

Implementation of Tally Convergence Tests in UNIST Monte Carlo Code MCS

Matthieu Lemaire, Hyunsuk Lee, Deokjung Lee*

Department of Nuclear Engineering, Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan,
44919, Republic of Korea

*Corresponding author: deokjung@unist.ac.kr

1. Introduction

Monte Carlo transport codes are commonly used to tally quantities of interest related to the design, performances and safety of nuclear systems. Due to the stochastic process, the output of a Monte Carlo code is never a single value, but instead a probability distribution characterized by a mean value m and an absolute standard deviation t . In the limit of the central limit theorem [1], for a large enough number of particle histories, this probability distribution is Gaussian and the probability that the “true tally mean” lies in the range $[m-\alpha t; m+\alpha t]$ is given by:

$$P(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{\alpha} \exp(-x^2/2) dx, \quad (1)$$

where commonly used values are $P(1) = 68.27\%$, $P(2) = 95.45\%$ and $P(3) = 99.73\%$.

The usual issue that users encounter when using a Monte Carlo code is to determine when the number of particles histories is “high enough” for the central limit theorem to apply. One wants to avoid running too few particles histories and obtaining results that “look like converged” but that are actually not converged. Another pitfall is, one also wants to avoid running too many particles histories to guarantee convergence and suffering from a high computational burden and potentially waste of computer time. Therefore, built-in diagnosis tests to assess if a Monte Carlo result is already converged, or adversely needs more particles histories to converge, go a long way helping users make more well-informed choices about their Monte Carlo calculations and increase the user-friendliness and user-efficiency of Monte Carlo codes overall.

This paper presents the tally statistical tests, i.e. tally convergence diagnosis tests, recently implemented in the Monte Carlo MCS [2] developed by Ulsan National Institute of Science and Technology. The MCS statistical tests are based on the well-known 10 statistical tests of the Monte Carlo code MCNP [3]. Theory behind the tests and a practical MCS output are presented successively.

2. Methods and Results

2.1 Definitions

For a tally X , the tally density function f is defined as the history score probability density function (PDF) for selecting a random particle walk that scores $x > 0$ to the tally. The PDF f is normalized so that $\int_0^{+\infty} f(x) dx = 1$. The true mean $E[x]$ and the true variance σ^2 of the population of x score values are defined respectively as:

$$E[x] = \int_0^{+\infty} x \cdot f(x) dx, \quad (2)$$

$$\sigma^2 = E[x^2] - E[x]^2. \quad (3)$$

For a Monte Carlo estimation with N particle histories, x_i being the score of the i^{th} particle history to the tally X , the true mean is approximated by the sample mean \bar{x} and an estimation $S_{\bar{x}}^2$ of the variance of this sample mean is given according to the expressions:

$$\bar{x} = \sum_{i=1}^N x_i / N, \quad (4)$$

$$S_{\bar{x}}^2 = \sum_{i=1}^N (x_i - \bar{x})^2 / (N(N - 1)). \quad (5)$$

The output of the Monte Carlo simulation for the tally X is the pair $(\bar{x}, R = S_{\bar{x}}/\bar{x})$ where R is the relative standard deviation associated to the sample mean \bar{x} . In the limit of the central limit theorem, R converges towards $\frac{\sigma}{\bar{x}\sqrt{N}}$ (i.e. R decreases in $1/\sqrt{N}$ for $N \rightarrow +\infty$).

The figure of merit (FOM) is defined as $FOM = (R^2 * T)^{-1}$ where T is the simulation time in minutes. Since $R \propto 1/\sqrt{N}$ and $T \propto N$, FOM is expected to remain constant for large enough N . The larger the FOM, the more efficient the Monte Carlo simulation and the applied variance reduction techniques are.

