)$$

The positively truncated Gaussian law will be denoted by  $\text{PG}(\mu, \sigma)$ .

### 3.2.2 Entropy

The differential entropy can be computed as a sum of  $\ln \mu$  and a function of  $\delta$ :

$$h(p) = \frac{1}{2} \left( 1 - \frac{1}{\delta\beta\sqrt{2\pi}} \exp\left(\frac{-1}{2\delta^2}\right) \right) + \ln(\delta\beta\sqrt{2\pi}) + \ln\mu. \quad (7)$$

### 3.2.3 Errors on moments

With this distribution, we modify the distribution moments, particularly for large relative uncertainties.

The truncated Gaussian distribution mean value is equal to:

$$E[p] = \mu + \sqrt{\frac{2}{\pi}} \frac{\sigma \exp\left(-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2\right)}{1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)}. \quad (8)$$

And the truncated Gaussian distribution variance is equal to:

$$\operatorname{Var}(p) = \sigma^2 + E[p](\mu - E[p]). \quad (9)$$

We obtain the following relative error on the expected value, which is a function of the relative standard deviation of the original input data (i.e. non-truncated Gaussian distribution  $\mu$  and  $\sigma$  parameters):

$$\frac{E[p] - \mu}{\mu} = \delta \sqrt{\frac{2}{\pi}} \frac{\exp\left(-\frac{1}{2}\left(\frac{1}{\delta}\right)^2\right)}{1 + \operatorname{erf}\left(\frac{1}{\delta\sqrt{2}}\right)}. \quad (10)$$

This bias is represented in [Figure 2](#) as a function of the parameter  $\delta$ .

The relative error of the standard deviation is also a function of  $\delta$  only:

$$\frac{\operatorname{Var}(p) - \sigma^2}{\sigma^2} = -\frac{\exp\left(-\frac{1}{2\delta^2}\right)}{\beta\sqrt{2\pi}} \left[ \frac{1}{\delta} + \frac{\exp\left(-\frac{1}{2\delta^2}\right)}{\beta\sqrt{2\pi}} \right]. \quad (11)$$

This bias is represented in [Figure 3](#) as a function of the parameter  $\delta$ .

The squared relative discrepancy between the relative standard deviations of the Gaussian distribution and the truncated Gaussian distribution is given in equation (12).

$$\frac{\frac{\operatorname{Var}(p)}{E[p]^2} - 1}{\delta^2} = \frac{\frac{1}{\delta^2} \left[ 1 + \frac{1}{\delta^2} + \frac{\exp\left(-\frac{1}{2\delta^2}\right)}{\delta\beta\sqrt{2\pi}} \right]}{\left( \frac{1}{\delta} + \frac{\exp\left(-\frac{1}{2\delta^2}\right)}{\beta\sqrt{2\pi}} \right)^2} - \frac{1 + \delta^2}{\delta^2}. \quad (12)$$

The relative standard deviation discrepancy is represented in [Figure 4](#). When choosing this truncated law, we obtain the numerical values for the errors shown in [Table 1](#).

We can conclude that if the truncation is totally acceptable up to 25% uncertainty, it begins to be problematic for a 50% uncertainty, and is totally unacceptable for 100% uncertainty.

