https://doi.org/10.61092/iaea.6ph2-gb9d

INDC(NDS)-0729
Distr. J,NM,SD,ST

INDC International Nuclear Data Committee

Testing the Goodness of Gaussian and Lognormal Emulators via Their Statistically Converged Probability Distribution Moments

Donald L. Smith
Argonne Associate of Seville
Argonne National Laboratory,
Coronado, CA 92118, U.S.A.

Denise Neudecker
Los Alamos National Laboratory
Los Alamos, NM 87545, U.S.A.

Roberto Capote Noy

March 2017

IAEA Nuclear Data Section
Vienna International Centre, P.O. Box 100, 1400 Vienna, Austria

Selected INDC documents may be downloaded in electronic form from
http://www-nds.iaea.org/publications
or sent as an e-mail attachment.
Requests for hardcopy or e-mail transmittal should be directed to
NDS.Contact-Point@iaea.org
or to:
Nuclear Data Section
International Atomic Energy Agency
Vienna International Centre
PO Box 100
1400 Vienna
Austria

Printed by the IAEA in Austria

March 2017

INDC(NDS)-0729
LA-UR-17-21103

Testing the Goodness of Gaussian and Lognormal Emulators via Their Statistically Converged Probability Distribution Moments

Donald L. Smith
Argonne Associate of Seville
NE Division, Argonne National Laboratory, Coronado, CA 92118, U.S.A.

Denise Neudecker
XCP Division, Los Alamos National Laboratory
Los Alamos, NM 87545, U.S.A.

Roberto Capote-Noy
Nuclear Data Section, International Atomic Energy Agency, A-1400 Vienna, Austria

ABSTRACT

Unified Monte Carlo (UMC) refers to a data evaluation technique that has been investigated during the past decade as a way to avoid the need to linearize relationships between primary and derived variables, as is necessary to apply least-squares procedures. UMC is based on Bayes' Theorem, whereby prior knowledge (generally theoretically derived) is embodied in a prior probability distribution function (PDF) while new experimental information is represented by a likelihood function. The product of these two functions forms a posterior PDF that can be used to provide estimates of evaluated mean values and uncertainties as well as additional quantities of applied interest. Probability functions can be characterized by their moments, and numerical values for these moments can be estimated by stochastically analyzing Markov chains of values of variables generated by sampling these PDFs using Monte Carlo techniques. In the present work we investigate two unrelated issues that are relevant to UMC: (1) Determine how many Monte Carlo histories are needed to obtain adequate estimates of the lowest-order moments of typical PDFs (mean values, standard deviations, skewness, and kurtosis). (2) Investigate whether Gaussian or lognormal functions could be used to emulate realistic model-generated PDFs in practical applications. For simplicity, issue (1) of this work considers only single-variable (1-D) situations. Both hypothetical data that are explicitly normally or lognormally distributed and prompt fission neutron spectra (PFNS) data calculated using the Los Alamos model are used in this work. We have shown that mean values and variances can often be estimated with adequate precision by considering only a few hundred sample histories. It requires larger numbers of sample histories to generate adequate estimates of skewness and kurtosis from stochastic analyses of Markov Chains. Depending on the circumstances, we have determined that no fewer than 1,000 to 5,000 sample histories are necessary to obtain acceptable estimates of skewness and excess kurtosis (kurtosis–3). Our present study also demonstrates that for PFNS data generated using the Los Alamos model, neither Gaussian nor lognormal probability functions can be used for a broad outgoing neutron energy range to adequately emulate computationally determined PDFs.

March 2017

Contents

1. Introduction

We recently investigated the effects of prior probability distribution function (PDF) shapes on the outcomes of simple evaluation exercises that involve only a single variable (1-D) to be evaluated from consideration of one hypothetical model-calculated value along with one hypothetical experimental value as input to the evaluation process. During this earlier study we discovered differences in the evaluated results that were produced by applying the traditional Generalized Least Squares (GLS) method as well as two specific formulations of Unified Monte Carlo (UMC-G and UMC-B). These differences could be attributed primarily to shape differences for the assumed prior PDFs, and these effects are likely to be due to the presence or absence of PDF skewness. The report that documents our earlier work includes extensive discussions of the background leading up to the development of contemporary UMC methods as well as of several general aspects of Bayesian data evaluation [1]. We will not repeat these discussions here.

When the UMC-G concept was originally introduced the prior PDF was taken to be normally distributed based on mean value and variance deduced from stochastic sampling data. This approximation discards much, potentially useful, shape information present in the true prior PDF, thereby suggesting limited applicability for UMC-G [2]. However, it was noted later that UMC-G in fact can be employed as an evaluation tool with greater flexibility than its original formulation if suitable analytic expressions for the prior PDF can be found that retain the essence of the true prior PDF shape, as reflected in the results of stochastic sampling of model-parameter space coupled with applications of nuclear models [3,4,5,6]. These methods, as well as several other approaches to nuclear data evaluation, including Monte Carlo ones, are discussed thoroughly in a thesis by Schnabel [7]. This thesis examines advantages and disadvantages of various evaluation concepts in their contemporary state of development.

The present work extends our earlier study that was conducted in one dimension (1-D) and reported in [1]. Here we consider statistical convergence issues related to stochastic determination of PDF moments, and we also explore whether analytic prior PDF shapes might be identified that could serve to facilitate the use of a UMC-G type approach as an alternative to UMC-B. True symmetric "averaging" of model-calculated and experimental data could then be accomplished rather than using experimental information only to assign weights to model-generated results. There is no *a priori* reason to expect that the UMC-G Bayesian approach might be feasible in all situations. However, it does enable non-linear effects between the variables to be treated while GLS does not. Therefore, it has the potential to be more accurate. This consideration provides ample justification for conducting such an exploratory exercise.

The basic concept of Bayesian data evaluation for a single variable is first outlined briefly in order to illustrate the role that an emulator prior PDF might play. We then summarize the key features of a single-variable lognormal function and indicate why it might be appealing to consider it as a potential emulator of prior PDFs. An extensive examination is then undertaken of convergence issues that would need to be considered in establishing whether or not an analytical PDF such as the lognormal function would be acceptable (or not) as an emulator of a true PDF that is embodied in a Markov Chain of values derived from stochastic sampling. The importance of this stochastic exercise is that any attempt to identify a potential analytical emulator PDF would need to rely on considerable Monte Carlo analysis. Therefore, understanding the level of computational effort that would be entailed in establishing the requisite numbers of stochastic histories needed to achieve reasonable convergence is crucial to making decisions regarding the practicality of selecting an emulator PDF. For convenience, lognormal PDFs are employed in this convergence test study. Finally, this approach is applied to actual sampling data corresponding to prior PDFs for theoretically calculated prompt fission neutron spectra (PFNS).

2. Basic Concepts

Let y be a single random variable that corresponds to a physical quantity that is to be evaluated. Also, let $p(\mathrm{y})$ be a PDF that defines the statistical properties of y. Furthermore, assume that $p(\mathrm{y})$ is a posterior PDF that is formed from the product of a prior PDF $p_0(\mathrm{y}|\mathrm{T})$ and a likelihood function $L(\mathrm{y}|\mathrm{E})$, according to Bayes Theorem, where "T" signifies theory and "E" signifies experiment. However, for convenience, use of "T" and "E" is dropped in the ensuing discussion since the point has now been made regarding typical dependence of the prior on theory and of the likelihood on experiment. Thus, we can write $p(\mathrm{y}) = \mathrm{C}_0\, p_0(\mathrm{y})\, L(\mathrm{y})$. We will assume that both p_0 and L are integrally normalized, while C_0 is a constant that is introduced to insure that the posterior PDF $p(\mathrm{y})$ is also normalized, i.e., that $\int p(\mathrm{y})\, \mathrm{dy} = 1$ when integration extends over the entire range of y where the magnitude of $p(\mathrm{y})$ is non-negligible.

Furthermore, let f(y) correspond to any well-defined function of the random variable y. The expectation value of f is defined as $\langle f\rangle = \int \mathrm{f(y)}\, p(\mathrm{y})\, \mathrm{dy}$ when $p(\mathrm{y})$ is normalized. If f(y) = y, $\langle f\rangle$ corresponds to the mean value of y. It can be written as $\langle y\rangle$ or MV{y}. If $\mathrm{f(y)} = \mathrm{y}^2 - \langle y\rangle^2$, Then $\langle f\rangle$ corresponds to the variance of y. This can be written as Var{y}. The standard deviation of y is defined as $[\mathrm{Var\{y\}}]^{½}$. It can be written as Std{y}. In addition, the skewness and kurtosis, i.e., Skew{y} and Kurt{y}, respectively, can be determined as discussed in [1] and repeated below for convenience.

The mathematical formulations described in the preceding paragraph are based on possessing an analytical expression for the prior PDF $p_0(\mathrm{y})$ so that the indicated integrals can be computed. In typical evaluation situations it might be the case symbolically that y = M(x), where x represents a model parameter of the model algorithm M that relates variables x and y. An important case that was examined in [1] is y = M(x) = exp(c x), with c acting as a scaling constant. Clearly the relationship between y and x is non-linear. If x is normally distributed, then y is precisely lognormally distributed, i.e., $p_0(\mathrm{y})$ is clearly a lognormal PDF.

In general this will rarely be the case for typical model algorithms M. More typically in nuclear data applications, the PDF for variable x is assumed (often for lack of any reason to choose an alternative PDF) to be Gaussian-distributed. This Gaussian PDF can be denoted by r(x). A Markov Chain of derived random values $\{\mathrm{y}_k\}$ can be generated using the expression $\mathrm{y}_k = \mathrm{M}(\mathrm{x}_k)$, for k=1,K (large K), where the collection $\{\mathrm{x}_k\}$ is generated by sampling the space of variable x according to PDF r(x). Knowledge of the true prior PDF $p_0(\mathrm{y})$ for derived y is embodied stochastically in the collection $\{\mathrm{y}_k\}$, especially when K is very large. It is unlikely that this PDF will correspond explicitly to any known analytical function.

However, specific numerical values for mean value, variance, standard deviation, skewness, and kurtosis, i.e., the most important population moments of $p_0(\mathrm{y})$, can be estimated stochastically from the collection $\{\mathrm{y}_k\}$. The following formulas are applicable only for very large K so that the distinction between sample and population moments is then negligible.

$$\mathrm{MV\{y\}} \approx \lambda_1 \equiv [\Sigma_{k=1,K}\, \mathrm{y}_k]/\mathrm{K}\,, \tag{1}$$

$$\mathrm{Var}\{y\} \approx \lambda_2 \equiv [\Sigma_{k=1,K} (y_k-\lambda_1)^2]/K \,, \tag{2}$$

$$\mathrm{Skew}\{y\} \approx \lambda_3 \equiv \{[\Sigma_{k=1,K} (y_k-\lambda_1)^3]/K\} \,/\, \lambda_2^{3/2} \,, \tag{3}$$

$$\mathrm{Kurt}\{y\} \approx \lambda_4 \equiv \{[\Sigma_{k=1,K} (y_k-\lambda_1)^4]/K\} \,/\, \lambda_2^{2} \;. \tag{4}$$

The symbol "≡" represents "defined as", and the λ_n (n=1,4) are stochastic estimators for mean value, variance, skewness, and kurtosis, respectively. Also, note that both skewness and kurtosis, as defined in Eqs. (3) and (4), are dimensionless. In the following discussion the terms MV{y}, Std{y} (or Std{y} in percent), Skew{y}, and Kurt{y} are used to represent both analytically derived and stochastically derived values, and no further use is made of λ_n (n=1,4). The distinction between these dual usages is discerned from the context.

The larger K becomes, the closer the estimates provided by Eqs. (1)–(4) will approach the correct values of the moments for the underlying prior PDF. It is proposed here to examine whether an analytical prior PDF, $p_{0S}(y)$, for which the corresponding mean value, variance, standard deviation, skewness, and kurtosis closely resemble those for the true $p_0(y)$, as deduced stochastically from the collection $\{y_k\}$, could serve as an effective emulator for the true prior PDF $p_0(y)$.

The challenge is to identify such an analytical function $p_{0S}(y)$ that could fulfill this requirement effectively in certain evaluation situations. The emulator posterior PDF is $p_S(y) = C_S\, p_{0S}(y)\, L(y)$, where, as mentioned earlier in this report, C_S is a constant that is introduced to insure that $p_S(y)$ is normalized. The intent here is to explore whether the emulator posterior PDF is able to yield acceptable values of skewness and kurtosis for evaluation purposes by comparison to those for the true prior PDF $p_0(y)$ when both functions exhibit the same mean value and standard deviation.

3. Single Variable Lognormal PDF

The single variable lognormal PDF is described in detail in [1]. This function possesses several desirable features worthy of its consideration as a potential prior PDF emulator.

Feature 1) It has only two variable parameters that are uniquely defined by the mean value and variance through analytical formulas. Furthermore, reciprocal formulas readily yield mean value and variance from the given parameters of the lognormal PDF.

Feature 2) The shape of the lognormal PDF is uniquely defined by the ratio of the variance to the mean value, or more intuitively, by the ratio of the standard deviation to the mean. This is also true for the Gaussian PDF. The ratio of the standard deviation to the mean value is the fractional standard deviation (or percent standard deviation). The scale of the lognormal function depends on the actual magnitude of the mean value, but this does not influence the PDF shape once the percent standard deviation is specified.

Feature 3) The dimensionless skewness and kurtosis for the lognormal PDF can be calculated easily from analytical formulas.

Feature 4) The magnitudes of both skewness and kurtosis vary predictably with the fractional standard deviation. Both skewness and kurtosis impact the PDF shape. This offers possibilities for this function to resemble arbitrary true PDF shapes. In contrast, all Gaussian PDFs are symmetric (zero skewness) with kurtosis that is always exactly equal to 3.