The variance of the variance (VOV), a.k.a. the estimated relative variance of R , is a statistical indicator involving the 3rd and 4th moment of the PDF f . It can be estimated during a Monte Carlo simulation with the following expression:

$$VOV = \frac{\sum_{i=1}^N (x_i - \bar{x})^4}{[\sum_{i=1}^N (x_i - \bar{x})^2]^2} - \frac{1}{N}. \quad (6)$$

The inverse power slope (short “slope”) of the PDF f is defined as an estimation of the asymptotic behavior of the PDF f for large score values: for $x \rightarrow +\infty$, $f(x) \approx x^{-slope}$. For the central limit theorem to apply, the true mean $E[x]$ and the true variance σ^2 must be well-defined, finite quantities; this is the case if the quantity $E[x^2] = \int_0^{+\infty} x^2 \cdot f(x) dx \approx \int_{\varepsilon}^{+\infty} x^{2-slope} dx$ with $\varepsilon > 0$, is finite, which requires that $slope > 3$.

2.2 Tally fluctuation chart

When the 10 statistical tests are turned on for a specific cell or mesh tally, MCS computes a tally fluctuation chart (TFC) of size n rows ($n \leq 20$) by 6 columns for every time bin, energy bin and cell/mesh of the mesh/cell tally. Each row of the TFC corresponds to a given number of particles histories (NPH), the n^{th} row corresponding to the total NPH specified by the user in the MCS input. For each line, six columns are provided: NPH, \bar{x} , R , VOV, slope and FOM. An example of TFC output of MCS is shown in Fig. 1.

```

ct3 TIME=1 CELL=3 ENERGY=2
ct3 | | nps | mean | error | vov | slope | fom |
ct3 | | 50200 | 1.05339E+02 | 2.79636E-04 | 0.0089 | 10.00 | 7.6787452E+08 |
ct3 | | 100400 | 1.05328E+02 | 1.99861E-04 | 0.0045 | 10.00 | 7.5075502E+08 |
ct3 | | 150600 | 1.05331E+02 | 1.65076E-04 | 0.0029 | 10.00 | 7.3234316E+08 |
ct3 | | 200800 | 1.05331E+02 | 1.42088E-04 | 0.0022 | 10.00 | 7.4006007E+08 |
ct3 | | 251000 | 1.05331E+02 | 1.26271E-04 | 0.0017 | 10.00 | 7.4917537E+08 |
ct3 | | 301200 | 1.05334E+02 | 1.14512E-04 | 0.0014 | 10.00 | 7.5797646E+08 |
ct3 | | 351400 | 1.05328E+02 | 1.06557E-04 | 0.0013 | 10.00 | 7.4970120E+08 |
ct3 | | 401600 | 1.05330E+02 | 9.96657E-05 | 0.0011 | 9.05 | 7.4976258E+08 |
ct3 | | 451800 | 1.05328E+02 | 9.39976E-05 | 0.0011 | 7.17 | 7.4914226E+08 |
ct3 | | 502000 | 1.05324E+02 | 9.01130E-05 | 0.0010 | 6.39 | 7.3345797E+08 |
ct3 | | 552200 | 1.05323E+02 | 8.58056E-05 | 0.0009 | 4.42 | 7.3495871E+08 |
ct3 | | 602400 | 1.05322E+02 | 8.22865E-05 | 0.0009 | 4.19 | 7.3247832E+08 |
ct3 | | 652600 | 1.05323E+02 | 7.91927E-05 | 0.0008 | 3.88 | 7.3019243E+08 |
ct3 | | 702800 | 1.05321E+02 | 7.65713E-05 | 0.0007 | 3.59 | 7.2507285E+08 |
ct3 | | 753000 | 1.05320E+02 | 7.38553E-05 | 0.0007 | 3.34 | 7.2743928E+08 |
ct3 | | 803200 | 1.05320E+02 | 7.16302E-05 | 0.0006 | 2.82 | 7.2505768E+08 |
ct3 | | 853400 | 1.05315E+02 | 6.96590E-05 | 0.0006 | 2.61 | 7.2117653E+08 |
ct3 | | 903600 | 1.05312E+02 | 6.75862E-05 | 0.0005 | 2.47 | 7.2337243E+08 |
ct3 | | 953800 | 1.05312E+02 | 6.58068E-05 | 0.0005 | 2.38 | 7.2274220E+08 |
ct3 | | 1000000 | 1.05315E+02 | 6.40848E-05 | 0.0005 | 2.14 | 7.2691319E+08 |

```

Fig. 1. Example of TFC output by MCS.