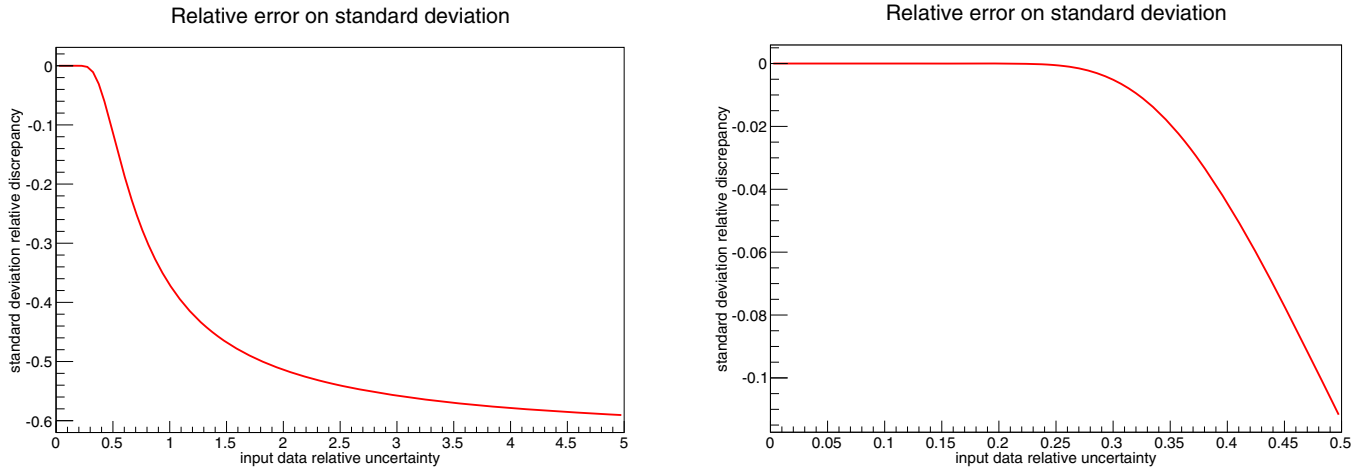


Fig. 3. Relative error on standard deviation value.

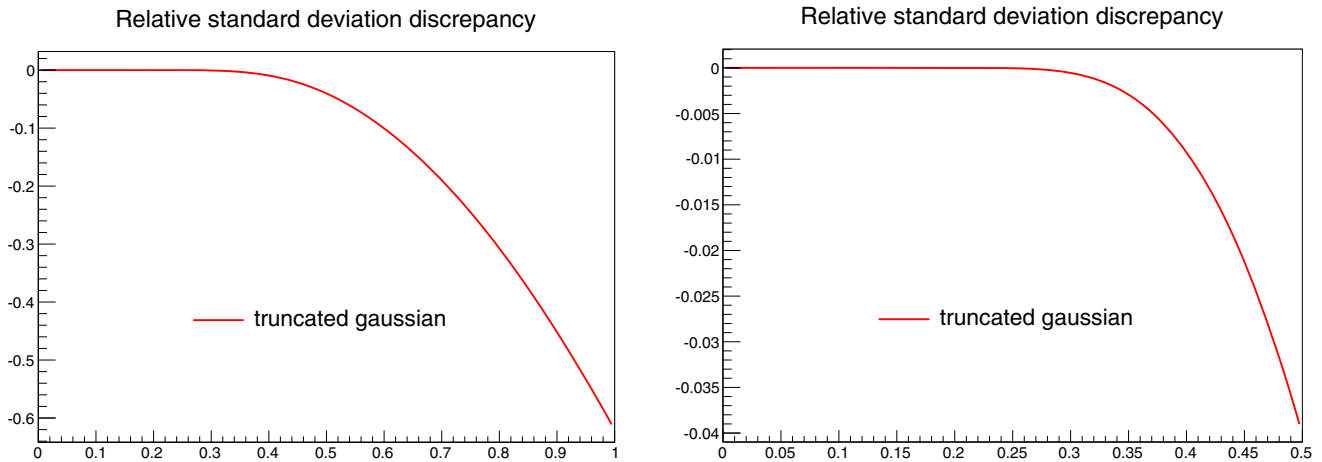


Fig. 4. Relative error on standard deviation value.

Table 1. Relative error between moments of the approximative law and expected moments.

Target rel. unc.	Rel. error on mean	Rel. error on std dev.	Abs. error on rel std dev.
10%	$<1 \cdot 10^{-10}$	$<1 \cdot 10^{-10}$	$<1 \cdot 10^{-10}$
25%	$3.4 \cdot 10^{-5}$	$5.4 \cdot 10^{-4}$	$3.8 \cdot 10^{-5}$
50%	2.7%	11%	4.0%
100%	28.8%	37.0%	62%

These formal results confirm the need to switch to another law, which has already been introduced heuristically in the URANIE/MENDEL sampling scheme.

### 3.3 Gaussian distribution with correct mean and variance

#### 3.3.1 Distribution law

We consider a random variable in an evaluated nuclear data file characterized by a mean value  $\mu$  and a standard deviation  $\sigma$ . We define a truncated Gaussian law  $q = \text{PG}(\tilde{\mu}, \tilde{\sigma})$  so that its mean value equals  $\mu$  and its

standard deviation equals  $\sigma$ .

$$q(x) = \begin{cases} \frac{1}{\tilde{\beta}\tilde{\sigma}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\tilde{\mu}}{\tilde{\sigma}}\right)^2\right), & x \geq 0 \\ 0, & x < 0. \end{cases} \quad (13)$$

Please note the changed notation compared to equation (5). The variables  $\tilde{\mu}$ ,  $\tilde{\sigma}$  here corresponds to  $\mu$ ,  $\sigma$  there.

