LA-UR-10-08203

*Approved for public release;
distribution is unlimited.*

Title: Statistical Assessment of Numerous Monte Carlo Tallies

Author(s): Brian C. Kiedrowski
Clell J. (CJ) Solomon

Intended for: 2011 ANS Math & Computation Topical Mtg.
May 8-11, 2011
Rio de Janeiro, RJ, Bazil

International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2011)
Rio de Janeiro, RJ, Brazil, May 8-12, 2011, on CD-ROM, Latin American Section (LAS) / American Nuclear Society (ANS)
ISBN 978-85-63688-00-2

STATISTICAL ASSESSMENT OF NUMEROUS MONTE CARLO TALLIES

Brian C. Kiedrowski* **and C.J. Solomon**
Los Alamos National Laboratory
P.O. Box 1663, MS A143 Los Alamos, NM 87545
bckiedro@lanl.gov; csolomon@lanl.gov

ABSTRACT

Four tests are developed to assess the statistical reliability of collections of tallies that number in thousands or greater. To this end, the relative variance density function is developed and its moments are studied using simplified, non-transport models. The statistical tests are performed upon the results of MCNP calculations of three different transport test problems and appear to show that the tests are appropriate indicators of global statistical quality.

Key Words: Statistical tests, mesh tallies, variance reduction.

1. INTRODUCTION

The Monte Carlo method is being used increasingly often to collect high-resolution data about radiation flux or dose distributions as functions of space and energy. The resolutions are such that millions of tallies and associated uncertainties are being collected in a single simulation. Metrics for statistical assessment have been developed and are successfully applied to small numbers of tally scores [1]; however, these same tests do not scale well given that the available computational memory resources are often already expended upon the tally data itself. To this end, new metrics need to be developed to assess the quality of the statistics of an entire distribution of tally results rather than the individuals separately with the primary requirement of limiting the memory footprint.

For statistical assessment, the relative variance density function is defined and certain useful properties are obtained. As proof of concept, simplified statistical (non-radiation transport) models are developed and the tests are performed upon the results. The statistical tests are implemented in a research version of MCNP [2] (prerelease version of MCNP6) and tested on three non-trivial, continuous-energy problems: fusion neutrons in a block of water, a three-legged duct benchmark [3, 4], and a modified k-effective of the World problem [5, 6]. The latter of these two problems require global weight-window maps [7] that are generated by an iterative linear tally combination (LTC) technique [8] coupled with the MCNP weight-window generator [9] (WWG).

*Corresponding author.

2. THE RELATIVE VARIANCE DENSITY FUNCTION

2.1 Definition

A metric that is commonly used for assessing calculational efficiency is analyzing a cumulative distribution of relative uncertainties (ratio of a tally's sample standard deviation to sample mean) corresponding to the fraction of mesh elements with a relative uncertainty less than some value [10]. Often curves of the cumulative distribution are plotted for different variance reduction parameters given a fixed computational time. By taking values of the cumulative distribution (typically at ten percent), the efficiency of each calculation can be assessed.

A related distribution can be employed for statistical assessment as well. First define v as the relative variance or the square of the relative uncertainty (range of zero to one). The rationale for using the relative variance as opposed to the relative uncertainty relates to convenient additive properties of variances. There is a density function $f(v;N)$ corresponding to the frequency of mesh elements having a particular relative variance v for a given number of histories N.

The relative variance is only defined when the sample mean exists and is non-zero. Since most Monte Carlo tally scores in radiation transport are either strictly positive or zero, this reduces to having a non-zero number of positive scores. For this reason, $f(v;N)$ only incorporates tallies that have positive scores, neglecting those that do not. Note that there are valid reasons to expect the number of scores to be zero such as in a time-dependent calculation for tally elements that are beyond a radiation front limited by the speed of light. For quantification, the factor ζ is defined as the faction of mesh elements having at least one positive score.

An assumption inherent in the use of a density function and moments is that the number of elements be sufficiently large to merit a statistical treatment. Should the number of mesh elements be few, then performing the statistical checks for individual tallies is tractable and preferred. While the exact number depends upon the problem, guidance is that it should surely be more than a hundred elements, which is typical for distributional tallies.

2.2 Moments and Related Quantities

The moments of $f(v;N)$ are meaningful in the limit of large N. Define V as a random variable taking on some relative variance v. The rth moment of $f(v;N)$ is

$$\mathbb{E}\left(V^r\right) = \int_0^1 dv\; v^r f(v;N). \tag{1}$$

Since $f(v;N)$ is usually unknown, it is necessary to approximate the true rth moment with the sample rth moment computed by

$$\overline{v^r} = \frac{1}{M}\sum_{m=1}^{M} v_m^r(N) \tag{2}$$

instead, where M is the number of elements in the mesh each with an index m with a relative variance v_m for a given number of N trials or histories. The density mean is the first moment, and the density variance (and standard deviation) can be derived from the second moment. Higher order moments can be used to derive the density skewness and density kurtosis as well.