Feature 5) The lognormal PDF converges to the Gaussian PDF in the limit of small fractional standard deviation. For practical purposes, it is likely that there is no advantage to considering a lognormal PDF rather than a Gaussian PDF when data uncertainties are less than 10%.

It is instructive to list the major formulas associated with the lognormal PDF in spite of the fact that this repeats information that appears in [1]. A normalized lognormal prior PDF is defined by:

$$p_0(\mathrm{y}) = \exp[-0.5\ (\mathit{log}\ \mathrm{y} - \mu)^2/\sigma^2]\ /\ (2\pi\sigma^2\mathrm{y}^2)^{1/2}\ . \tag{5}$$

Here, "*log*" signifies "natural logarithm function". For convenience, let m = MV{y} and v = Var{y}. If m and v are provided for a lognormal PDF of Eq. (5), and μ and σ are the parameters of this function, the following formulas define the relationships between m, v, μ, and σ [8, 9]:

$$\mu = \mathit{log}[\mathrm{m}^2/(\mathrm{v} + \mathrm{m}^2)^{1/2}]\ , \tag{6}$$

$$\sigma = \{\mathit{log}[(\mathrm{v}/\mathrm{m}^2) + 1]\}^{1/2}\ , \tag{7}$$

$$\mathrm{m} = \exp[\mu + (\sigma^2/2)]\ , \tag{8}$$

$$\mathrm{v} = \exp(2\mu + \sigma^2)\ [\exp(\sigma^2) - 1]\ . \tag{9}$$

Analytical expressions for skewness and kurtosis of a lognormal PDF are as follows:

$$\mathrm{Skew\{y\}} = [\exp(\sigma^2) + 2]\ [\exp(\sigma^2) - 1]^{1/2}\ , \tag{10}$$

$$\mathrm{Kurt\{y\}} = \exp(4\sigma^2) + 2\ \exp(3\sigma^2) + 3\ \exp(2\sigma^2) - 3\ . \tag{11}$$

The so-called excess kurtosis is defined as "Kurt{y}–3" where Kurt{y} is given by Eq. (11). This factor is a useful measure of shape difference between a lognormal PDF and a Gaussian PDF since the latter always has kurtosis equal to 3. In the limit of very small fractional standard deviation, the excess kurtosis approaches zero since the lognormal PDF approaches the Gaussian PDF under these circumstances.

Table 1 gives values of the dimensionless lognormal PDF skewness, kurtosis, and excess kurtosis calculated using the formulas above for various selected values of standard deviation Std{y}, expressed in percent, over the nominal range 0.1–1,000%. Figure 1 shows the same information as Table 1 in graphical form, to aid in comprehension, for percent Std{y} values up to 500%.

Table 1: *Calculated values of skewness, kurtosis, and excess kurtosis of a lognormal PDF for selected percent standard deviations.*

Std{y} (%)	Skew{y}	Kurt{y}	Kurt{y}-3
0.1	3.0000E-03	3.0000E+00	1.6000E-05
0.15	4.5000E-03	3.0000E+00	3.6000E-05
0.2	6.0000E-03	3.0001E+00	6.4000E-05
0.3	9.0000E-03	3.0001E+00	1.4400E-04
0.5	1.5000E-02	3.0004E+00	4.0001E-04
0.75	2.2500E-02	3.0009E+00	9.0005E-04
1	3.0001E-02	3.0016E+00	1.6002E-03
1.5	4.5003E-02	3.0036E+00	3.6008E-03
2	6.0008E-02	3.0064E+00	6.4024E-03
3	9.0027E-02	3.0144E+00	1.4412E-02
5	1.5012E-01	3.0401E+00	4.0094E-02
7.5	2.2542E-01	3.0905E+00	9.0476E-02
10	3.0100E-01	3.1615E+00	1.6151E-01
15	4.5338E-01	3.3677E+00	3.6766E-01
20	6.0800E-01	3.6644E+00	6.6439E-01
30	9.2700E-01	4.5659E+00	1.5659E+00
50	1.6250E+00	8.0352E+00	5.0352E+00
75	2.6719E+00	1.7914E+01	1.4914E+01
100	4.0000E+00	4.1000E+01	3.8000E+01
150	7.8750E+00	2.0891E+02	2.0591E+02
200	1.4000E+01	9.4700E+02	9.4400E+02
300	3.6000E+01	1.2297E+04	1.2294E+04
500	1.4000E+02	4.9415E+05	4.9415E+05
750	4.4438E+02	1.1128E+07	1.1128E+07
1000	1.0300E+03	1.0615E+08	1.0615E+08

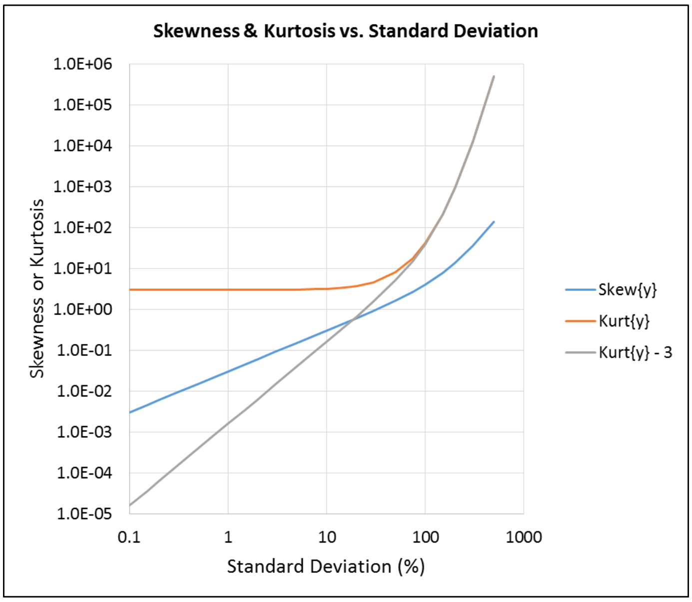

Fig. 1: *Analytically calculated values of skewness, kurtosis, and excess kurtosis for a lognormal PDF as a function of percent standard deviation. These moments are characteristic of the inherent shape of the function and they do not depend on scale, i.e., on the actual magnitudes of the mean value MV{y} or standard deviation Std{y}.*

It is apparent that the magnitude of kurtosis changes very slowly as a function of nominal percent standard deviation below about 10%. Both skewness and excess kurtosis change rapidly with percent standard deviation over the range considered, especially for percent

Std{y} values above 100%. Figure 1 suggests that over short ranges of percent Std{y} the curves are approximately linear on a log-log scale. Thus Table 1 might be useful, along with log-log interpolation, to provide approximate values of Skew{y} and Kurt{y} as an alternative to employing Eqs. (6)–(11) for percent Std{y} values not listed in Table 1.

4. Stochastic Convergence Tests for Lognormal PDFs

It is of interest to determine whether an arbitrary prior PDF $p_0(y)$ generated in the form of a Markov Chain of sample values from nuclear modeling can be shown to resemble a particular analytical PDF. One way to approach this matter is to examine whether the known moments of an analytical PDF can be determined stochastically with sufficient accuracy from a Markov Chain of sample values generated in accordance with that known PDF.

***Table 2**: Comparison of analytically and stochastically derived values of skewness, kurtosis, and excess kurtosis for percent standard deviations Std{y} ranging from 0.1–300%. The cells shaded in light green indicate regions of percent Std{y} in which skewness and excess kurtosis can be determined stochastically with adequate reliability for $K=10^6$ samples. For present purposes, "adequate reliability" is defined arbitrarily as agreement of stochastically determined values with comparable analytically determined ones to within 10%.*

c	Std{y} Stoch (%)	Std{y} Det (%)	Skew{y} Det	Kurt{y} Det	Kurt{y}-3 Det	Skew{y} Stoch	Kurt{y} Stoch	Kurt{y}-3 Stoch	Skew Ratio	Kurt-3 Ratio
0.01	0.10	0.10	3.0000E-03	3.0000E+00	1.6000E-05	5.7863E-03	3.0065E+00	6.4614E-03	1.93	403.84
0.015	0.15	0.15	4.5000E-03	3.0000E+00	3.6000E-05	6.1673E-03	3.0014E+00	1.3734E-03	1.37	38.15
0.02	0.20	0.20	6.0000E-03	3.0001E+00	6.4000E-05	8.1061E-03	2.9940E+00	Negative	1.35	Negative
0.03	0.30	0.30	9.0000E-03	3.0001E+00	1.4400E-04	1.0292E-02	3.0014E+00	1.4248E-03	1.14	9.89
0.05	0.50	0.50	1.5000E-02	3.0004E+00	4.0001E-04	1.3205E-02	3.0082E+00	8.2088E-03	0.88	20.52
0.075	0.75	0.75	2.2500E-02	3.0009E+00	9.0005E-04	2.0602E-02	3.0085E+00	8.4529E-03	0.92	9.39
0.1	1.00	1.00	3.0001E-02	3.0016E+00	1.6002E-03	2.9868E-02	3.0000E+00	4.2259E-05	1.00	0.03
0.15	1.50	1.50	4.5003E-02	3.0036E+00	3.6008E-03	4.7947E-02	3.0024E+00	2.3888E-03	1.07	0.66
0.2	2.00	2.00	6.0008E-02	3.0064E+00	6.4024E-03	5.9869E-02	3.0084E+00	8.3900E-03	1.00	1.31
0.3	3.00	3.00	9.0027E-02	3.0144E+00	1.4412E-02	9.2118E-02	3.0093E+00	9.2558E-03	1.02	0.64
0.5	5.00	5.00	1.5012E-01	3.0401E+00	4.0094E-02	1.4948E-01	3.0392E+00	3.9212E-02	1.00	0.98
0.75	7.51	7.50	2.2542E-01	3.0905E+00	9.0476E-02	2.2715E-01	3.0964E+00	9.6428E-02	1.01	1.07
0.998	10.00	10.00	3.0100E-01	3.1615E+00	1.6151E-01	2.9896E-01	3.1614E+00	1.6137E-01	0.99	1.00
1.493	15.01	15.00	4.5338E-01	3.3677E+00	3.6766E-01	4.5555E-01	3.3731E+00	3.7310E-01	1.00	1.01
1.98	19.99	20.00	6.0800E-01	3.6644E+00	6.6439E-01	6.0709E-01	3.6624E+00	6.6243E-01	1.00	1.00
2.94	30.06	30.00	9.2700E-01	4.5659E+00	1.5659E+00	9.3169E-01	4.5789E+00	1.5789E+00	1.01	1.01
4.73	50.07	50.00	1.6250E+00	8.0352E+00	5.0352E+00	1.6486E+00	8.3639E+00	5.3639E+00	1.01	1.07
6.663	75.00	75.00	2.6719E+00	1.7914E+01	1.4914E+01	2.7303E+00	1.9274E+01	1.6274E+01	1.02	1.09
8.31	99.81	100.00	4.0000E+00	4.1000E+01	3.8000E+01	3.9122E+00	3.5985E+01	3.2985E+01	0.98	0.87
10.88	150.18	150.00	7.8750E+00	2.0891E+02	2.0591E+02	7.5014E+00	1.4641E+02	1.4341E+02	0.95	0.70
12.78	200.34	200.00	1.4000E+01	9.4700E+02	9.4400E+02	1.1152E+01	3.1049E+02	3.0749E+02	0.80	0.33
15.2	301.25	300.00	3.6000E+01	1.2297E+04	1.2294E+04	2.5520E+01	1.6690E+03	1.6660E+03	0.71	0.14
Det => Deterministic		Stoch => Stochastic		Skew Ratio = Skew{y} Stoch / Skew{y} Det			Kurt-3 Ratio = Kurt{y}-3 Stoch / Kurt{y}-3 Det			

An ideal opportunity to test this conjecture is by considering the lognormal PDF. The convergence tests reported in this section were performed using a procedure described in [1]. Markov Chains of values $\{y_k\}$ (k=1,K) were generated to resemble what would be encountered by sampling from a lognormal PDF. A collection of sample values $\{x_k\}$ was produced from a Gaussian PDF with MV{x} = 1 and Std{x} = 0.1 (10% standard deviation). It was assumed that y = M(x) = exp(c x), where various value for constant "c" were considered in the exercise.