The following specification of  $\tilde{\beta}$  ensures the proper normalization of the distribution:

$$\tilde{\beta} = \frac{1}{2} \left( 1 + \operatorname{erf}\left(\frac{1}{\tilde{\sigma}}\sqrt{2}\right) \right). \quad (14)$$

#### 3.3.2 Coefficient determination

We obtain from equation (8):

$$\mu = \tilde{\mu} + \tilde{\sigma} \sqrt{\frac{2}{\pi}} \frac{\exp\left[-\frac{1}{2}\left(\frac{\tilde{\mu}}{\tilde{\sigma}}\right)^2\right]}{1 + \operatorname{erf}\left(\frac{\tilde{\mu}}{\tilde{\sigma}\sqrt{2}}\right)}. \quad (15)$$

And from equation (9):

$$\sigma^2 = \tilde{\sigma}^2 + \mu(\tilde{\mu} - \mu).$$

Leading to:

$$\frac{\tilde{\mu} = \mu^2 + \sigma^2 - \tilde{\sigma}^2}{\mu} \quad (16)$$

which leads to (17) by substituting  $\tilde{\mu}$  as given in (16) in (15):

$$\sigma^2 = \tilde{\sigma}^2 - \mu \sqrt{\frac{2}{\pi}} \frac{\tilde{\sigma} \exp\left(-\frac{1}{2} \left(\frac{\mu^2 + \sigma^2 - \tilde{\sigma}^2}{\mu \tilde{\sigma}}\right)^2\right)}{1 + \operatorname{erf}\left(\frac{\mu^2 + \sigma^2 - \tilde{\sigma}^2}{\mu \tilde{\sigma} \sqrt{2}}\right)}. \quad (17)$$

Equation (17) contains one unknown variable  $\tilde{\sigma}$ , but its complexity does not enable a formal analytical solution.

Hence, we compute the solutions for several values of  $(\mu, \sigma)$  tuples numerically.

To do so, we introduce a pseudo relative standard deviation  $\frac{\tilde{\sigma}}{\mu}$  in equation (17).

With  $X = \delta = \frac{\sigma}{\mu}$  and  $\tilde{X} = \frac{\tilde{\sigma}}{\mu}$  we obtain:

$$X^2 = \tilde{X}^2 - \tilde{X} \sqrt{\frac{2}{\pi}} \frac{\exp\left[-\frac{1}{2} \left(\frac{1 + X^2 - \tilde{X}^2}{\tilde{X}}\right)^2\right]}{1 + \operatorname{erf}\left(\frac{1 + X^2 - \tilde{X}^2}{\tilde{X} \sqrt{2}}\right)}. \quad (18)$$

The distribution parameters are summarized in Table 2 for different values of  $X = \frac{\sigma}{\mu}$ .

Problematic values, i.e. negative and unreasonably large values, appear in bold letters.

The reader should note that  $\tilde{\mu}$  and  $\tilde{\sigma}$  are not the mean value and standard deviation of the truncated Gaussian distribution, which are respectively equal to  $\mu$  and  $\sigma$ . A negative value of  $\tilde{\mu}$  means that more than half of the distribution is truncated and the mode is positioned at zero, which may not be desirable.

For this reason, it is reasonable to not use the truncated Gaussian distribution for relative standard deviations larger than 75%.

### 3.3.3 Entropy

The differential entropy of the fitted Gaussian distribution is computed in the same way as in equation (7).

## 3.4 Lognormal distribution

### 3.4.1 Distribution law

The probability density function of a lognormal distribution characterized by parameters  $m$  and  $s$  reads:

$$l(x) = \frac{1}{xs\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln x - m}{s}\right)^2\right). \quad (19)$$