In the limit of large N the rth moment of $f(v;N)$ has the property of converging to zero as $1/N^r$, so long as the underlying sampling functions meet certain conditions. This is a direct consequence of the central limit theorem that asserts that if the mean and variance of a distribution exist and the samples are independent, identically distributed, and sufficient in number, then the sample mean is approximately distributed about the true mean as a normal distribution with a standard deviation converging as $1/\sqrt{N}$.

Each individual mesh element has an underlying scoring density function that is unknown. Should every scoring density function satisfy the central limit theorem, the individual relative variances should all converge as $1/N$, and, therefore, their sum should converge as $1/N$. This result is easily extended to the rth moment as converging as $1/N^r$. Except for special cases that are incredibly unlikely to occur in realistic problems (specifically, where the scoring probabilities are identical for all elements), the variance of $f(v;N)$ converges as $1/N^2$.

2.3 Numerics of Simplified Models

The actual scoring density functions encountered in a radiation transport simulation are going to be complicated and unknown. For the sake of convenience, simplified non-transport models are used to study the relative variance density function.

2.3.1 Binomial mixture

A model allowing for analytical expressions is a mixture of binomial distributions. Suppose there are M elements in the mesh and that the mesh may be decomposed into a finite number of domains J such that elements in the jth domain have a probability of scoring p_j and there are $w_j M$ elements within this domain. The probability of having k scores is distributed as binomial:

$$s_j(k;N) = \binom{N}{k} p_j^k \left(1-p_j\right)^{N-k}. \tag{3}$$

If all scores are identical, it is possible to derive a form for the relative variance. Suppose X represents a random variable for the cumulative score at history N. The relative variance of the mean for k scores is

$$\begin{aligned} V(k;N) &= \frac{\mathrm{Var}\left(\mathbb{E}\left(X\right)\right)}{\mathbb{E}\left(X\right)^2} = \frac{1}{N}\left[\frac{\mathrm{Var}\left(X\right)}{\mathbb{E}\left(X\right)^2}\right] = \frac{1}{N}\left[\frac{\mathbb{E}\left(X^2\right)}{\mathbb{E}\left(X\right)^2} - 1\right] = \frac{1}{N}\left[\frac{N}{k} - 1\right] \\ &= \frac{1}{k} - \frac{1}{N}. \end{aligned} \tag{4}$$

The relative variance is also distributed as binomial, but slightly modified via renormalization to account for the fact that it is undefined for $k=0$,

$$f_j(v(k);N) = \frac{1}{1-\left(1-p_j\right)^N} \binom{N}{k} p_j^k \left(1-p_j\right)^{N-k}. \tag{5}$$

The density functions for two hypothetical regions with sampling probabilities of 0.05 and 0.10 respectively, given $N = 100$, are plotted in Fig. 1.a. Note that the connecting lines are merely for illustrative purposes; the density function is only defined at the data points. Intuitively, the peak of the distribution has a lower center and is thinner for a higher sampling probability. This is extended to a mixture of binomial distributions to account for different domains having