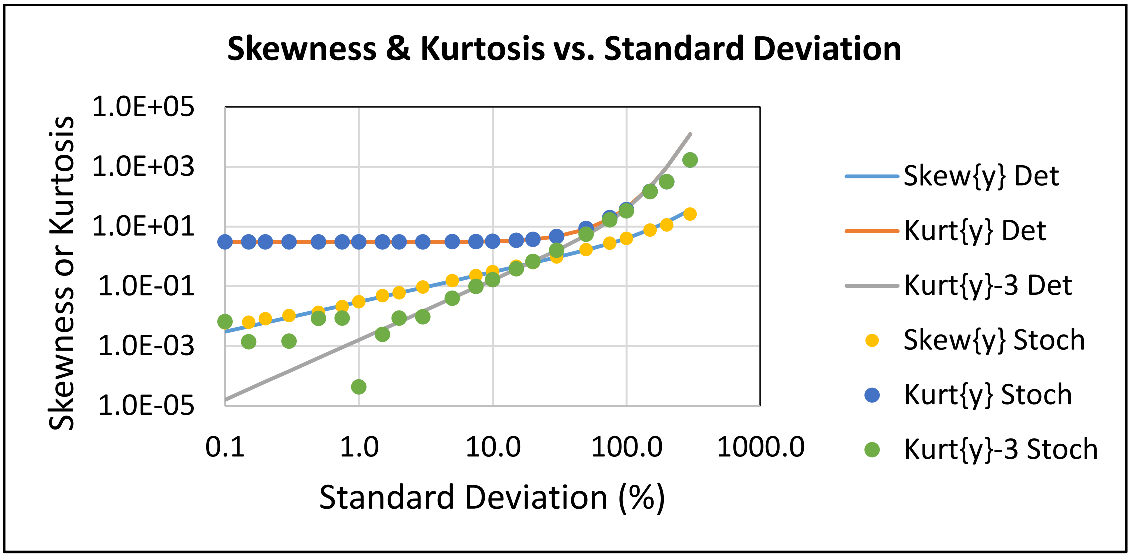

***Fig. 2**: Deterministically and stochastically generated values of Skew{y}, Kurt{y}, and Kurt{y}–3 (excess kurtosis) for $K=10^6$ samples.*

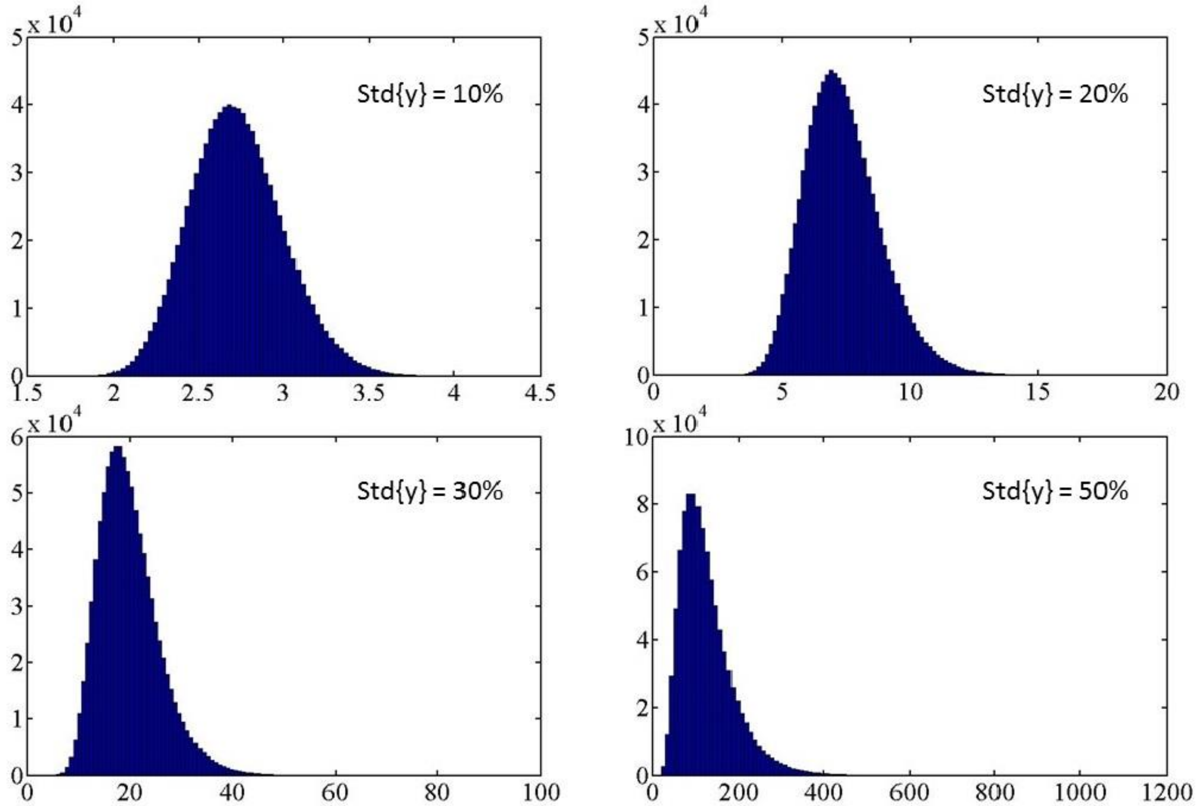

***Fig. 3**: Plots of 100-bin histograms of $K=10^6$ collections of lognormally distributed values $\{y_k\}$ for nominal Std{y}=10, 20, 30, and 50%. The abscissa scale corresponds to values of variable y while the ordinate scale indicates sample events per histogram bin. Note that the PDF for Std{y} = 10% appears to be only slightly skewed toward larger values of y_k whereas the PDF for Std{y} = 50% is severely skewed. PDFs for values of Std{y} smaller than 10% would not be visually distinguishable from symmetric Gaussian PDFs.*

Values y_k in these Markov Chains were derived from the expression $y_k = M(x_k)$. The values of constant "c" were selected so that the stochastically calculated MV{y} and Var{y} yielded corresponding stochastically determined fractional Std{y} values very close to those values given in Table 1 when a very large number K of Monte Carlo histories was traced. In particular, this analysis was performed on a PC using MATLAB [9] with $K=10^6$ samples for each selected value of "c". It was possible to use "c" as a tool to accomplish this objective because of standard deviation amplification (if c > 1) or suppression (if c < 1) effects, as described in [1]. Table 2 provides the results of this analysis over the range 0.1–300% in nominal percent Std{y}. This information is also represented graphically in Figs. 2 and 3.

<u>Conclusion</u>: *It is evident that even for as many as $K=10^6$ samples, the results from stochastic determinations of skewness and excess kurtosis for a lognormal PDF appear to be reliable (e.g., differing by no more than 10% from the known analytically determined values) <u>only</u> for values of Std{y} within the ranges 0.75–150% (for Skew{y}) and 5–75% (for Kurt{y}–3). As discussed in [1], it is very difficult to calculate values of skewness and excess kurtosis accurately by stochastic means when Std{y} in percent is either very small or relatively large.*

In practical situations, it is unlikely that there would be a need to consider data with uncertainties larger than those indicated by the upper limits of these ranges. In fact, for practical applications the range 10–50% of nominal Std{y} is probably the most important one. Data with uncertainties larger than 50% are unlikely to play significant roles in realistic evaluations. Realistic nuclear model analyses involve algorithms that propagate multiple model parameters to multiple derived observables. This is a far more complex and computationally intensive process than the simple 1-D analysis involving an exponential function discussed above. Therefore, the emulator approach described in the present work can be shown to be practical only if reasonable results can be obtained for data sets with typical uncertainties when fewer sampling histories are considered, e.g., for K not exceeding 1,000 to 5,000 histories. The following analyses explore this issue in some detail by once again employing lognormal PDFs. Two hundred sets of Monte Carlo simulations were performed using the lognormal PDF. They involve ten distinct values of c = 0.3, 0.5, 0.75, 0.998, 1.493, 1.98, 2.94, 4.73, 6.663, and 8.31, with K=1,000 or 5,000 histories. These selected values of c form a subset of the list given in Table 1. They correspond to the nominal standard deviations of 3, 5, 7.5, 10, 15, 20, 30, 50, 75, and 100% that are included in Table 1. Representative PDF histograms from $K=10^6$ and 5,000 and 1,000 sampling histories for Std{y} = 30% are shown in Fig. 4.

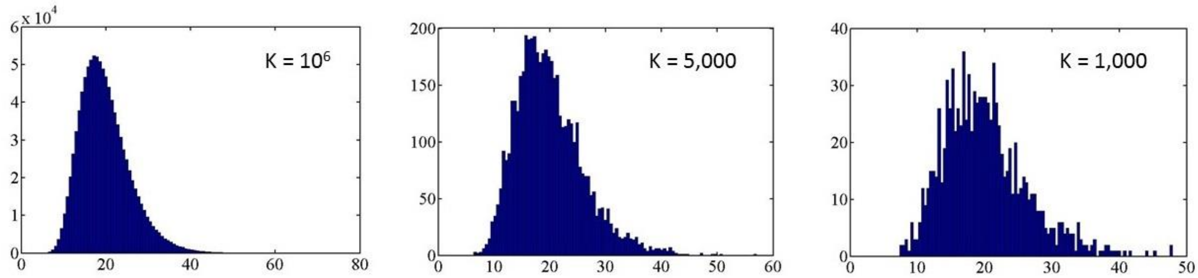

***Fig. 4**: Plots of 100-bin PDF histograms from $K=10^6$ and 5,000 and 1,000 sampling histories for nominal Std{y} = 30%. The abscissa scale corresponds to values of variable y while the ordinate scale indicates sample events per histogram bin. Although the histograms for K=1,000 and 5,000 are considerably more irregular than the one for $K=10^6$, nevertheless they do reflect the essential shape of the lognormal PDF that is shown much more smoothly in the $K=10^6$ history profile.*

Ten independent but repetitive Monte Carlo simulations were performed for each value of "c" and K. This accounts for the abovementioned 200 simulations. For each such simulation set, sample averages of the ten stochastically derived values for MV{y}, Std{y}, Std{y} in percent, Skew{y}, and Kurt{y}–3 were determined along with corresponding sample standard deviations. These averages of ten values are denoted by "AVG" while the sample standard deviations for the ten are denoted by "STD".

These calculations were performed using algorithms that are available in MATLAB and Microsoft Excel. The numerical results from this statistical exercise are tabulated in the Appendix. Keep in mind when considering the tabulated values of AVG and STD that, e.g., ten independent repetitions of 5,000 Monte Carlo simulations correspond to performing 50,000 simulations.

The objective of performing ten independent simulations for each choice of "c" and K was to assess stochastic scatter in the obtained results. This provides a reasonable measure of the reliability of results for the PDF moments that might be obtained for any single simulation from the collection of ten. This scatter is clearly reflected in STD (%). The reader is reminded that this particular exercise was conceived as a means of assessing how many Monte Carlo samplings of an arbitrary prior PDF would be needed in order to provide a meaningful comparison of its moments with those of a potential lognormal PDF that might serve as an emulator prior PDF.

The purpose for performing a stochastic analysis of an arbitrary prior PDF generated from nuclear modeling would be to determine whether its skewness and kurtosis agree reasonably well (or not) with a corresponding emulator PDF (e.g., a lognormal function) having the same percent standard deviation. It is important to establish just how well the mean values and standard deviations can be determined stochastically for any stochastically generated prior PDF by considering a limited number of sampling histories. Obviously, $K=10^6$ is beyond practical feasibility limits. In the present investigation K=1,000 and 5,000 histories are considered, and these correspond to representative PDFs (which happen to be lognormal).

Table 3 suggests that the ratio Std{y}/MV{y}, which corresponds to Std{y} in percent, can be determined quite accurately with modest numbers of Monte Carlo histories, even for rather large values of nominal standard deviation. The entries in Table 3 can be understood as follows: The role of "c" in yielding lognormal PDFs with desired nominal standard deviations is clear from the preceding discussion. The quantity "AVG" is based on averages of sets of 10 simulations of K=1,000 or 5,000, respectively. The quantity "STD" is a measure of the scatter and thus reliability of any single stochastic simulation.

The most severe case is 100% nominal standard deviation. It is seen that a determination of the standard deviation from 1,000 samples might be unreliable by at most 10%. It is unlikely that an evaluator will be considering very many data with uncertainties exceeding 50%. The scatter in this case would be expected to be no more than about 3%, even for as few as K=1,000 sampling histories. So, for practical reasons the focus in the following analysis is on the range 10–50% nominal standard deviation.

Table 3: *Estimates are given of the scatter "STD" for stochastically derived percent Std{y} values based on ten trials of K=1,000 or 5,000 histories performed to estimate the true values of Std{y} by sampling known lognormal PDFs. These results are taken from the Appendix.*

Std{y} Percent		**K = 1,000**		**K = 5,000**	
Nominal Std{y} (%)	c	AVG (%)	STD (%)	AVG (%)	STD (%)
3%	0.3	3.00%	2.60%	2.98%	1.05%
5%	0.5	4.99%	1.10%	4.99%	1.07%
7.5%	0.75	7.42%	2.33%	7.52%	1.62%
10%	0.998	9.79%	2.69%	10.03%	1.36%
15%	1.493	14.95%	1.93%	15.04%	1.30%
20%	1.98	19.68%	3.61%	19.91%	1.44%
30%	2.94	30.10%	2.88%	30.04%	1.73%
50%	4.73	50.26%	3.22%	49.99%	1.79%
75%	6.663	73.87%	5.01%	74.63%	2.20%
100%	8.31	102.20%	10.30%	99.17%	3.78%

What this table demonstrates is that in the range of nominal PDF percent standard deviation from 10–50%, the scatter in stochastically determined values for percent Std{y} can range from 1.0–3.6% for K=1,000 histories. The corresponding scatter in the case for K=5,000 histories is 1.3–1.8%. These results suggest that acceptable (emphasized by light green filled cells) estimates of Std{y} can be achieved with modest scatter from analyses involving as few as K=1,000 sampling histories for the range of nominal standard deviation of interest for most applications. The present investigation also indicates that in such practical situations acceptable estimates of MV{y} most likely can be obtained from only a few hundred sample histories.

Similar analyses were performed to determine the scatter likely to be encountered in statistically estimated skewness and excess kurtosis values derived from stochastic sampling data. The results are shown in Tables 4 and 5. In these tables, light green filled cells indicate acceptable values of scatter less than 10%, light yellow filled cells indicate marginally acceptable values of scatter in the nominal standard deviation range 10–50%, while orange-brown filled cells indicates unacceptable values of scatter greater than 50% for that range.