**Table 2.** Gaussian distribution parameters for correct mean and standard deviation.

Truncated Gaussian	Modified Gaussian law		
	$\frac{\tilde{\sigma}}{\mu}$	$\frac{\tilde{\mu}}{\mu}$	$\frac{\tilde{\sigma}}{\mu}$
0.00	0.00	1	0.00
0.05	0.05	1	0.05
0.1	0.1	1	0.1
0.15	0.15	1	0.15
0.2	0.20000	1	0.200001
0.25	0.25007	0.99997	0.25008
0.3	0.30080	0.99952	0.30095
0.35	0.35378	0.99734	0.35472
0.4	0.41117	0.99094	0.41493
0.45	0.47535	0.97655	0.48676
0.5	0.54897	0.94864	0.57869
0.55	0.63526	0.89895	0.70666
0.6	0.73845	0.81469	0.90641
0.65	0.86452	0.67510	1.28059
0.7	1.02249	0.44452	2.30020
0.75	1.22698	0.057029	<b>21.5150</b>
0.8	1.50413	<b>-0.62242</b>	<b>-2.41659</b>
0.85	1.90733	<b>-1.91541</b>	<b>-0.99578</b>
0.9	2.57152	<b>-4.80274</b>	<b>-0.53543</b>
0.95	4.01884	<b>-14.2486</b>	<b>-0.28205</b>
1	7.68560	<b>-57.0684</b>	<b>-0.13467</b>

### 3.4.2 Coefficient determination

The mean value of a lognormal distribution is equal to:

$$E[q] = \mu = \exp\left(m + \frac{s^2}{2}\right). \quad (20)$$

The variance of a lognormal distribution is equal to:

$$\operatorname{Var}(q) = \sigma^2 = \left(e^{s^2} - 1\right) e^{2m+s^2}. \quad (21)$$

which leads to:

$$s = \sqrt{\ln\left(1 + \left(\frac{\sigma}{\mu}\right)^2\right)}. \quad (22)$$

$$m = \ln \mu - \frac{1}{2} \ln\left(1 + \left(\frac{\sigma}{\mu}\right)^2\right). \quad (23)$$

**Table 3.** Entropy values for several laws.

$\sigma/\mu$	$h(\text{PG}(\mu, \sigma))$	$h(\text{PG}(\tilde{\mu}, \tilde{\sigma}))$	$h(\text{LN}(\mu, \sigma))$
0.05	-1.5768	-1.5768	-1.5787
0.10	-0.8836	-0.8836	-0.8911
0.15	-0.4782	-0.4782	-0.4949
0.20	-0.1905	-0.1905	-0.2200
0.25	0.0323	0.0326	-0.0129
0.30	0.2120	0.2145	0.1502
0.35	0.3573	0.3668	0.2822
0.40	0.4744	0.4956	0.3909
0.45	0.5692	0.6044	0.4814
0.50	0.6475	0.6955	0.5574
0.55	0.7140	0.7709	0.6214
0.60	0.7721	0.8323	0.6755
0.65	0.8240	0.8815	0.7213
0.70	0.8714	0.9198	0.7599
0.75	0.9153	0.9489	0.7924
0.80	0.9565	0.9701	0.8197
0.85	0.9955	0.9847	0.8424
0.90	1.0326	0.9938	0.8613
0.95	1.0682	0.9986	0.8767
1.00	1.1024	1.0001	0.8891

### 3.4.3 Entropy

The differential entropy of the lognormal distribution reads:

$$h(q) = \frac{1}{2} + \ln(s\sqrt{2\pi}) + m$$

$$h(q) = \frac{1}{2} \left[ 1 + \ln \left( 2\pi \ln \left( 1 + \left( \frac{\sigma}{\mu} \right)^2 \right) \right) - \ln \left( 1 + \left( \frac{\sigma}{\mu} \right)^2 \right) \right] + \ln \mu$$

$$h(q) = \frac{1}{2} \left[ 1 + \ln \left( \frac{2\pi \ln(1 + \delta^2)}{1 + \delta^2} \right) \right] + \ln \mu. \quad (24)$$

## 4 Choice of the law

### 4.1 Entropy principle

The different distribution laws will now be compared based on their differential entropies.

We show in Table 3 the entropy values for the truncated Gaussian law  $\text{PG}(\mu, \sigma)$ , the fitted truncated Gaussian law  $\text{PG}(\tilde{\mu}, \tilde{\sigma})$  – constant with the mean value and the standard deviation provided in the evaluated nuclear data file – and the log-normal law  $\text{LN}(\mu, \sigma)$ . The

values were obtained for  $\mu = 1$ . The differences between the entropies are independent from the choice of  $\mu$ .

Inspecting equations (7) and (24) shows that the difference between the truncated Gaussian law and the lognormal law is independent of  $\mu$ .