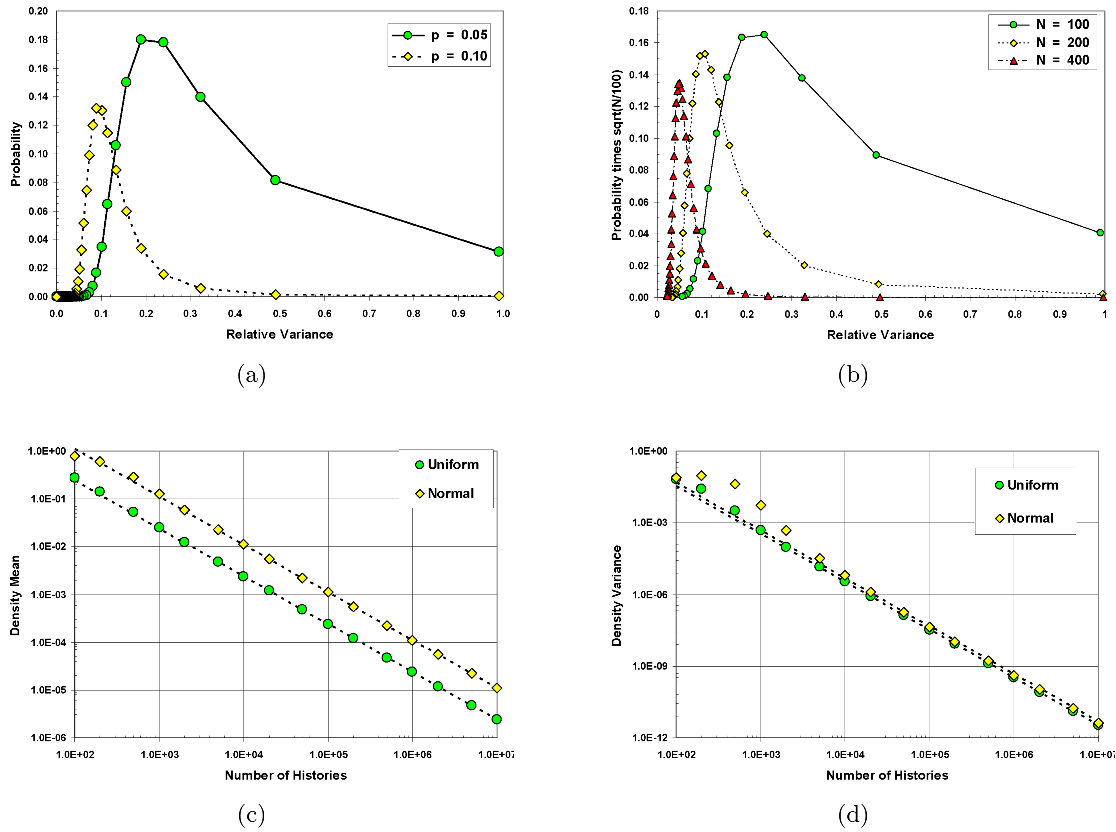


Figure 1: Binomial model of relative variance density. (a) Varied p for fixed N, (b) Varied N for fixed distribution of p, (c) mean of relative density function for uniform and normal sampling models, and (d) variance of relative density function for uniform and normal sampling models.

different sampling probabilities. Scores are still the same between domains for simplicity. The density function for this overall system is simply

$$f(v(k); N) = \sum_{j=1}^{J} w_j f_j, \tag{6}$$

where the w_j is a weighting factor corresponding to the fraction of elements within the jth domain. As an example, the w_j are linked with the sampling probabilities that are distributed normally with mean 0.05 and standard deviation 0.01 (the distribution is truncated at $p = 0.001$ and $p = 0.10$ with minimal impact). The density functions for $N = 100$, 200, and 400 are plotted in Fig. 1.b. As expected, increasing N shifts and thins the peak. Insightful is the fact the distributions are positive skewed with the infrequently sampled regions producing a long-thin tail (i.e., the distribution is leptokurtic). This behavior is observed empirically for realistic radiation transport problems.

The mean and variance of $f(v(k);N)$ can also be computed to determine if the distribution converges at respective rates of $1/N$ and $1/N^2$ for large N. Two cases are considered and results are computed numerically for N ranging from 100 to 10 million. The first involves 1000 regions where (randomly chosen) p_j are uniformly distributed from 0.01 to 0.10, and the second are normally distributed with mean 0.01 with standard deviation 0.002. The mean and variance of these two cases are plotted in Figs. 1.c and 1.d respectively. Linear regression using a least squares fit for $1/N$ and $1/N^2$ curves (denoted by $h(x)$) are plotted. To assess the goodness of fit, the coefficient of determination (R^2 value),

$$R^2 = 1 - \left(\sum_{i=1}^{N} (x_i - h_i)^2 \right) \Big/ \left(\sum_{i=1}^{N} (x_i - \bar{x})^2 \right), \tag{7}$$

is determined. In this case, both match having coefficients of determination approximately unity for $N > 100$ thousand. This is evidence that, for these simplified models, that the mean and variance converge at their expected asymptotic rates.

2.3.2 Variation in scoring

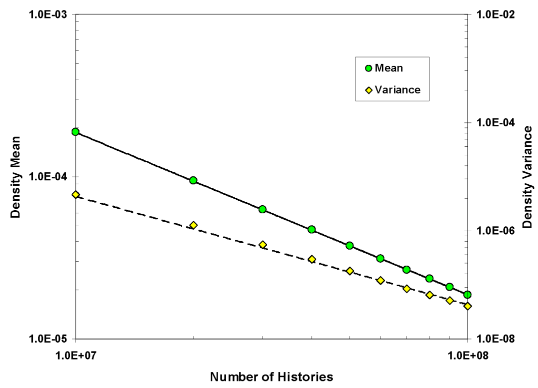
Unfortunately, the logical extension of allowing for variation in the scoring density function is not amenable to an analytic calculation of $v(k)$. Numerical simulation via Monte Carlo techniques is therefore required. The model is similar: a mesh has M elements, each having a probability of scoring p_m that are distributed with some prescribed distribution. Should the mesh have a score in a history, there is an underlying scoring density function that is sampled – for simplicity, an exponential distribution with a fixed parameter with mean score of unity. The scoring probabilities are distributed as two distributions: uniform and log-normal.

For the uniform case, the scoring probabilities are sampled with a range from 1 x 10^{-4} to 0.01. The mesh has 100,000 elements and 10 million histories are run. The mean and variance of the relative variance distribution are computed periodically throughout the simulation. A plot of these values can be found in Fig. 2.a. Linear regression is performed upon the last half of the histories for the mean and variance curves on a log-log scale. The regression should match lines with slopes of -1 and -2 respectively. The corresponding R^2 values are 0.9999 and 0.9985.