Table 4: *Estimates are given of the percent scatter "STD" in stochastically derived Skew{y} for ten repeated trials of K=1,000 or 5,000 histories to determine the true value of Skew{y} from sampling known lognormal PDFs. These results are taken from the Appendix. Differences between AVGs and analytical values are also presented.*

Skew{y}		**K = 1,000**		Analy	Differ
Nominal Std{y} (%)	c	AVG (%)	STD (%)	Skew{y}	in % *
3%	0.3	0.1066356	58.59%	0.090027	-18.45%
5%	0.5	0.2098224	37.64%	0.150125	-39.77%
7.5%	0.75	0.2350123	26.03%	0.2254219	-4.25%
10%	0.998	0.2881653	32.66%	0.301	4.26%
15%	1.493	0.4652225	23.07%	0.453375	-2.61%
20%	1.98	0.5551559	18.01%	0.608	8.69%
30%	2.94	0.9971676	11.03%	0.927	-7.57%
50%	4.73	1.633515	18.58%	1.625	-0.52%
75%	6.663	2.4444728	18.41%	2.671875	8.51%
100%	8.31	3.809321	45.52%	4	4.77%
* Differ = (Analy Skew{y} - AVG)/Analy Skew{y}					

Skew{y}		**K = 5,000**		Analy	Differ
Nominal Std{y} (%)	c	AVG (%)	STD (%)	Skew{y}	in % *
3%	0.3	0.105898	30.00%	0.090027	-17.63%
5%	0.5	0.1686167	14.92%	0.150125	-12.32%
7.5%	0.75	0.2204237	9.04%	0.2254219	2.22%
10%	0.998	0.2914944	12.35%	0.301	3.16%
15%	1.493	0.4617234	9.53%	0.453375	-1.84%
20%	1.98	0.6203352	9.62%	0.608	-2.03%
30%	2.94	0.9179036	5.08%	0.927	0.98%
50%	4.73	1.5531083	6.95%	1.625	4.42%
75%	6.663	2.5605645	11.00%	2.671875	4.17%
100%	8.31	3.9885878	29.11%	4	0.29%
* Differ = (Analy Skew{y} - AVG)/Analy Skew{y}					

Conclusion: *Stochastic sampling with K=5,000 histories appears to be capable of yielding acceptable values of skewness for most of the range of nominal standard deviation from 10–50%, but the corresponding results for K=1,000 histories are marginally acceptable.*

***Table 5**: Estimates are given of the percent scatter "STD" in stochastically derived Kurt{y}–3 (excess kurtosis) for ten repeated trials of K=1,000 or 5,000 histories to determine the true value of excess kurtosis from sampling known lognormal PDFs. These results are taken from the Appendix. Differences between the AVGs and analytical values are also presented.*

Kurt{y} -3		K = 1,000		Analy	Differ
Nominal Std{y} (%)	c	AVG	STD (%)	Kurt{y}-3	in % *
3%	0.3	0.0028923	4267.78%	0.0144122	79.93%
5%	0.5	0.1801931	121.12%	0.0400938	-349.43%
7.5%	0.75	0.1121771	253.42%	0.0904757	-23.99%
10%	0.998	0.1592911	227.74%	0.161506	1.37%
15%	1.493	0.3875443	96.01%	0.3676624	-5.41%
20%	1.98	0.4354172	60.29%	0.6643866	34.46%
30%	2.94	2.049494	37.16%	1.5659396	-30.88%
50%	4.73	4.609746	46.87%	5.0351563	8.45%
75%	6.663	10.278251	44.21%	14.914078	31.08%
100%	8.31	28.824201	100.41%	38	24.15%
* Differ = (Analy Kurt{y}-3 - AVG)/Analy Kurt{y}-3					

Kurt{y} -3		K = 5,000		Analy	Differ
Nominal Std{y} (%)	c	AVG	STD (%)	Kurt{y}-3	in % *
3%	0.3	0.0289514	288.94%	0.0144122	-100.88%
5%	0.5	0.0617619	127.36%	0.0400938	-54.04%
7.5%	0.75	0.0723711	126.88%	0.0904757	20.01%
10%	0.998	0.1025939	93.45%	0.161506	36.48%
15%	1.493	0.3893479	31.03%	0.3676624	-5.90%
20%	1.98	0.7384295	34.21%	0.6643866	-11.14%
30%	2.94	1.5156075	21.70%	1.5659396	3.21%
50%	4.73	4.2649052	20.25%	5.0351563	15.30%
75%	6.663	12.53725	36.50%	14.914078	15.94%
100%	8.31	37.587754	85.66%	38	1.08%
* Differ = (Analy Kurt{y}-3 - AVG)/Analy Kurt{y}-3					

Conclusion: *Stochastic sampling with K=5,000 histories appears to be capable of yielding marginally acceptable values of excess kurtosis for most of the range of nominal standard deviation 10–50%, but the corresponding results for K=1,000 histories are generally unacceptable over most of this range.*

It is well-established, e.g., from [1], that skewness in the data PDFs can influence the outcome of an evaluation relative to what would be obtained for symmetric Gaussian PDFs with the same standard deviations. However, what is not well known is what the comparable effect of kurtosis might be on evaluated outcomes. Intuitively, one would expect that PDFs with no skewness, the same standard deviations, but differing values of kurtosis should yield rather similar results. However, this is an issue that needs to be investigated more thoroughly before deciding whether the sole focus in identifying potential emulator PDFs should be on equating mean value, standard deviation, and skewness for the true and emulator PDFs. That should be the topic of a future investigation.

5. Statistical Analyses of Data from PFNS Model Calculations

In this section, we analyze stochastic data produced by realistic nuclear model calculations. In particular, we consider prompt fission neutron spectrum (PFNS) values generated using the Los Alamos model [10] and emphasize the determination of skewness and kurtosis for the corresponding PDFs. These investigations strive to answer three questions:

(1) Are 1,000 or 5,000 samples of model generated values marginally sufficient to approximate the skewness and kurtosis calculated from a distinctly larger number of samples? Contrary to the studies reported in the preceding sections, which involved only Gaussian and lognormal probability functions, in this section we do not know the exact skewness and kurtosis of the model calculated values as there are no analytical expressions for the skewness and kurtosis that pertain to the PFNS model used. Therefore, the applicable numerical values of skewness and kurtosis of these model-calculated values must be approximated by those calculated stochastically from a large number of sampled model-generated values.

(2) Can the PDFs generated by the particular PFNS model used in this work be well-described by either a Gaussian or a lognormal emulator? If this happens to be true, then UMC-G, involving either a Gaussian or lognormal prior PDF, or GLS in either ordinary or log-space, could be used for PFNS evaluations with this particular model.

(3) Can the skewness and kurtosis derived from actual PFNS data calculated from this model be used to quantify how well these data are described by a chosen emulator PDF?

We investigated these questions by examining numerous statistically generated ^{239}Pu PFNS values predicted by the Los Alamos model [10] for an incident neutron energy of 0.5 MeV and the following collection of outgoing neutron energies *E*={0.0001, 0.001, 0.01, 0.1, 0.5, 1, 2, 2.24, 3, 4.5, 6, 8, 10, 15, 17.6} MeV. The PFNS at these particular outgoing neutron energies were chosen since they serve well to represent the characteristics of a PFNS, where the neutron yield is non-negligible, and the associated relative uncertainties span a large range from 2.4–132%, as can be seen in Fig. 5. A total of K=34,241 Monte-Carlo generated PFNS samples were produced for each outgoing neutron energy *E*. These samples were obtained in two steps. In the first step the Los Alamos model parameters were sampled within their uncertainties K=34,241 times. K=34,241 model-parameter vectors were then generated from the sampled parameters. These model-parameter vectors were used in the second step as input for the Los Alamos model that was used to calculate the K=34,241 PFNS vectors. The interested reader is referred to Ref. [11] for a more detailed description of how these PFNS vectors were calculated. Here we concentrate on answering the questions stated above.

5.1. Required Number of Sampling Histories

Question: Are 1,000 or 5,000 samples of model-generated values marginally sufficient to approximate the skewness and kurtosis calculated from a distinctly larger number of samples?

This question is examined by calculating mean values MV{y}, standard deviation Std{y}, skewness Skew{y}, and kurtosis Kurt{y} for each outgoing neutron energy *E*, first for K=1,000, 5,000 and 34,241 samples, and then for K=1,000*i* (*i*=1–34) samples. These values were calculated with the NumPy functions “mean” and “var” and the SciPy.stats functions “skew” and “kurtosis” implemented in the programming language Python [12]. The MV{y} and Std{y} values calculated from K=1,000 and 5,000 samples in Table 6 are acceptably close to those calculated from K=34,241 samples. The deviation from values calculated using K=34,241 samples is less than 3% for MV{y} and less than 5% for Std{y}. The deviation is largest where the PFNS neutron yield is very small, as is seen in Fig. 5. It is also obvious from Figs. 6 and 7 that MV{y} and Std{y} converge very rapidly towards those values calculated from K=34,241 samples.

However, Skew{y} values calculated from K=1,000 samples are only marginally acceptable compared to those values calculated from K=34,241; that is the deviation is in the range of 10–50%. K=5,000 samples yield acceptable values of Skew{y} with a deviation of less than 10%. The only exception is Skew{y} for *E*=2 MeV. At this *E*, Skew{y} converges very slowly towards the values calculated from K=34,241 samples, as is illustrated in Fig. 8, while Skew{y} converges rapidly at all other *E*. At *E*=2 MeV, Skew{y} assumes its lowest value. This slow convergence is an artifact that stems from observing the convergence as a ratio to Skew{y} calculated from K=34,241. Very small absolute changes in the values of two small numbers being divided can produce large effects on their ratio. In fact the Skew{y} value for K=5,000 at *E*=2 MeV is close to the value calculated from K=34,241 samples.

***Table 6**: MV{y} and Std{y} values given in this table were calculated from K=1,000, 5,000 and 34,241 model-predicted PFNS sample vectors that are dependent on the outgoing neutron energy E. MV{y} and Std{y} values for K=1,000 and 5,000 are also provided as ratios to those corresponding quantities calculated from K=34,241 PFNS samples.*

E (MeV)	MV{y} (MeV^{-1}) K=34,241	Ratio MV{y} K=1,000/34,241	Ratio MV{y} K=5,000/34,241	Std{y} (%) K=34,241	Ratio Std{y} K=1,000/34,241	Ratio Std{y} K=5,000/34,241
0.0001	0.005861	0.9965400	0.996733	12.59158	0.966198	0.982005
0.001	0.018530	0.9965398	0.996730	12.58064	0.966344	0.981937
0.01	0.058452	0.9966096	0.996772	12.51384	0.967419	0.982392
0.1	0.179822	0.9971190	0.997142	12.31883	0.977585	0.987052
0.5	0.333744	0.9982805	0.998093	13.38499	1.004679	1.001228
1	0.330704	0.9987725	0.998260	10.83695	1.006354	1.004709
2	0.241238	0.9999686	0.999830	3.541333	1.008290	1.011618
2.24	0.215008	1.0001882	1.000154	2.403518	0.985124	0.996894
3	0.140439	1.0008057	1.001070	8.234878	1.007308	0.996884
4.5	0.056421	1.0026311	1.003330	23.75321	1.002724	0.999851
6	0.022386	1.0050163	1.005853	37.67255	0.996269	0.998431
8	0.006158	1.0077442	1.009212	54.85443	0.988505	0.997986
10	0.001683	1.0097186	1.013572	70.88279	0.980820	0.998398
15	6.728299E-05	1.0105917	1.023952	109.8787	0.971508	0.993320
17.6	1.278100E-05	1.0091402	1.028958	131.5706	0.968145	0.989924

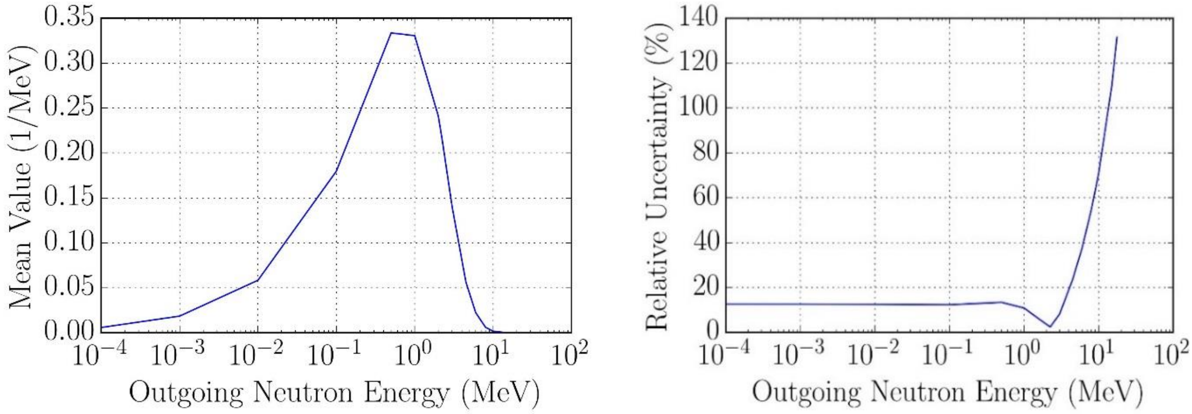

***Fig. 5**: MV{y} and Std{y} (in percent) calculated from K=34,241 samples are shown as a function of outgoing neutron energy E. These plots illustrate information provided in Table 6.*

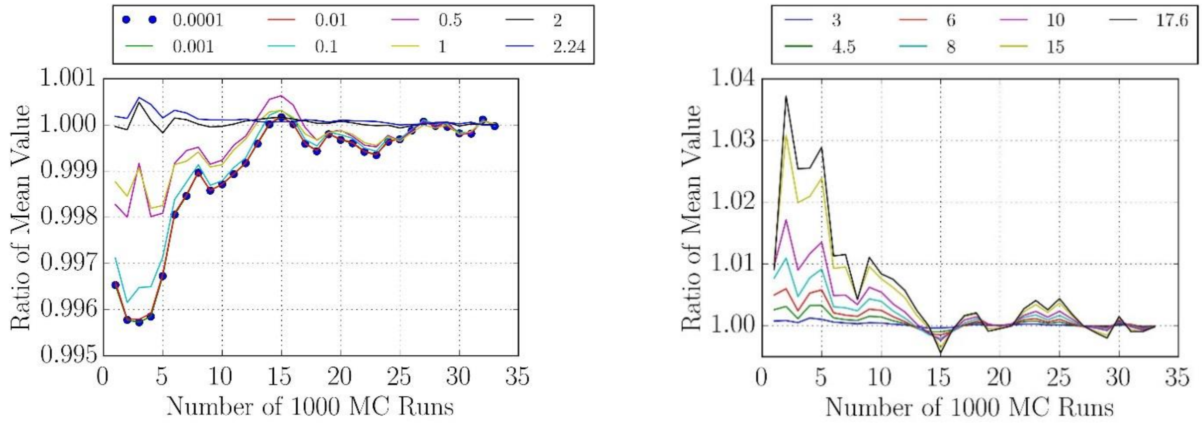

Fig. 6: *MV{y} values are shown for each E as a function of sample numbers K from which they were calculated. These values are given as ratios to MV{y} calculated from K=34,241 samples. The values of E in the legend are given in units of MeV. MV{y}(K) converges rapidly towards MV{y}(K=34,241), and K=1,000/5,000 samples yield acceptable values for MV{y}.*