Even though the mode value  $\tilde{\mu}$  of the fitted Gaussian distribution which has to be used in equation (7) differs from  $\mu$ , the difference to the lognormal distribution or the truncated Gaussian distribution is still a function of the relative standard deviation  $\delta$  only. Equation (7) for the fitted Gaussian law reads:

$$h(p) = \frac{1}{2} \left( 1 - \frac{1}{\tilde{\delta}} \frac{1}{\tilde{\beta}} \sqrt{2\pi} \exp \left( \frac{-1}{2\tilde{\delta}^2} \right) \right) + \ln(\tilde{\delta}\tilde{\beta}\sqrt{2\pi}) + \ln\tilde{\mu}.$$

In fact,  $\tilde{\delta}$  is a function of  $\delta$  only as  $\tilde{\delta} = \frac{\tilde{\sigma}}{\tilde{\mu}} = \frac{\tilde{\sigma}}{\mu} \frac{\mu}{\tilde{\mu}}$  and:

–  $\frac{\tilde{\sigma}}{\mu}$  depends of  $\delta$  only as it is the solution of equation (18);  
– and  $\frac{\tilde{\mu}}{\mu} = 1 + \delta^2 + \left( \frac{\tilde{\sigma}}{\mu} \right)^2$  is also a function of  $\delta$  only.

In conclusion, the difference between the fitted truncated Gaussian law entropy and another candidate law entropy will be a function of  $\delta$  plus  $\ln\tilde{\mu} - \ln\mu = \ln\frac{\tilde{\mu}}{\mu}$ , which is also a function of  $\delta$  only.

Consequently, the difference in differential entropies is a function of the relative standard deviation only.

Despite entropy considerations, large discrepancies between the mode and the mean value are not desirable, which is an argument against the fitted truncated Gaussian in case of large relative uncertainties.

Using the entropy principle, we can see in Table 3 that – between those three laws – the modified Gaussian law is optimal, when the relative standard deviation is less than 80%.

When relative standard deviation is bigger than 80%, the truncated Gaussian distribution entropy is optimal. Nevertheless, for truncated Gaussian with high level of uncertainties, users need to take into account the huge discrepancy between objective and effective moments. Truncated Gaussian laws is not to be used in this context.

The comparison of the candidate laws for several values of  $\frac{\sigma}{\mu}$  in Figures 5 and 6 shows the similarity of the truncated Gaussian distribution and the Gaussian distribution in the case of small relative uncertainties. The fitted truncated Gaussian distribution is the first distribution to diverge from the Gaussian distribution, and resembles the lognormal distribution for large relative uncertainties (100% uncertainty, right part of Fig. 6).

### 4.2 Mode and mean

For a non-truncated Gaussian, mean value and mode are equal.

For the truncated Gaussian law (resp. the fitted truncated Gaussian law), the parameter  $\mu$  (resp.  $\tilde{\mu}$ ) indicates the distribution mode. In general, it is preferable to use distributions where the mode does not differ too much from the mean value.



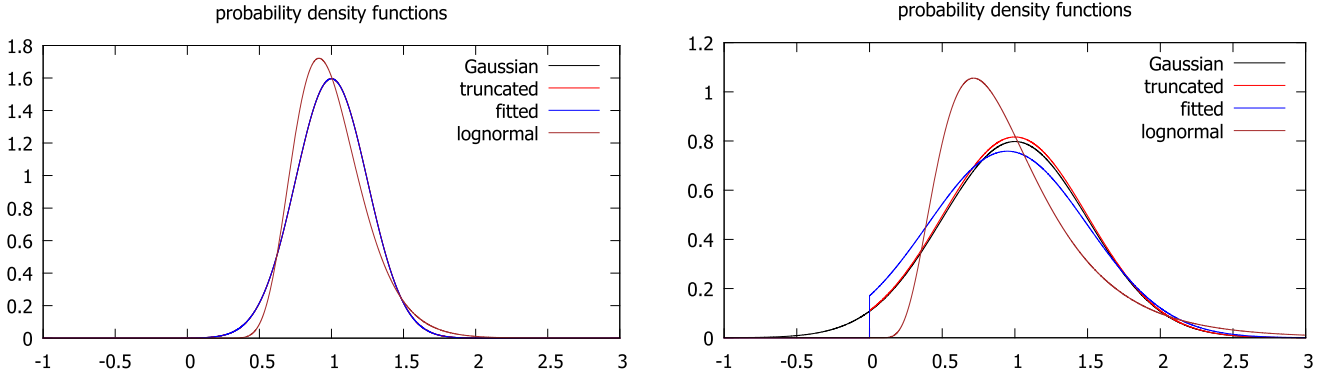


Fig. 5. Candidate law probability density function for relative standard deviation of 25% (left) and 50% (right).