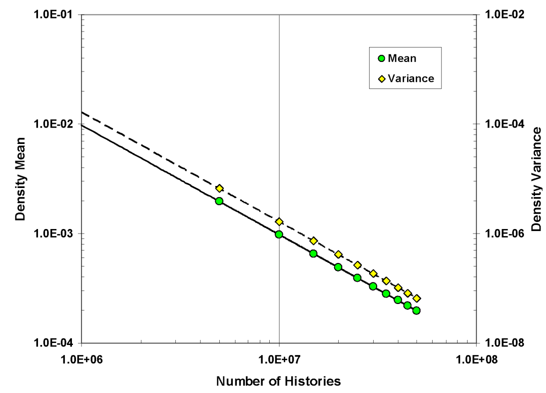
The log-normal case samples the scoring probabilities from the exponentiation of a normal distribution with mean -8 and standard deviation 1. The mesh has 100,000 elements and 50 million histories are run. Fig. 2.b displays periodically computed means and variances of the relative variance distribution. The R^2 values of the linear regression are 0.9999 and 0.9935 for the respective mean and variance curves.

2.3.3 Rare event simulation

Statistical problems often arise in a Monte Carlo simulation because of rare events contributing an unusually large score to a tally. For individual tally results, the variance of the variance (VOV) is particularly useful for this. The analogous quantity is the variance of the relative variance density (or density variance).



(a)



(b)

Figure 2: Simplified model with variable scoring probabilities with: (a) a uniform density, and (b) a log-normal density.

The question is whether or not the distribution variance is able to pick out a poorly behaved mesh element. To assess this, a simulation is run with 100 thousand mesh elements with some defined scoring probability distribution. In one of these elements, there exists a small probability of sampling a rare event where the usually tally score is multiplied by some large factor. In this case, only one element of a large mesh will exhibit rare, large scores.

Two cases considered are a simple model where all the scoring probabilities are identical (0.01) and one where the scoring probabilities are distributed as log-normal (exponentiation of a normal distribution with mean -5 and standard deviation 0.5). The scoring density itself is distributed as exponential with mean score of unity for both. For the former, the assigned rare scoring element has a rare event probability of one part in 100 thousand where the score is multiplied by one thousand. The latter case takes the least likely element to score; given that it scores, with probability 0.01, the score is multiplied by one hundred.

To easily visualize that the relative variance density can spot misbehavior in a single element, the tally standard deviation of the misbehaved element and the standard deviation of the density function are obtained at specific intervals within the simulation. The ratio of the standard deviation at an interval to the one computed in the previous is taken representing the amount of change of the quantity from one interval to the next. Should there be a large fluctuation in either the standard deviations of the tally or density function, the ratio should show a spike. For a large number of histories, these fluctuations will be mostly from sampling the rare event. Figs. 3.a and 3.b for the respective uniform and log-normal cases illustrate this. Qualitatively, the spikes appear to mostly follow each other.

This effect can be quantified by computing the correlation coefficient between these ratios for intervals greater than 10 million histories. The uniform density has a correlation coefficient of 0.811 and the log-normal has a correlation coefficient of 0.758. This indicates that there is a fairly strong positive correlation between the interval ratios of the misbehaved tally standard deviation and the density standard deviation implying that the the fluctuations in the density function can detect fluctuations in a single misbehaved tally.

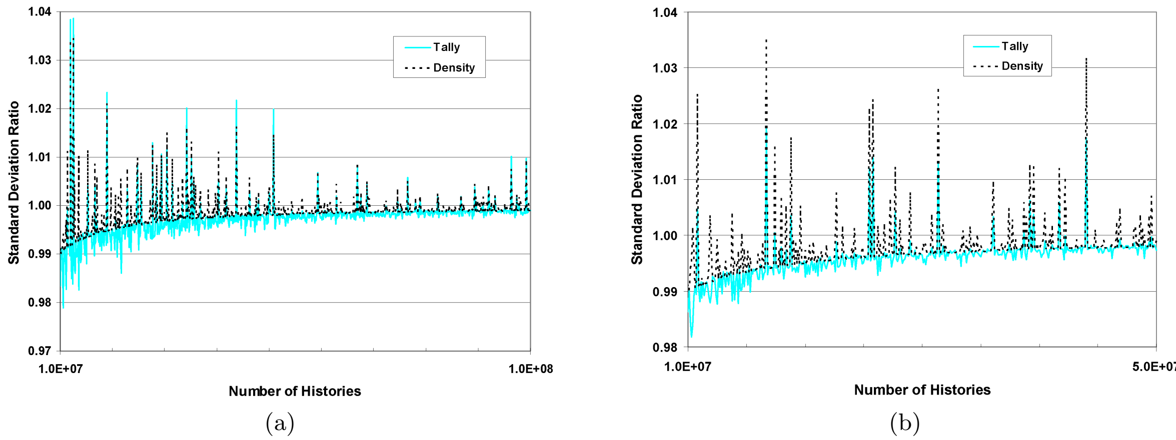


Figure 3: Ratio of standard deviations between history intervals for: (a) a uniform density, and (b) a log-normal density.