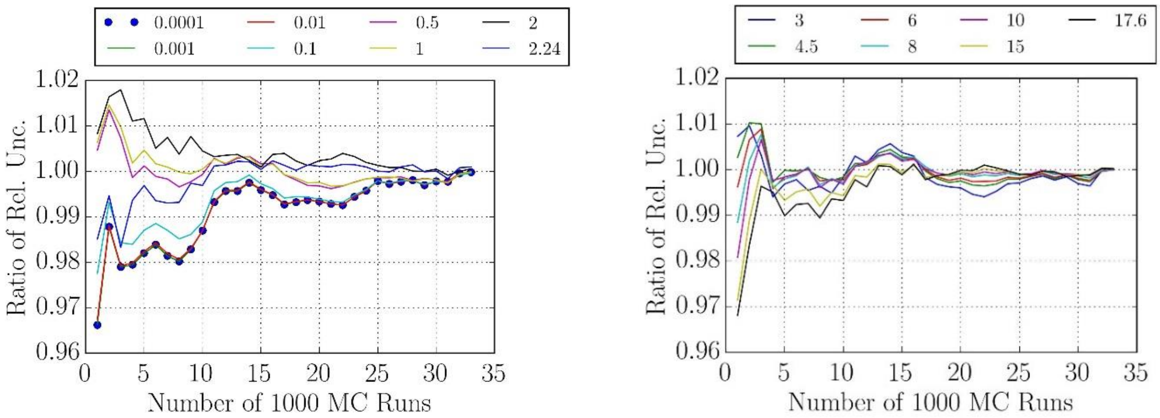

Fig. 7: *Std{y} values are shown as a function of sample numbers K from which they were calculated for each E. These values are given as ratios to Std{y} calculated from K=34,241 samples. The values of E in the legend are given in units of MeV. Std{y}(K) converges rapidly towards Std{y}(K=34,241), and K=1,000/5,000 samples yield acceptable values for Std{y}.*

Table 7: *Skew{y} and Kurt{y}-3 values for the indicated outgoing neutron energies E are given as calculated from K=1,000, 5,000 and 34,241 model-predicted PFNS sample vectors. Skew{y} and Kurt{y}-3 for K=1,000 and 5,000 samples are given as ratios to the K=34,241 values.*

E (MeV)	Skew{y} K=34,241	Ratio Skew{y} K=1,000/34,241	Ratio Skew{y} K=5,000/34,241	Kurt{y}-3 K=34,241	Ratio Kurt{y}-3 K=1,000/34,241	Ratio Kurt{y}-3 K=5,000/34,241
0.0001	0.497412	0.913935	0.981389	0.025455	-2.085910	0.853052
0.001	0.498720	0.918133	0.981238	0.026375	-1.697841	0.872524
0.01	0.500605	0.947218	0.981847	0.019130	0.335094	0.864214
0.1	0.567234	1.103421	0.995549	0.068218	4.747358	1.138593
0.5	0.816338	1.092405	1.040896	0.387228	1.325284	1.285237
1	0.746625	1.101503	1.035387	0.157048	1.417523	1.444791
2	-0.081892	0.115315	0.562504	-0.37252	1.066587	1.297996
2.24	-0.569154	0.951257	0.986697	0.668951	0.658609	0.883216
3	-1.448442	1.046990	1.034933	2.109777	1.090294	1.134511
4.5	-0.661709	1.131863	1.044541	-0.143029	0.259311	0.622340
6	-0.220612	1.351971	1.121561	-0.660567	0.956307	1.024191
8	0.262882	0.715627	0.893443	-0.596928	1.125889	1.118824
10	0.689622	0.866435	0.961030	0.017014	-12.67771	-5.648806
15	1.735932	0.923811	0.960369	3.949938	0.767580	0.873710
17.6	2.367074	0.914027	0.951492	8.180630	0.738951	0.861306

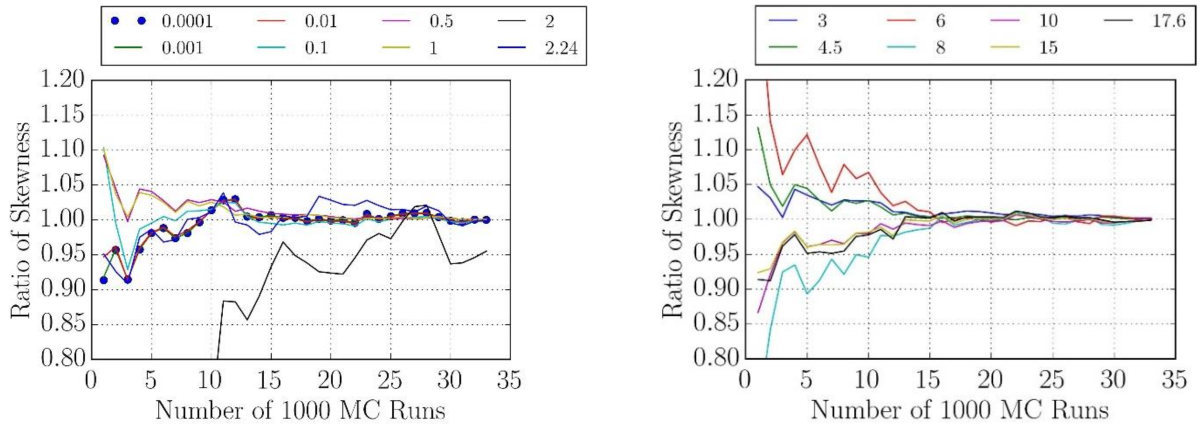

***Fig. 8**: Calculated Skew{y} values are shown for each E as a function of sample numbers K. These results are given as ratios to Skew{y} calculated from K=34,241 samples. The values of E in the legend are given in units of MeV. Skew{y}(K) converges considerably more slowly towards Skew{y}(K=34,241) than MV{y}(K) and Std{y}(K). However, K=1,000/5,000 samples yield marginally acceptable values for Skew{y}. An exception is Skew{y} at E=2 MeV where the skewness is near zero. There, exhibiting the difference in percent is misleading.*

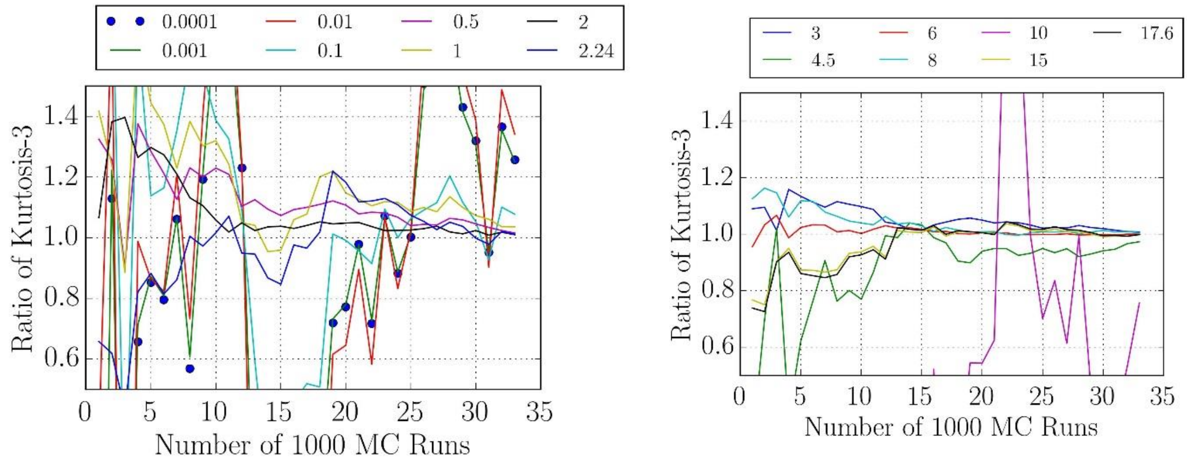

***Fig. 9**: Calculated Kurt{y}–3 values for each E are shown as a function of number of samples K from which they were calculated. They are given as ratios to Kurt{y}–3 calculated from K=34,241 samples. The values of E in the legend are in units of MeV. Kurt{y}–3 (K) converges much more slowly towards Kurt{y}–3 (K=34,241) than MV{y}(K), Std{y}(K), and Skew{y}(K). It does not converge at all for E ={0.0001, 0.001, 0.01, 0.1, 10} MeV since Kurt{y}–3 is close to 0 for those particular energies. For all other E, K=5,000 samples yield marginally acceptable values for Skew{y} and Kurt{y}–3.*

It is obvious from Fig. 9 that Kurt{y}–3(K) has not converged toward Kurt{y}–3 for K=34,241, whenever the values are close to zero, if one chooses to observe the convergence by means of a ratio to Kurt{y}–3 for K=34,241. However, it is also seen in Fig. 10 that the Kurt{y}–3 calculated from K=5,000 or 10,000 samples seem to be close to that calculated from K=34,241 samples. Thus, the results in Table 7 for *E*< 0.5 MeV are misleading. One might be led to believe that those values have converged to marginally acceptable results by considering the ratio to Kurt{y}(*K*=34,241), but this is just an anomaly associated with this particular *E*.

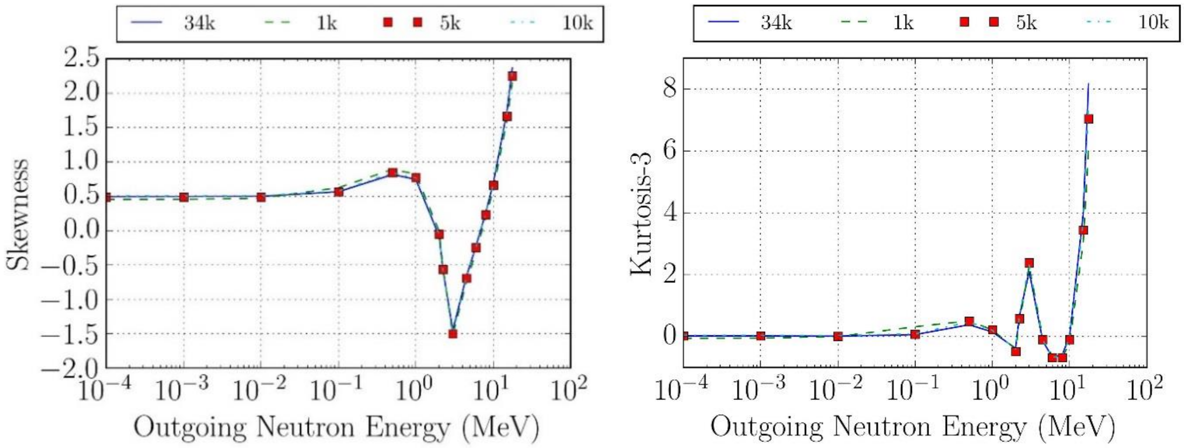

Fig. 10*: Skew{y} and Kurt{y}–3 calculated from K={1,000, 5,000, 10,000 and 34,241} samples are shown as a function of E. K=5,000 and 10,000 samples yield Skew{y} and Kurt{y}–3 values close to those calculated with K=34,241 samples.*

Conclusion: *Stochastic sampling with K=5,000 histories appears to be capable of yielding at least marginally acceptable values of skewness and excess kurtosis for PFNS predicted by the Los Alamos model. In cases where the skewness and excess kurtosis assume values close to 0, the ratio to those values calculated from many more samples is not a good measure of convergence.*

5.2. Use of Lognormal or Gaussian Functions as PDF Emulators

Question: Can the PDFs generated by the particular PFNS model used in this work be well-described by either a Gaussian or a lognormal emulator? If this happens to be true, then UMC-G, involving either a Gaussian or lognormal prior PDF, or GLS in either ordinary or log-space, could be used for PFNS evaluations with this particular model.

This question is studied by using the mean values MV{y} and standard deviation Std{y} calculated from K=34,241 samples of the true PDF as input for Gaussian or lognormal distributions. For the Gaussian, MV{y} and standard deviation Std{y} were taken as given, while the parameters that characterize the lognormal PDFs were derived by considering Eqs. (6) – (9) and using the stochastically generated MV{y} and Std{y} values from the true PDF. The resulting Gaussian and lognormal PDFs approximate the actual model-generated PFNS PDFs reasonably well only for *E*=2 MeV, as shown in Fig. 11. They provide a fair description for model PFNS PDFs with *E*<0.5 keV and *E*=2.24 MeV. The lognormal PDF is marginally better in describing the model PFNS PDFs in these energy ranges. However, both Gaussian and lognormal PDFs fail to describe model PFNS PDFs for *E*={0.5, 1, 3, 4.5, 6, 8, 10, 15, 17.6} MeV. At 15 and 17.6 MeV, the lognormal PDF describes the model PFNS PDFs better, while at 6 and 8 MeV, the Gaussian PDFs are better suited to describe the model PFNS PDFs. The Gaussian and lognormal PDFs determined on the basis of MV{y} and Std{y} values from the model-generated PDFs differ substantially for several *E*, as seen in

Fig. 11. The question therefore arose as to whether GLS evaluated results in PFNS space might differ in a similar manner from those obtained by a GLS evaluation in corresponding *log*(PFNS) space. To investigate this point, covariances and mean values were calculated from K=34,241 samples to form prior covariance matrices and mean values as input for a GLS evaluation. They were updated with experimental ^{239}Pu PFNS described in [12,13], first applying GLS in PFNS space and then applying it in *log*(PFNS) space. The evaluated results in Fig. 12 differ noticeably for most *E*. So, the difference between using a Gaussian or lognormal approximation for Los Alamos model values can influence the evaluated results.