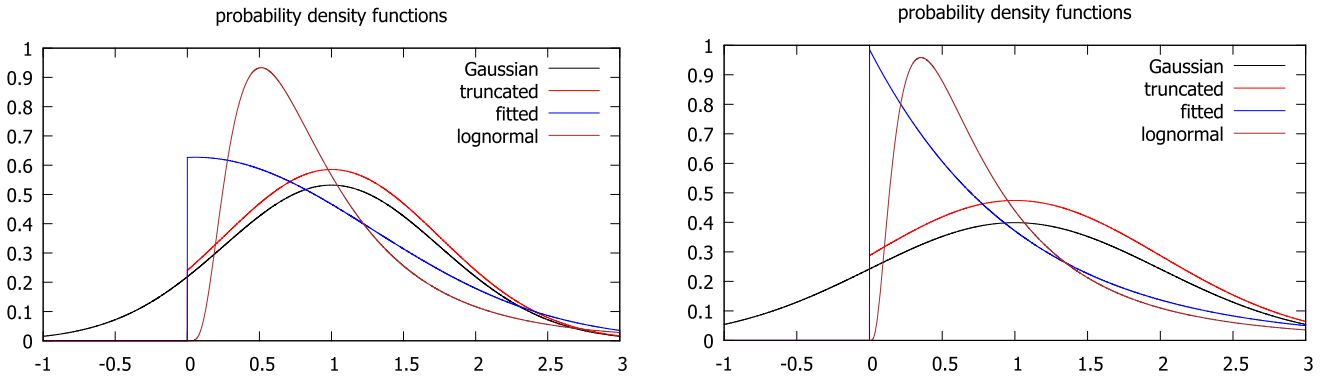


Fig. 6. Candidate law probability density function for relative standard deviation of 75% (left) and 100% (right).

If we want to limit to 10% the discrepancy between mode and mean value when using fitted truncated Gaussian distribution, we cannot employ it for  $\frac{\tilde{\mu}}{\mu} > 0.90$ . According to Table 2, this constraint is equivalent to:

$$\frac{\sigma}{\mu} \leq \frac{1}{2}.$$

### 5 Conclusion

The Gaussian distribution is not adequate for positive physical quantities, especially for large relative uncertainties, as it leads to excessive amount of negative occurrences.

For this reason, we investigated several positive distribution: the truncated Gaussian distribution, the fitted truncated Gaussian distribution and the lognormal distribution. All three laws exhibit zero probability density for negative values.

First, we studied the impact on the mean value and the standard deviation of the use of the truncated Gaussian distribution. We then compared the three distributions in terms of differential entropy. Despite the fact that the differential entropy is not scale invariant, the difference between two differential entropies is a function of the relative standard deviation only.

Both the maximum entropy principle and physical considerations have been considered in this work.

In summary, we suggest the following distribution recipe for choosing among the three distributions, depending on the relative standard deviation:

- for small values of relative uncertainties ( $\frac{\sigma}{\mu} < \frac{1}{4}$ ), the mean and standard deviation of the three laws are nearly identical. The differential entropy is slightly better for Gaussian laws. Hence, users can choose indifferently between the truncated Gaussian law  $PG(\mu, \sigma)$  and the fitted truncated Gaussian law  $PG(\tilde{\mu}, \tilde{\sigma})$ ;
- for intermediate values of relative uncertainties, id est ( $\frac{1}{4} \leq \frac{\sigma}{\mu} < \frac{1}{2}$ ), the principle of maximum entropy and favors the fitted truncated Gaussian law  $PG(\tilde{\mu}, \tilde{\sigma})$ ;
- for large values of relative uncertainties ( $\frac{\sigma}{\mu} \geq \frac{1}{2}$ ), positiveness of the mode and accuracy of the moments impose the choice of a lognormal law.