3. STATISTICAL ASSESSMENT

3.1 Development of Tests

Using the density function and related quantities, a set of statistical tests can be developed to provide guidance to an analyst as to the statistical pedigree of the results within a distribution. These tests are meant to be low cost in terms of memory and computational time. As such, they can only capture broad behavior of the distribution, but as previously seen, moments of the density function (particularly the variance) are particularly sensitive to even one misbehaved element.

All of the calculations are performed in a research version of MCNP6. For the sake of simplicity, it is assumed that the user cares equally about all elements in the mesh. Often the mesh overlaid is chosen for convenience, and this is not a valid assumption. In this case, a mechanism can be implemented for the user to assert the degree of relevance of particular regions of the problem that act as weighting functions when calculating various quantities. While this extension is indeed possible and fairly straightforward, it is not employed here.

It is worth stating that the conditions of a test are necessary but not sufficient to prove statistical robustness. No test can definitively answer the question of whether or not a region has been sufficiently sampled should not enough data be available to make the assessment. However, tests are still valuable in providing the user with some confidence in this respect.

An easy test to perform is to count the fraction of elements with non-zero tally ζ. Ideally, this fraction should converge to unity; however, for aforementioned reasons, this need not be the case. A reasonable assertion for its behavior is that this value should be constant for much of the problem; should new regions be added, this could invalidate the conclusions of the statistical tests of the density function because it is undefined for zero score elements. Therefore, a reasonable requirement is that ζ be constant for the last half of the histories. While this does not ensure that the phase space has been appropriately sampled, it does give clues as to such.

Two other tests can be derived from moments of the density function. The first is that the mean of the relative variance density be decreasing monotonically as $1/N$ in the last half of the histories. The second is that the variance of the relative variance density decrease monotonically as $1/N^2$ in the last half of the histories. To assess both of these, a linear regression is performed using a least squares fitting scheme. The functions A/N and B/N^2 are fit using the data collected in the last half of the histories using iterative techniques such as Newton-Raphson. Next, the R^2 value is obtained for the fit. Empirically, it seems that $R^2_{\bar{v}} \geq 0.99$ and $R^2_{\sigma^2_v} \geq 0.95$ for the mean and variance tests seem reasonable to define as passing. The reason for the difference being related to the higher statistical fluctuation in the variance since it is related to the fourth moment of the scoring densities.

Another often used metric counts the fraction of elements with a relative uncertainty less than some prescribed value. An accepted threshold that is common is ten percent, which is based upon tests of statistical reliability for distributions that are statistically problematic [11]. Unfortunately, demanding that every element have a relative uncertainty of less than ten percent (i.e., the infinity norm) is very stringent, and it may be acceptable for a some elements to be less precise depending upon the requirements of the application.

Rather, for the purposes of automation, it may be beneficial to think of which elements being important in a statistical way (again, if the user cares about specific elements, then those need to be checked manually). Suppose an element is randomly selected (no sampling is actually done; this is only for the sake of argument) from the mesh with relative variances distributed as $f(v;N)$. Observed properties of $f(v;N)$ are that they are (for large N) positively skewed and leptokurtic, implying that even for a well converged peak, there exists a risk (sometimes termed kurtosis risk) of selecting an element significantly worse than the mean of $f(v;N)$. Such a selection, should it occur by chance, is considered to be far worse than one with slightly higher than average relative variance (i.e., elements far on the tail are considered the most problematic). Of course, if $\bar{v}$ is large, then the risk is dominated by the peak itself and this must be taken into account with this statistical selection treatment.

To handle the issue of the tail, consider the index of dispersion for the density variance

$$D_{\sigma^2_v} = \frac{\mathrm{Var}\left(\sigma^2_v\right)}{\sigma^2_v}. \tag{8}$$

The variance of σ^2_v is

$$\mathrm{Var}\left(\sigma^2_v\right) = \frac{1}{n}\left(\mu_4 - \sigma^4_v\right) + O(n^{-2}), \tag{9}$$

where κ_v is the density kurtosis and n is the number of elements selected in this hypothetical sampling and is considered to be a significant fraction such that a statistical treatment is valid (not to be confused with the number of Monte Carlo histories N). The dispersion coefficient is approximately

$$D_{\sigma^2_v} \approx \frac{\sigma^2_v}{n}\left(\kappa_v - 1\right). \tag{10}$$

Now define a statistical risk parameter

$$\Theta = \bar{v} + nD_{\sigma^2_v} = \bar{v} + \sigma^2_v\left(\kappa_v - 1\right). \tag{11}$$

The first term considers the peak. Suppose the tallies somehow had no dispersion in their variances, then of relevance is the location of the peak or $\bar{v}$. The second term accounts for the added risk of sampling an element far on the tail that is poorly converged. The factor of n is present to scale it by the number of elements hypothetically sampled and exists to eliminate that term so no sampling of $f(v; N)$ is actually done. The first term goes asymptotically as $1/N$ and the second $1/N^2$, so that the second term vanishes for large N leaving $\Theta = \bar{v}$ when the dispersion of the density function becomes small. Therefore, when the impact of the density tail is deemed insignificant, $\Theta \leq 0.01$ corresponds to an average relative uncertainty of ten percent.