Since the agreement between Gaussian, lognormal, and the PDF of actual model values is unsatisfactory for some *E*, this shortcoming raises the question of whether using GLS in either PFNS or logarithmic PFNS space for evaluations might therefore bias the evaluated results. GLS implicitly assumes that all data entering the algorithm are Gaussian distributed. Hence, evaluating with GLS in PFNS or *log*(PFNS) space leads to a truncation of model predicted PDFs, as can be observed from Fig. 11. The same would be true for using UMC-G with a Gaussian or lognormal emulator.

Conclusion: *The Los Alamos model predicted PFNS are neither well-described by a Gaussian nor by a lognormal emulator. Therefore, it should be studied whether evaluating with GLS in PFNS or log(PFNS) space biases the evaluated results, since model values are assumed to be Gaussian or lognormal in this case. The same is true for UMC-G if a Gaussian or lognormal emulator is used.*

5.3. Choice of a PDF Emulator Based on Skewness and Kurtosis

Question: Can the skewness and excess kurtosis calculated from actual model-generated data be used to quantify how well these data might be described by a chosen analytical emulator PDF?

The skewness and kurtosis of Gaussian distributed data are zero and 3, respectively, with the excess kurtosis being exactly zero. The skewness and kurtosis–3 of lognormal distributed data can be calculated using MV{y} and Std{y} via Eqs. (7), (10) and (11). Skewness and kurtosis can be calculated from sample data for arbitrary PDF by Eqs. (3) and (4). These operations are straightforward to implement and perform for large data sets. The question to consider now is whether the differences of Skew{y} and Kurt{y} for a Gaussian or lognormal PDF from those calculated from actual model-generated sample data that involve an arbitrary PDF are indicators of how well model-generated PDFs might possibly be emulated by a Gaussian or lognormal PDF. This matter also has been investigated in the present work using the K=34,241 PFNS samples for the 15 emitted neutron energies *E* mentioned earlier.

The difference between the skewness of a lognormal PDF and the skewness calculated from actual PFNS data, shown in Fig. 13, is smallest at *E*=2 MeV where the model results are well-described by a lognormal PDF, as seen in Fig. 11. The difference is largest where the model PDFs clearly differ from a lognormal PDF. The difference between Kurt{y}–3 calculated from a lognormal and model data is also larger for *E* where the lognormal PDF does not describe the model value PDF well. However, this effect is not as obvious as when Skew{y} is examined. A similar effect can be observed when comparing the Skew{y} and Kurt{y}–3 values of a Gaussian and the model-generated PDF: Skew{y} in Fig. 10 is close to zero, where the model-generated PDF is nearly Gaussian, and it assumes increasingly larger absolute values when the model-generated PDF deviates considerably from a Gaussian. This tendency is not as pronounced when observing Kurt{y}–3.

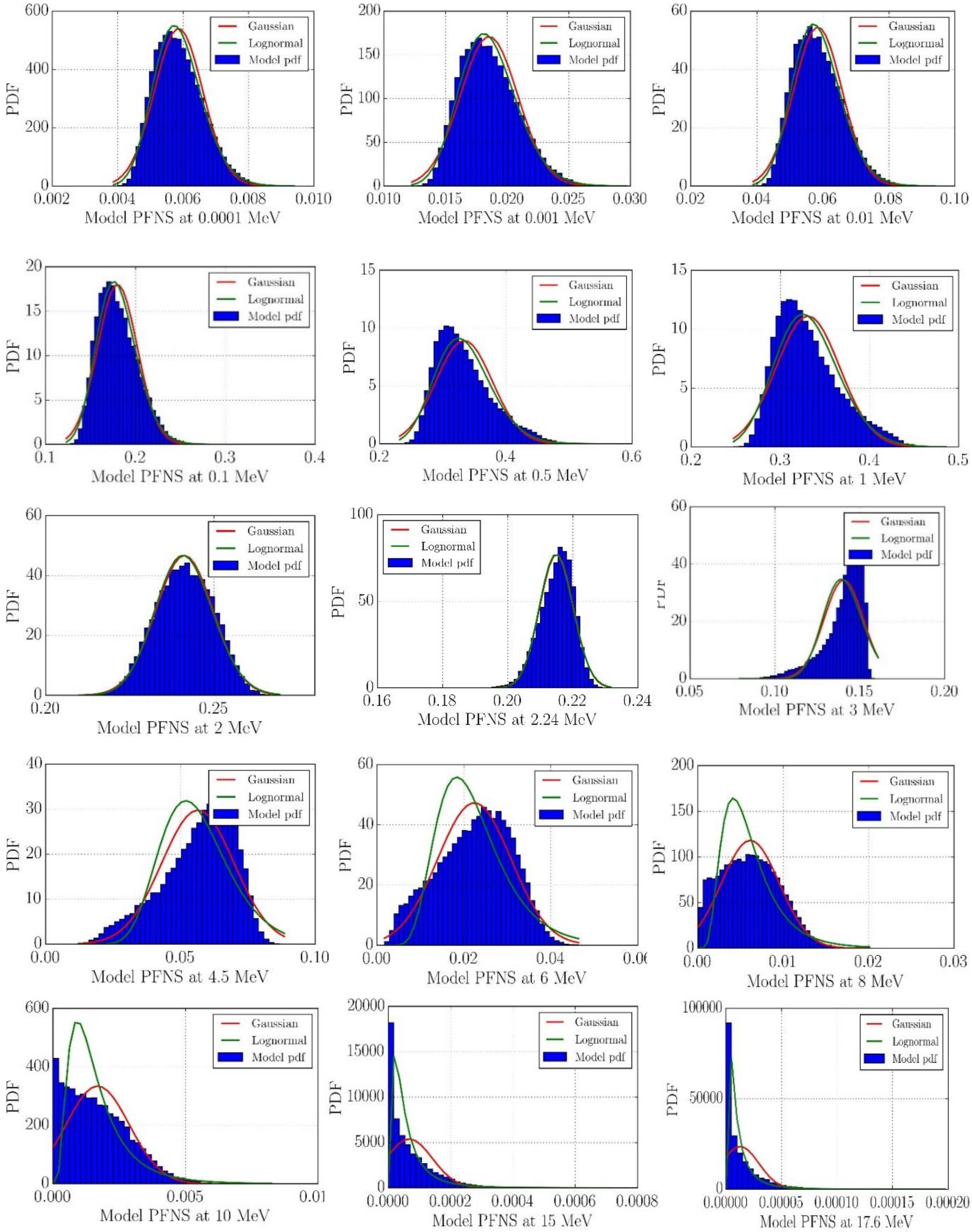

Fig. 11*: The histograms shown here represent sets of K=34,241 stochastically determined PFNS values for each of the indicated outgoing energies E that were calculated using the Los Alamos model. The MV{y} and Std{y} values derived from these sample data were used to generate the shown Gaussian and lognormal PDF shapes to compare with these histograms.*

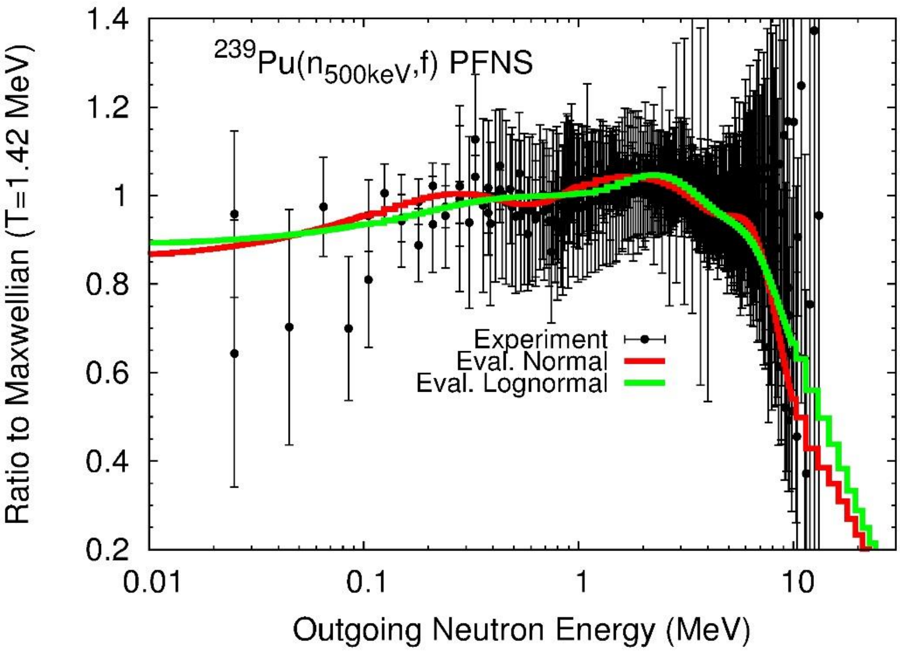

***Fig. 12**: Evaluated results obtained using GLS in PFNS space are compared to those obtained for the same data using GLS in log(PFNS) space. Noticeable differences between these two evaluated PDF shapes can be observed for most of the outgoing neutron energies E.*

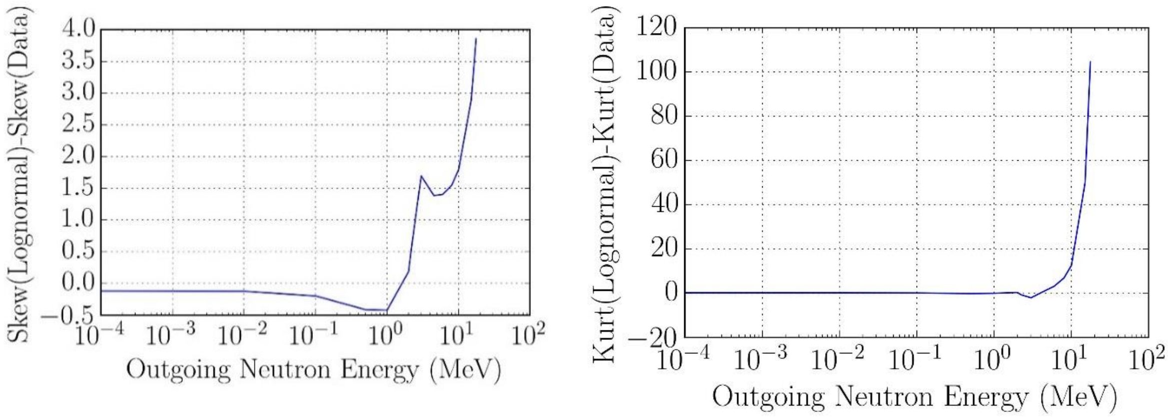

***Fig. 13**: The difference between skewness and kurtosis calculated from a lognormal PDF using MV{y} and Std{y} values obtained from the model-generated PFNS are compared to Skew{y} and Kurt{y} calculated directly from model-generated PFNS.*

Conclusion: *The difference between the skewness calculated from a Los Alamos model-generated PDF and a Gaussian or lognormal PDF based on MV{y} and Std{y} taken from the model-generated values can be used as an indicator of how well these data are described by a Gaussian or lognormal PDF.*

6. Summary

While mean values typically can be estimated with reasonable reliability from PDF sample data collections corresponding to perhaps a few hundred Monte Carlo histories, at least 1,000 sampling histories are needed to obtain acceptable estimates of skewness while no less than

5,000 sampling histories are needed to estimate kurtosis reliably. These sampling numbers apply equally well to lognormal PDFs and arbitrary smooth PDFs, e.g., for those applicable to PFNS data generated using the Los Alamos model. Differences between skewness values, and to a lesser extent kurtosis values, obtained from realistic nuclear model analyses and simple lognormal PDFs having the same mean values and standard deviations, can be used as a way to ascertain whether lognormal or Gaussian distributions might serve effectively as analytical surrogates for actual PDFs. In the case of PFNS data generated by the Los Alamos model, we have determined that it is not possible to do this for all the emitted neutron energies corresponding to 0.5 MeV incident neutron energy.

Acknowledgments

Two of the authors (D.N. and D.L.S) are grateful for assistance from the IAEA Nuclear Data Section in preparing the present manuscript for publication as an IAEA report. One author (D.N.) thanks D. Vaughan for her feedback and interest in this work. The work performed at Los Alamos National Laboratory was partly carried out under the auspices of the National Nuclear Security Agency of the U.S. Department of Energy under Contract No. DE-AC52-06NA25396.