In conclusion we propose the following two laws:

$$p(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{\beta\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & , 0 \leq \frac{\sigma}{\mu} < \frac{1}{4} \quad x \geq 0 \\ \frac{1}{\beta\tilde{\sigma}\sqrt{2\pi}} \exp\left(-\frac{(x-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right) & , \frac{1}{4} \leq \frac{\sigma}{\mu} < \frac{1}{2}, \quad x \geq 0 \\ \frac{1}{xs\sqrt{2\pi}} \exp\left(-\frac{(\ln x - m)^2}{2s^2}\right) & , \frac{\sigma}{\mu} \geq \frac{1}{2}, \quad x \geq 0 \end{cases}$$



or

$$p(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{\beta\tilde{\sigma}\sqrt{2\pi}} \exp\left(-\frac{(x - \tilde{\mu})^2}{2\tilde{\sigma}^2}\right) & 0 \leq \frac{\sigma}{\mu} < \frac{1}{2}, x \geq 0 \\ \frac{1}{xs\sqrt{2\pi}} \exp\left(-\frac{(\ln x - m)^2}{2s^2}\right) & \frac{\sigma}{\mu} \geq \frac{1}{2}, x \geq 0 \end{cases}$$

Future work can be the study of other distribution laws, such as asymmetric Gaussian and mixed Gaussian laws [15]. Perspectives are the use of the proposed distribution laws in uncertainty quantification problems and the uncertainty propagation in nuclear reactor fuel cycle studies.

## References

1. Cross Sections Evaluation Working Group, *ENDF-6 Formats Manual* (National Nuclear Data Center Brookhaven National Laboratory, 2012)
2. M.A. Kellet, O. Bersillon, R.W. Mills, *The JEFF-3.1/-3.1.1 Radioactive Decay Data and Fission Yields Sub-libraries* (OECD, 2009)
3. M.B. Chadwick, M. Herman, ENDF/B-VII.1 Nuclear data for science and technology: cross sections, covariances, fission product yields and decay data, *Nucl. Data Sheets* **112**, 2887 (2011)
4. J. Katakura, Data/Code, Technical Report 2011-025, JAEA, 2012
5. S. Lahaye, P. Bellier, H. Mao, A. Tsilanizara, Y. Kawamoto, First verification and validation steps of MENDEL release V1.0 cycle code system, in *PHYSOR 2014 – The role of reactor physics toward a sustainable future* (Kyoto, Japan, 2014)
6. H. Golfier, R. Lenain, J.J. Lautard, P. Fougeras, P. Magat, E. Martinolli, Y. Dutheillet, APOLLO3: a common project of CEA, AREVA and EDF for the development of new deterministic multi-purpose code for physics analysis, in *M&C 2009* (New York, USA, 2009)
7. E. Brun, F. Damian, C.M. Diop, E. Dumonteil, F.X. Hugot, C. Jouanne, Y.K. Lee, F. Malvagi, A. Mazzolo, O. Petit, J.C. Trama, T. Visonneau, A. Zoia, TRIPOLI-4<sup>®</sup>, CEA, EDF and AREVA reference Monte Carlo code, *Ann. Nucl. Energy* **82**, 151 (2015)
8. A. Tsilanizara, C.M. Diop, B. Nimal, M. Detoc, L. Luneville, M. Chiron, T.D. Huynh, I. Bresard, M. Eid, J.C. Klein, DARWIN: an evolution code system for a large range of applications, *Nucl. Sci. Technol. Suppl.* **1**, 845 (2000)
9. A. Tsilanizara, N. Gilardi, T.D. Huynh, C. Jouanne, S. Lahaye, J.M. Martinez, C.M. Diop, Probabilistic approach for decay heat uncertainty estimation under URANIE platform by using MENDEL depletion code, *Ann. Nucl. Energy* **90**, 62 (2016)
10. F. Gaudier, URANIE: the CEA/DEN Uncertainty and Sensitivity platform, *Procedia Soc. Behav. Sci.* **2**, 7660 (2010)
11. M.D. McKay, R.J. Beckmana, W.J. Conover, Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* **21**, 239 (1979)
12. S. Lahaye, T.D. Huynh, A. Tsilanizara, Comparison of deterministic and stochastic approaches for isotopic concentration and decay heat uncertainty quantification on elementary fission pulse, *EPJ Web Conf.* **111**, 09002 (2016)
13. C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 379 (1948)
14. E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev. Ser. II* **106**, 620 (1957)
15. J.V. Michalowicz, J.M. Nichols, F. Bucholtz, Calculation of differential entropy for a mixed Gaussian distribution, *Entropy* **10**, 200 (2008)

Cite this article as: Sébastien Lahaye, Choice of positive distribution law for nuclear data, EPJ Nuclear Sci. Technol. 4, 38 (2018)