The four statistical tests are summarized as follows:

1. The fraction of mesh elements with scores ζ must be constant for the second half of the problem.
2. The quantity $\Theta \leq 0.01$ at the end of the simulation.
3. The density mean must monotonically decrease as $1/N$ (an $R^2 \geq 0.99$) in the second half of the problem.
4. The density variance must monotonically decrease as $1/N^2$ (an $R^2 \geq 0.95$) in the second half of the problem.

3.2 Verification of Statistical Checks

To test whether the checks apply to radiation transport problems, a simple, non-trivial problem is used. The problem is a 20 cubic cm block of water (1.0 g/cc density) with a centrally located monoenergetic 14.1 MeV neutron source. A 50 x 50 x 50 element mesh tally with equal sized elements is overlaid on the problem and the neutron flux is estimated. The problem is run for varied N and the statistical checks are computed. The results of these tests are displayed in Table I (ζ is given at the halfway through the calculation and should be 1.0 to pass) where passing is indicated by italicization.

All tests are satisfied sometime between 500 and 750 thousand histories, and the tests continue to pass for a larger number of histories. 200 thousand histories are required before all elements are sampled by halfway through the problem. Also, at this time, the relative variances go as $1/N$ with a fit of $R^2 > 0.99$ suggesting the means of the density function satisfy the central limit theorem. Twice as many histories are required before the density variance goes as $1/N^2$ indicating it takes this long to believe the uncertainties are also well behaved. The parameter Θ appears the most difficult to converge; for this problem, at 750 thousand histories, just over 80 percent of the mesh elements have a relative uncertainty less than 10 percent; however, Θ is within about three percent of $\bar{v}$ indicating a low dispersion.

Further verification is performed with more difficult problems. The two problems selected require global importance functions to achieve convergence in a reasonable amount of time. There are techniques to generate these global importance maps involving either deterministic methods to generate weight windows from forward-weighted adjoint solutions or iterative Monte Carlo sampling via a linear tally combination with the weight-window generator (LTC-WWG). For these comparisons, the iterative technique is used. For a fixed N that is sufficiently large, the statistical tests should perform better with a better importance map. This is a check on the

Table I: Statistical test quantities for the water block problem.

N	ζ	Θ	$R^2_{\bar{v}}$	$R^2_{\sigma^2_v}$
50 K	0.9904	0.3659	*0.9971*	0.8682
100 K	0.9992	0.1416	0.9666	0.9007
150 K	0.9999	0.1146	0.9541	0.8693
200 K	*1.0000*	0.0805	*0.9960*	0.9091
250 K	*1.0000*	0.0311	*0.9964*	0.9085
300 K	*1.0000*	0.0191	*0.9977*	0.9273
400 K	*1.0000*	0.0135	*0.9985*	*0.9576*
500 K	*1.0000*	0.0106	*0.9990*	*0.9646*
750 K	*1.0000*	*0.0069*	*0.9997*	*0.9909*
1 M	*1.0000*	*0.0051*	*0.9998*	*0.9961*
1.5 M	*1.0000*	*0.0034*	*0.9999*	*0.9975*
2 M	*1.0000*	*0.0025*	*0.9999*	*0.9983*

tests, because better variance reduction parameters should lead to better behaved results more efficiently.

The first test problems is a fixed-source duct streaming problem that incorporate the physics of optically thick regions with long streaming paths. Specifically, a three-legged duct benchmark is selected. A global neutron mesh tally is overlaid with resolution 15 x 143 x 66, making each mesh element having a side length of about 9 cm. It should be noted that implicit capture and forced collisions in air are also used to help with the calculation. Eight iterations are conducted and weight-window maps are generated each iteration.

The second problem is an eigenvalue calculation; namely, a variant on the k-effective of the World problem. This version uses a 9 x 9 x 9 array of 2 cm radius spheres of 20 g/cc plutonium-239 spaced 10 cm apart in a bath of water at 1 g/cc. The entire problem is surrounded by a 10 cm thick water reflector at the same density. As is, this configuration is subcritical. The center sphere is then replaced with a 4 cm sphere of plutonium-239 at the same density, making this problem supercritical. This type of problem has been classically used to demonstrate the dangers arising from bias in eigenvalue calculations. The goal of this problem will be to compute the neutron flux throughout this entire system on a uniform 50 x 50 x 50 grid.