References

1. D.L. Smith, D. Neudecker, and R. Capote Noy, "A Study of UMC in One Dimension", Report INDC(NDS)-0709, International Atomic Energy Agency, Vienna, Austria (2016).
2. D.L. Smith, "A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation", *Proceedings of the Eighth International Topical Meeting on Nuclear Applications and Utilization of Accelerators (AccApp'07)*, Pocatello, Idaho, July 29 – August 2, 2007, American Nuclear Society, LaGrange Park, IL, 736 (2007).
3. R. Capote and D.L. Smith, "An Investigation of the Performance of the Unified Monte Carlo Method of Neutron Cross Section Data Evaluation", *Nuclear Data Sheets* **109**, 2768 (2008).
4. R. Capote and D.L. Smith, "Unified Monte Carlo and Mixed Probability Functions", Proceedings of the *International Conference on Nuclear Data for Science and Technology*, April 26-30, 2010, Jeju, Korea. *Journal of the Korean Physical Society* **59**, 1284 (2011).
5. R. Capote, D.L. Smith, and A. Trkov, "Nuclear Data Evaluation Methodology Including Estimates of Covariances", *EPJ Web of Conferences* **8**, 04001 (2010).
6. R. Capote, D. Smith, A. Trkov, and M. Meghzifene, "A New Formulation of the Unified Monte Carlo Approach (UMC-B) and Cross-Section Evaluation for the Dosimetry Reaction ^{55}Mn(n,γ)^{56}Mn", Fourth International Symposium on Reactor Dosimetry (ISRD-14), Bretton Woods, ASTM STP-1550, 179 (2012). See also: R. Capote, D. Smith, A. Trkov, and M. Meghzifene, *Journal of the ASTM International* **9**(4), JAI 104119 (2012).
7. Georg Schnabel, "Large Scale Bayesian Nuclear Data Evaluation with Consistent Model Defects", Doctoral Thesis, Technische Universität Wien, Vienna, Austria, June 2015.
8. Wikipedia: Internet https://en.wikipedia.org/wiki/Log-normal_distribution, December 2015.

9. MATLAB: The Language of Technical Computing, Mathworks Corporation: Internet https://www.mathworks.com/help/pdf_doc/matlab/matlab_prog.pdf, December 2015.
10. D.G. Madland, and J.R. Nix, "New Calculation of Prompt Fission Neutron Spectra and Average Prompt Neutron Multiplicities", *Nucl. Sci. Eng.* **81**, 213 (1982).
11. D. Neudecker, P. Talou, T. Kawano, D.L. Smith, R. Capote, M.E. Rising, and A.C. Kahler, "Evaluation of the ^{239}Pu Prompt Fission Neutron Spectrum Induced by Neutrons of 500 keV and Associated Covariances", *Nucl. Inst. Methods in Phys. Res. Sect. A* **791**, 80 (2015).
12. S. van der Walt, S.C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation", *Comp. Sci. Eng.* **13**, 22 (2011).
13. D. Neudecker, T.N. Taddeucci, R.C. Haight, H.Y. Lee, M.C. White, and M.E. Rising, "The Need for Precise and Well-documented Experimental Data on Prompt Fission Neutron Spectra from Neutron-induced Fission of ^{239}Pu", *Nucl. Data Sheets* **131**, 289 (2016).

Appendix

The results from K=5,000 and K=1,000 histories, respectively, of stochastic calculations of MV{y}, Std{y}, percent Std{y}, Skew{y}, and Kurt{y}, for selected values of the factor “c” corresponding to nominal standards deviations 3, 5, 7.5, 10, 15, 20, 30, 50, 75, and 100%, are tabulated in detail below. Each stochastic calculation was repeated ten times for every value of “c” and K, resulting in 200 independent sets of calculated statistical data. Averages (AVG), sample standard deviations (STD), and sample percent standard deviations, i.e., STD (%) = STD/AVG of each set of ten repetitions were generated for the accumulated results. These sample-specific values, indicated by the yellow filled cells, are also provided in the following tables. The principal results from analyses of these raw statistical results are summarized in Section 4 of the main body of this paper.

Std{y} = 3% ==> c = 0.3		K = 1,000			
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	1.351588	0.0410211	3.035%	0.0947382	2.9738179	-0.0261821
2	1.3501613	0.0417032	3.089%	0.0190899	2.8695943	-0.1304057
3	1.349273	0.0393399	2.916%	0.0403701	2.8736806	-0.1263194
4	1.351166	0.0426811	3.159%	0.1581693	3.0289641	0.02896413
5	1.3512169	0.0407678	3.017%	0.1911391	3.2224255	0.22242546
6	1.349605	0.0401868	2.978%	0.0866795	3.1666127	0.16661271
7	1.3490654	0.0402477	2.983%	0.0725035	2.9166302	-0.0833698
8	1.3501556	0.0394838	2.924%	0.1593147	3.0789909	0.07899089
9	1.3502714	0.039407	2.918%	0.0577106	2.8894264	-0.1105736
10	1.349244	0.0403779	2.993%	0.1866408	3.0087804	0.00878041
AVG	1.3501747	0.0405216	3.001%	0.1066356	3.0028923	0.00289229
STD	0.0009013	0.0010695	0.078%	0.0624734	0.1234367	0.12343671
STD (%)	0.067%	2.639%	2.603%	58.586%	4.111%	4267.782%

Std{y} = 3% ==> c = 0.3		K = 5,000				
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	1.3512744	0.0403688	2.987%	0.1433266	3.1194561	0.1194561
2	1.3497725	0.0400162	2.965%	0.111585	3.0013153	0.0013153
3	1.3501872	0.0397661	2.945%	0.0900796	3.023051	0.023051
4	1.3505584	0.0404332	2.994%	0.0678465	2.9280661	-0.071934
5	1.3500623	0.0408945	3.029%	0.1279727	3.0770381	0.0770381
6	1.3505655	0.040128	2.971%	0.1119076	2.9842032	-0.015797
7	1.3505409	0.0395525	2.929%	0.0430174	2.8702326	-0.129767
8	1.3503524	0.0403576	2.989%	0.1145861	3.0932773	0.0932773
9	1.3498464	0.0408042	3.023%	0.1053321	3.0734185	0.0734185
10	1.3512744	0.0403688	2.987%	0.1433266	3.1194561	0.1194561
AVG	1.3504435	0.040269	2.982%	0.105898	3.0289514	0.0289514
STD	0.0005212	0.0004193	0.031%	0.0317726	0.0836528	0.0836528
STD (%)	0.039%	1.041%	1.046%	30.003%	2.762%	288.942%

Std{y} = 5% ==> c = 0.5		K = 1,000				
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	1.6480893	0.082542	5.008%	0.2675525	3.4349181	0.43491812
2	1.6538292	0.0829045	5.013%	0.3147773	3.378778	0.37877801
3	1.6539675	0.083238	5.033%	0.1906684	3.1239396	0.12393962
4	1.6530186	0.0822107	4.973%	0.1100428	2.8855383	-0.1144617
5	1.6532087	0.0806976	4.881%	0.1416169	2.9748432	-0.0251568
6	1.6459538	0.0831321	5.051%	0.2002684	3.1771207	0.1771207
7	1.6510281	0.0811137	4.913%	0.1002995	2.8891575	-0.1108425
8	1.6480893	0.082542	5.008%	0.2675525	3.4349181	0.43491812
9	1.6538292	0.0829045	5.013%	0.3147773	3.378778	0.37877801
10	1.6539675	0.083238	5.033%	0.1906684	3.1239396	0.12393962
AVG	1.6514981	0.0824523	4.993%	0.2098224	3.1801931	0.18019312
STD	0.0030277	0.0008851	0.055%	0.0789689	0.2182434	0.21824342
STD (%)	0.183%	1.073%	1.097%	37.636%	6.863%	121.116%

Std{y} = 5% ==> c = 0.5		K = 5,000				
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	1.6524226	0.0823526	4.984%	0.2059987	3.163129	0.163129
2	1.6493503	0.0815461	4.944%	0.1705141	3.0318658	0.0318658
3	1.6501846	0.0810355	4.911%	0.1494525	3.0605644	0.0605644
4	1.6509673	0.0823928	4.991%	0.125419	2.9494109	-0.050589
5	1.6499759	0.0833637	5.052%	0.1903892	3.1220319	0.1220319
6	1.6509691	0.0818063	4.955%	0.1704895	3.0215195	0.0215195
7	1.6495328	0.0831517	5.041%	0.1675893	3.0947126	0.0947126
8	1.6505446	0.0822687	4.984%	0.1767785	3.1390833	0.1390833
9	1.6503822	0.0839938	5.089%	0.1920391	3.1062173	0.1062173
10	1.6511333	0.0823937	4.990%	0.1374967	2.9290841	-0.070916
AVG	1.6505463	0.0824305	4.994%	0.1686167	3.0617619	0.0617619
STD	0.0008921	0.0008811	0.054%	0.0251586	0.0786619	0.0786619
STD (%)	0.054%	1.069%	1.074%	14.921%	2.569%	127.363%

Std{y} = 7.5% ==> c = 0.75		K = 1,000				
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	2.1272525	0.1587401	7.462%	0.1798591	2.9383733	-0.0616267
2	2.1275468	0.1558838	7.327%	0.212934	2.9912991	-0.0087009
3	2.1136896	0.1603592	7.587%	0.2808842	3.2286241	0.22862406
4	2.123364	0.1565067	7.371%	0.1694316	2.9199272	-0.0800728
5	2.123374	0.1561697	7.355%	0.2193226	3.1625258	0.16252579
6	2.11217	0.1594132	7.547%	0.3110323	3.016626	0.01662602
7	2.1281601	0.1538471	7.229%	0.2559239	3.125354	0.12535405
8	2.1212846	0.1582211	7.459%	0.2870129	3.8568143	0.85681432
9	2.1235694	0.1516193	7.140%	0.1334141	2.8813126	-0.1186874
10	2.1240109	0.1640163	7.722%	0.3003079	3.0009147	0.00091466
AVG	2.1224422	0.1574776	7.420%	0.2350123	3.1121771	0.11217711
STD	0.0054854	0.0034904	0.173%	0.0611797	0.2842808	0.28428079
STD (%)	0.258%	2.216%	2.329%	26.033%	9.134%	253.421%

Std{y} = 7.5% ==> c = 0.75		K = 5,000				
Trial No.	**MV{y}**	**Std{y}**	**Std{y} in %**	**Skew{y}**	**Kurt{y}**	**Kurt{y}-3**
1	2.1199266	0.1609091	7.590%	0.2385228	3.0759243	0.0759243
2	2.1223089	0.1603042	7.553%	0.2234667	3.0008433	0.0008433
3	2.1247525	0.1585325	7.461%	0.1791234	3.0926871	0.0926871
4	2.1223904	0.1549256	7.300%	0.2265291	3.2415563	0.2415563
5	2.1211821	0.1621017	7.642%	0.214055	3.0645329	0.0645329
6	2.124209	0.1552005	7.306%	0.2074765	2.8997957	-0.100204
7	2.1240775	0.1613267	7.595%	0.2382953	3.0676338	0.0676338
8	2.126564	0.1605298	7.549%	0.2400181	3.108015	0.108015
9	2.1250594	0.1615861	7.604%	0.2014474	3.0149916	0.0149916
10	2.1255786	0.160605	7.556%	0.2353026	3.157731	0.157731
AVG	2.1236049	0.1596021	7.516%	0.2204237	3.0723711	0.0723711
STD	0.0020877	0.0025737	0.122%	0.0199244	0.0918214	0.0918214
STD (%)	0.098%	1.613%	1.619%	9.039%	2.989%	126.876%

Std{y} = 10% ==> c = 0.998 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	2.7106248	0.2741268	10.113%	0.3610319	3.302819	0.30281898
2	2.7269591	0.2675737	9.812%	0.2382173	2.9657463	-0.0342537
3	2.7269597	0.2671514	9.797%	0.295872	3.2416511	0.24165114
4	2.73503	0.2634827	9.634%	0.3290872	3.2217774	0.22177741
5	2.7270251	0.2589069	9.494%	0.2166692	3.1307147	0.13071468
6	2.7182917	0.2653574	9.762%	0.373204	2.8560201	-0.1439799
7	2.723479	0.2707825	9.943%	0.388373	4.0738749	1.07387488
8	2.7236947	0.2608922	9.579%	0.1072788	2.8193762	-0.1806238
9	2.7283741	0.2809212	10.296%	0.3727206	3.0738049	0.07380493
10	2.7271113	0.2591073	9.501%	0.1991987	2.9071264	-0.0928736
AVG	2.724755	0.2668302	9.793%	0.2881653	3.1592911	0.15929109
STD	0.0065259	0.0069852	0.264%	0.0941086	0.362773	0.36277297
STD (%)	0.240%	2.618%	2.692%	32.658%	11.483%	227.742%

Std{y} = 10% ==> c = 0.998 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	2.7251326	0.2748878	10.087%	0.3054826	3.0457738	0.0457738
2	2.7225944	0.2703527	9.930%	0.2403474	2.933898	-0.066102
3	2.7250856	0.2726131	10.004%	0.2759039	3.0815697	0.0815697
4	2.7257138	0.2814305	10.325%	0.3197127	3.1619192	0.1619192
5	2.726469	0.2729047	10.009%	0.3265394	3.1936148	0.1936148
6	2.7237114	0.2719635	9.985%	0.2958936	3.1997555	0.1997555
7	2.7217305	0.2726499	10.018%	0.2448659	3.0189948	0.0189948
8	2.7290582	0.2696386	9.880%	0.2508665	3.1487314	0.1487314
9	2.7297998	0.2772492	10.156%	0.3198942	3.0219064	0.0219064
10	2.7302061	0.2695656	9.873%	0.3354374	3.2197757	0.2197757
AVG	2.7259501	0.2733256	10.027%	0.2914944	3.1025939	0.1025939
STD	0.0029517	0.0036982	0.136%	0.0359952	0.0958778	0.0958778
STD (%)	0.108%	1.353%	1.356%	12.349%	3.090%	93.454%

Std{y} = 15% ==> c =1.493 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	4.519868	0.6864476	15.187%	0.4465634	3.1838894	0.18388943
2	4.4975353	0.6918212	15.382%	0.3680839	3.1298637	0.12986366
3	4.4782669	0.6510557	14.538%	0.3737417	3.2162018	0.2162018
4	4.4785424	0.6774826	15.127%	0.6300705	4.1012216	1.10122164
5	4.5253651	0.686667	15.174%	0.6640173	4.0279116	1.0279116
6	4.5181065	0.6727732	14.891%	0.3927526	3.1984185	0.19841848
7	4.5184638	0.6614524	14.639%	0.4228208	3.1408846	0.14088463
8	4.5010815	0.6617765	14.703%	0.3764861	3.1031529	0.10315293
9	4.5010198	0.6616273	14.699%	0.4514896	3.4643613	0.46436128
10	4.4551188	0.6742158	15.134%	0.5261993	3.3095379	0.30953789
AVG	4.4993368	0.6725319	14.947%	0.4652225	3.3875443	0.38754433
STD	0.0227495	0.0133812	0.289%	0.1073301	0.372098	0.37209801
STD (%)	0.506%	1.990%	1.934%	23.071%	10.984%	96.014%

Std{y} = 15% ==> c =1.493 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	4.4907522	0.6717264	14.958%	0.5001802	3.554683	0.554683
2	4.5080881	0.6668716	14.793%	0.4519295	3.2897034	0.2897034
3	4.5080692	0.6823116	15.135%	0.4762565	3.4707368	0.4707368
4	4.4984424	0.6756614	15.020%	0.4546689	3.6207815	0.6207815
5	4.4902269	0.6855277	15.267%	0.5074774	3.3052522	0.3052522
6	4.4945176	0.6627089	14.745%	0.3684831	3.2851684	0.2851684
7	4.4899923	0.6680288	14.878%	0.4085037	3.3337945	0.3337945
8	4.5027905	0.6858382	15.231%	0.4674216	3.3096834	0.3096834
9	4.4981182	0.6763266	15.036%	0.4814032	3.3117304	0.3117304
10	4.5228868	0.6916534	15.292%	0.5009103	3.4119459	0.4119459
AVG	4.5003884	0.6766654	15.036%	0.4617234	3.3893479	0.3893479
STD	0.0104234	0.0094869	0.195%	0.0439832	0.1208045	0.1208045
STD (%)	0.232%	1.402%	1.297%	9.526%	3.564%	31.027%

Std{y} = 20% ==> c =1.98 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	7.4273638	1.4319523	19.279%	0.6334346	3.8402257	0.84022575
2	7.3823957	1.3953694	18.901%	0.509766	3.5215191	0.52151907
3	7.3392176	1.4361915	19.569%	0.6191155	3.2262717	0.22627166
4	7.4009859	1.5281208	20.648%	0.6649477	3.525393	0.52539302
5	7.3656709	1.3970108	18.967%	0.3633688	3.0420759	0.04207588
6	7.4131927	1.5209216	20.516%	0.6792135	3.8109271	0.81092709
7	7.3829543	1.3950687	18.896%	0.4594087	3.1408091	0.1408091
8	7.4295562	1.5025819	20.224%	0.5899752	3.3894734	0.38947342
9	7.3819473	1.5091439	20.444%	0.5146394	3.3498173	0.3498173
10	7.3360667	1.4179635	19.329%	0.5176894	3.5076602	0.50766021
AVG	7.3859351	1.4534324	19.677%	0.5551559	3.4354172	0.43541725
STD	0.0328396	0.0554079	0.710%	0.0999615	0.2625298	0.26252984
STD (%)	0.445%	3.812%	3.610%	18.006%	7.642%	60.294%

Std{y} = 20% ==> c =1.98 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	7.4145634	1.483641	20.010%	0.6900323	3.9586052	0.9586052
2	7.3583043	1.4559041	19.786%	0.6200657	3.6598504	0.6598504
3	7.3715671	1.446731	19.626%	0.611854	3.8294718	0.8294718
4	7.3887434	1.470403	19.901%	0.561868	3.4809244	0.4809244
5	7.3740965	1.4941858	20.263%	0.6781038	4.004275	1.004275
6	7.387411	1.4648966	19.830%	0.6239341	3.7412799	0.7412799
7	7.3827301	1.4334253	19.416%	0.5070266	3.2892824	0.2892824
8	7.381173	1.4739707	19.969%	0.6639822	4.0167446	1.0167446
9	7.3829104	1.5070505	20.413%	0.6762754	3.9169497	0.9169497
10	7.3917085	1.4720007	19.914%	0.5702096	3.4869119	0.4869119
AVG	7.3833208	1.4702209	19.913%	0.6203352	3.7384295	0.7384295
STD	0.0147075	0.0217744	0.286%	0.0597067	0.2525986	0.2525986
STD (%)	0.199%	1.481%	1.437%	9.625%	6.757%	34.208%

Std{y} = 30% ==> c =2.94 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	19.939646	6.3752859	31.973%	1.0381947	4.7831159	1.78311595
2	19.874301	6.1143841	30.765%	1.1922069	6.2171055	3.21710545
3	19.579778	5.891928	30.092%	0.9002992	4.0922047	1.09220468
4	19.616932	5.8748984	29.948%	1.1124721	6.5773512	3.57735124
5	19.672496	5.814498	29.556%	1.01736	4.9221745	1.92217454
6	19.680129	5.7145576	29.037%	0.8608355	4.6761557	1.67615567
7	19.578272	5.9397995	30.339%	1.0373775	4.9325495	1.93254954
8	19.759116	6.0281129	30.508%	0.9898464	4.711743	1.71174302
9	20.06859	5.935941	29.578%	0.8369357	4.5477016	1.54770161
10	19.820407	5.7832896	29.178%	0.9861476	5.0348383	2.03483833
AVG	19.758967	5.9472695	30.098%	0.9971676	5.049494	2.049494
STD	0.1647268	0.189976	0.866%	0.109981	0.7616049	0.7616049
STD (%)	0.834%	3.194%	2.876%	11.029%	15.083%	37.161%

Std{y} = 30% ==> c =2.94 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	19.654559	5.9753123	30.402%	0.9289258	4.374954	1.374954
2	19.734522	5.9568302	30.185%	0.9018107	4.3905094	1.3905094
3	19.719894	5.8391661	29.611%	0.910998	4.8361514	1.8361514
4	19.695626	5.7591938	29.241%	0.9690703	4.9526465	1.9526465
5	19.708253	6.0177657	30.534%	0.919616	4.5185946	1.5185946
6	19.762281	5.7512719	29.102%	0.8139803	3.8736878	0.8736878
7	19.806661	6.0256433	30.422%	0.9390404	4.5200591	1.5200591
8	19.88987	6.0166595	30.250%	0.9449759	4.5642096	1.5642096
9	19.842896	6.016459	30.320%	0.8781156	4.226602	1.226602
10	19.854843	6.0129831	30.285%	0.9725026	4.898661	1.898661
AVG	19.766941	5.9371285	30.035%	0.9179036	4.5156075	1.5156075
STD	0.0779111	0.1107329	0.519%	0.0466583	0.3289529	0.3289529
STD (%)	0.394%	1.865%	1.729%	5.083%	7.285%	21.704%

Std{y} = 50% ==> c =4.73 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	128.19739	63.610404	49.619%	1.2534523	5.0643741	2.0643741
2	125.01267	65.542759	52.429%	2.1744565	11.845131	8.84513125
3	129.28032	68.143627	52.710%	2.0927708	10.551724	7.55172438
4	129.25892	65.783062	50.892%	1.6590462	7.5753143	4.57531425
5	128.11106	62.892813	49.092%	1.4893225	6.9835908	3.98359075
6	127.8436	61.785946	48.329%	1.366364	5.6177151	2.61771506
7	123.56422	63.384198	51.297%	1.7139388	7.9681349	4.96813493
8	126.34706	60.942657	48.234%	1.3528732	5.715837	2.715837
9	126.37617	62.026378	49.081%	1.6486023	8.0041938	5.00419376
10	122.90358	62.524039	50.872%	1.5843237	6.7714445	3.77144449
AVG	126.6895	63.663588	50.256%	1.633515	7.609746	4.609746
STD	2.262458	2.202958	1.618%	0.3034376	2.1604464	2.16044642
STD (%)	1.786%	3.460%	3.219%	18.576%	28.391%	46.867%

Std{y} = 50% ==> c =4.73 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	126.60767	63.691324	50.306%	1.5040727	6.5176508	3.5176508
2	125.64279	61.29765	48.787%	1.3474162	5.7391859	2.7391859
3	126.38782	62.834714	49.716%	1.5307027	7.2390162	4.2390162
4	127.28863	66.044823	51.886%	1.6491614	7.6070838	4.6070838
5	126.75713	63.801122	50.333%	1.6434839	7.7800343	4.7800343
6	126.06889	63.056687	50.018%	1.6695943	8.3903017	5.3903017
7	125.64993	62.129749	49.447%	1.478906	7.1410412	4.1410412
8	126.97368	62.184274	48.974%	1.5210992	7.2004275	4.2004275
9	127.80657	64.817596	50.715%	1.4960373	6.5100834	3.5100834
10	127.29253	63.332349	49.753%	1.69061	8.5242277	5.5242277
AVG	126.64756	63.319029	49.994%	1.5531083	7.2649052	4.2649052
STD	0.7224472	1.3799461	0.897%	0.107947	0.8635192	0.8635192
STD (%)	0.570%	2.179%	1.795%	6.950%	11.886%	20.247%

Std{y} = 75% ==> c =6.663 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	962.8548	782.54214	81.273%	3.5274189	24.423087	21.423087
2	1007.0591	767.32897	76.195%	2.5541937	14.112229	11.1122286
3	990.92426	719.39062	72.598%	2.3728039	12.98837	9.98836992
4	986.35393	699.88465	70.957%	2.0124052	9.0381056	6.03810557
5	945.89003	729.15741	77.087%	2.6548304	14.948498	11.948498
6	970.05927	685.96978	70.714%	2.0313019	9.1425714	6.14257137
7	971.52106	713.29648	73.421%	2.6286403	15.326426	12.3264264
8	938.11785	711.00878	75.791%	2.3528144	11.260841	8.26084063
9	959.36331	661.49547	68.952%	2.2582719	11.808906	8.80890582
10	950.77809	682.23297	71.755%	2.0520474	9.7334735	6.7334735
AVG	968.29217	715.23073	73.874%	2.4444728	13.278251	10.2782507
STD	21.559331	37.336178	3.700%	0.4500096	4.5442114	4.54421137
STD (%)	2.227%	5.220%	5.008%	18.409%	34.223%	44.212%

Std{y} = 75% ==> c =6.663 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	968.05917	736.16867	76.046%	2.9891913	21.612675	18.612675
2	980.1784	717.13483	73.164%	2.4695347	14.267431	11.267431
3	970.83557	711.30341	73.267%	2.2121486	11.06304	8.0630401
4	988.21314	756.32398	76.534%	3.1023243	25.516141	22.516141
5	984.83064	744.1219	75.558%	2.3878537	12.207976	9.2079759
6	988.74488	747.47836	75.599%	2.6862113	16.411884	13.411884
7	974.14259	743.64236	76.338%	2.434922	12.545907	9.545907
8	964.67031	692.85111	71.823%	2.43531	14.300603	11.300603
9	963.83555	705.34272	73.181%	2.479303	14.557674	11.557674
10	976.27155	729.83607	74.757%	2.4088461	12.889167	9.8891671
AVG	975.97818	728.42034	74.627%	2.5605645	15.53725	12.53725
STD	9.2892711	20.817219	1.645%	0.2815863	4.5755201	4.5755201
STD (%)	0.952%	2.858%	2.204%	10.997%	29.449%	36.495%

Std{y} = 100% ==> c =8.31 K = 1,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	5741.8897	7031.058	122.452%	7.7791095	100.56768	97.5676812
2	5493.0063	4701.1262	85.584%	2.6125634	16.029843	13.0298426
3	5893.4922	5994.7636	101.718%	2.9655132	15.412426	12.4124257
4	5920.8297	6191.8607	104.578%	2.3474797	10.830469	7.83046904
5	5910.4446	5744.0083	97.184%	2.7776196	13.993103	10.9931026
6	5770.1745	5608.2102	97.193%	2.9418094	16.455888	13.455888
7	5473.6269	5164.0047	94.343%	3.9969941	35.28534	32.2853404
8	6069.767	6593.1553	108.623%	3.9512291	30.882989	27.8829887
9	5924.4862	6716.7693	113.373%	5.8663744	63.598602	60.5986021
10	5601.9989	5433.2886	96.988%	2.8545174	15.185666	12.1856659
AVG	5779.9716	5917.8245	102.204%	3.809321	31.824201	28.8242006
STD	201.11735	732.13053	10.527%	1.733843	28.941825	28.9418245
STD (%)	3.480%	12.372%	10.300%	45.516%	90.943%	100.408%

Std{y} = 100% ==> c =8.31 K = 5,000

Trial No.	MV{y}	Std{y}	Std{y} in %	Skew{y}	Kurt{y}	Kurt{y}-3
1	5975.222	6181.6179	103.454%	4.1095119	36.826379	33.826379
2	5821.8973	5889.6105	101.163%	4.1108855	36.158635	33.158635
3	5486.5815	5071.27	92.430%	2.9316203	16.965623	13.965623
4	5831.5489	5748.2246	98.571%	3.2281217	20.189508	17.189508
5	5856.8835	5890.8084	100.579%	3.8061767	31.533434	28.533434
6	5783.4758	5644.6314	97.599%	3.3787489	24.795828	21.795828
7	5685.0802	5890.3669	103.611%	5.736654	90.232269	87.232269
8	5782.2944	5504.7089	95.199%	2.9780034	17.596917	14.596917
9	5684.2084	5829.2856	102.552%	6.319802	108.96409	105.96409
10	5712.8497	5516.6003	96.565%	3.2863538	22.614861	19.614861
AVG	5762.0042	5716.7124	99.172%	3.9885878	40.587754	37.587754
STD	130.81098	303.12151	3.746%	1.1611754	32.198565	32.198565
STD (%)	2.270%	5.302%	3.777%	29.112%	79.331%	85.662%

Nuclear Data Section	E-mail: nds.contact-point@iaea.org
International Atomic Energy Agency	Fax: (43-1) 26007
Vienna International Centre, P.O. Box 100	Telephone: (43-1) 2600 21725
A-1400 Vienna, Austria	Web: http://www-nds.iaea.org