The global importance maps are generated with the LTC technique. The three-legged duct problem calculates fluxes for 7.5 minutes in the first iteration and increases it by that amount of time for each subsequent iteration (e.g., the fourth iteration runs for 30 minutes). The k-effective of the World problem is based upon batch sizes with the base iteration being 5000 particles per cycle for 100 active cycles (100 are skipped) with the batch size increasing by that amount in subsequent iterations. Once a weight-window map for each iteration I is obtained, production calculations are run to test the validity of the map toward achieving a well-behaved solution. For the three-legged duct problem, one billion neutron histories are run corresponding to about two hours of computation time. For the k-effective of the World problem, 20000 neutrons per cycle for 500 active cycles (100 cycles are skipped for convergence) are used.

The results of the four tests are shown in Table II with the italicized numbers indicating passage of the test. For both problems without global importance maps, none of the statistical checks pass for the given number of histories. For the duct problem, five or six LTC-WWG iterations are required to pass the tests given N of one billion histories. The k-effective of the World problem requires about three LTC-WWG iterations. Note that a negative R^2 value implies that the fit is very poor and not applicable; additionally, the result of the test is not influenced by the possibility of a negative R^2 value.

For the three-legged duct problem, Figs. 4.a and 4.b show the mean and variance of the density function with the number of histories for various weight-window generator iterations. Likewise,

Table II: Statistical test quantities for various LTC-WWG iterations for two test cases.

	Three-Legged Duct				k-effective of the World			
I	ζ	Θ	$R^2_{\bar{v}}$	$R^2_{\sigma^2_{\bar{v}}}$	ζ	Θ	$R^2_{\bar{v}}$	$R^2_{\sigma^2_{\bar{v}}}$
0	0.9947	0.6787	-1.558	-53.03	0.9987	0.4628	0.8204	-0.587
1	0.9956	0.6759	-12.76	-70.66	*1.0000*	0.0948	0.9539	-2.8651
2	0.9999	0.5655	-114.8	-476.9	*1.0000*	0.0511	*0.9939*	0.6497
3	*1.0000*	0.4246	-91.57	-206.7	*1.0000*	*0.0087*	*0.9995*	*0.9568*
4	*1.0000*	0.2521	*0.8132*	-10.84	*1.0000*	*0.0048*	*0.9991*	*0.9869*
5	*1.0000*	*0.0026*	*0.9994*	-5.296	*1.0000*	*0.0025*	*0.9999*	*0.9989*
6	*1.0000*	*0.0001*	*0.9996*	*0.9825*	*1.0000*	*0.0050*	*0.9998*	*0.9787*

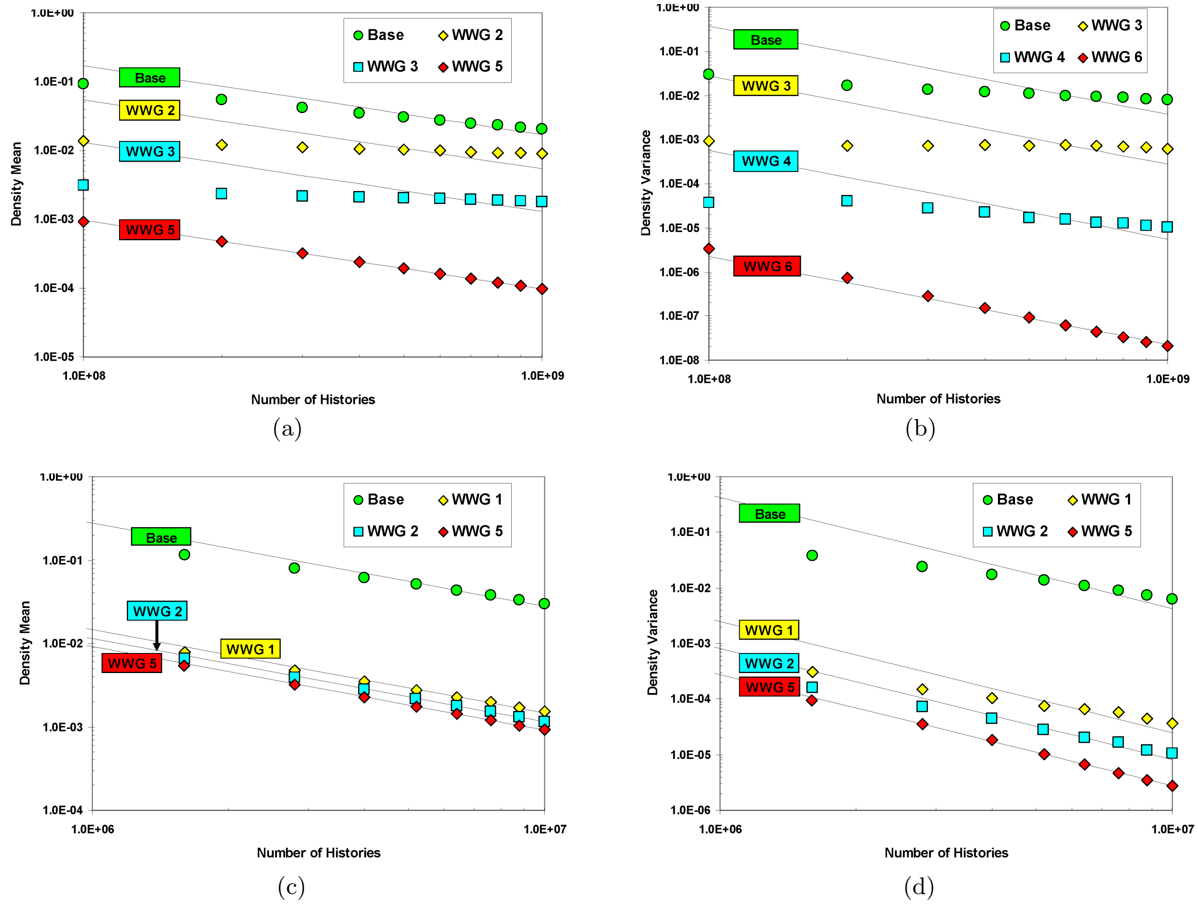


Figure 4: Density mean and variance with curve fits for the test problems: (a) Density mean for the three-legged duct, (b) density variance for the three-legged duct, (c) density mean for the k-effective of the World problem, and (d) density variance for the k-effective of the World problem.

Figs. 4.c and 4.d show the mean and variance of the density function with the number of histories for various weight-window generator iterations for the k-effective of the World problem. On all figures, the fits of $1/N$ and $1/N^2$ are plotted for each curve. Visually, the early iterations do not follow the fit well, whereas better importance maps show a better match with the theoretical fit.

4. CONCLUSIONS

Properties of the relative variance density function are developed and demonstrated on simplified statistical models. Specifically, the convergence rates of the density mean and variance are demonstrated numerically along with the ability of those to detect the misbehavior of individual elements. Four statistical checks for mesh distributions are developed based upon this density function and the associated moments and demonstrated on three different problems; two of these are difficult enough to require global importance functions generated by some means, namely an iterative LTC technique.

Further evidence for the validity of these tests; namely an entire suite of problems demonstrating their applicability will be required before these tests should be implemented in production level software. Also, the issue of efficiency of variance reduction parameters towards achieving a converged solution (i.e., figures of merit) is not addressed, but it is quite likely moments of the relative variance density function are applicable to that as well.

ACKNOWLEDGMENTS

Funding for this work was provided by the U.S. Department of Energy Criticality Safety Program.

REFERENCES

1. S. P. Pederson, R. A. Forster, and T. E. Booth. Confidence Interval Procedures for Monte Carlo Transport Simulations. *Nucl. Sci. Eng.*, **127**, pp. 54-77 (1997).

2. X-5 Monte Carlo Team. MCNP – A General N-Particle Transport Code, Version 5 – Volume I: Overview and Theory. Los Alamos National Laboratory Report, LA-UR-03-1987 (2003).

3. R. E. Maerker and F. J. Muckenthaler. Neutron Fluxes in Concrete Ducts Arising from Incident Thermal Neutrons: Calculations and Experiment. *Nucl. Sci. Eng.*, **29**, pp. 444-454 (1967).

4. C. J. Solomon and A. Sood. Duct Streaming Validation Benchmark Calculations with a Global Importance Map. *Trans. Am. Nuc. Soc.*, **101**, pp. 638-642 (2009).

5. G. E. Whitesides. A Difficulty in Computing the k-effective of the World. *Trans. Am. Nuc. Soc.*, **14**, No. 2, pp. 680 (1971).

6. F. B. Brown. Revisiting the 'K-effective of the World' Problem. *Trans. Am. Nuc. Soc.*, **102**, pp. 302-304 (2010).

7. J. C. Wagner, E. D. Blakeman, and D. E. Peplow. Forward-Weighted CADIS Method for Global Variance Reduction. *Trans. Am. Nuc. Soc.*, **97**, pp. 630-633 (2007).

8. C. J. Solomon, A. Sood, and T. E. Booth. A Weighted Adjoint Source for Weight-Window Generation by Means of a Linear Tally Combination. *Proc. International Conference on Mathematics, Computational Methods, and Reactor Physics (M&C 2009)*, Saratoga Springs, NY, May 3-7, 2009, CD-ROM (2009).

9. T. E. Booth. Automatic Importance Estimation in Forward Monte Carlo Calcuations. *Trans. Am. Nuc. Soc.*, **41**, pp. 308 (1982).

10. D. E. Peplow, E. D. Blakeman, and J. C. Wagner. Advanced Variance Reduction Strategies for Optimizing Mesh Tallies in MAVRIC. *Trans. Am. Nuc. Soc.*, **97**, pp. 595-597 (2007).

11. S. P. Pederson. Mean Estimation in Highly Skewed Samples. Los Alamos National Laboratory Report, LA-12114-MS (1991).