2.3 Implementation of 10 tests

MCS performs 10 checks on each generated TFC to help the user assess the precision of the output result (\bar{x}, R) . We note js the index of the first TFC row in the last half of the problem (i.e. $js \geq n/2$) for which $\bar{x} \neq 0$. We define jr , jv as the first indexes $\geq js$ for which $R \neq 0$, respectively $VOV \neq 0$.

The 1st test consists in checking that the sample mean \bar{x} has a nonmonotonic behavior (i.e. no up or down trend) as a function of NPH for the last half of the problem. If the sample mean stays constant, i.e. $\left(1 - \frac{\bar{x}(jx)}{\bar{x}(n)}\right) \leq 10^{-7}$, then the test is considered as passed. If the sample mean does not stay constant, then if the sample mean keeps on decreasing or keeps on increasing from row js to row n , the test is failed, else it is passed.

The 2nd test consists in checking if the final R value (R value of the n^{th} row of the TFC) has an acceptable magnitude, i.e. is below some arbitrary low boundary. A value of 0.1 is adopted for this test: if $R(n) < 0.1$ the test is passed, else it is failed.

The 3rd test consists in checking whether R decreases monotonically as a function of NPH for the last half of the problem. For each index j from $js+1$ to n , if $R(j) \leq R(j-1)$ then the test is passed. If $R(j) > R(j-1)$, the test is not automatically failed. Instead it is checked if the increase of the R value is still compatible with stochastic fluctuations. In practice, if $R(j) > R(j-1)$, the test is considered failed only if

$$\sqrt{\frac{NPH(j)}{NPH(j-1)} \frac{R(j)}{R(j-1)}} > \max(1.1, C3), \quad (7)$$

where the coefficient C3 is a measure of how much the R value can reasonably be expected to fluctuate:

$$C3 = 1 + \sqrt{VOV(j-1)} + 2.5 * VOV(j-1). \quad (8)$$

The 4th test consists in checking that R decreases as $1/\sqrt{NPH}$ (as should theoretically be the case) for the last half of the problem. This 4th test fails automatically if the 3rd test is failed. Otherwise, it is investigated whether

the quantity $Q4 = \sqrt{\frac{NPH(jr) R(jr)}{NPH(n) R(n)}}$ is reasonably close to 1.

In practice the test is passed if and only if

$$\max\left(Q4, \frac{1}{Q4}\right) \leq \max(1.05, C4), \quad (9)$$

where the coefficient C4 is defined using $VM = \max(VOV(jr), VOV(n))$ as

$$C4 = \sqrt{1 + 5 * \sqrt{VM} + 12.5 * VM}. \quad (10)$$

The 5th test consists in checking if the final VOV value (VOV value of the n^{th} row in the TFC) has an acceptable magnitude, i.e. is below some arbitrary low boundary. A value of 0.1 is adopted for this test: if $VOV(n) < 0.1$ the test is passed, else it is failed.

The 6th test consists in checking if VOV decreases monotonically as a function of NPH for the last half of the problem. This could be done by checking if the inequality $VOV(j) \leq VOV(j-1)$ is valid for each row j from $js+1$ to n . In practice, the variations of VOV are subject to strong statistical fluctuations, and this test is therefore relaxed by only demanding that $VOV(j) \leq 2 * VOV(j-1)$ for each row j from $js+1$ to n .

The 7th test consists in checking that VOV decreases as $1/NPH$ (as should theoretically be the case) for the last half of the problem. This 7th test fails automatically if the 6th test is failed. Otherwise, it is investigated whether the quantity $Q7 = \frac{NPH(jv) VOV(jv)}{NPH(n) VOV(n)}$ is reasonably close to 1. In practice the test is passed if and only if

$$\max\left(Q7, \frac{1}{Q7}\right) < 1.5. \quad (11)$$

The 8th test consists in checking whether the FOM remains constant (as should theoretically be the case) for the last half of the problem. It is therefore investigated if the quantity $Q8 = \frac{FOM(js)}{FOM(n)}$ is reasonably close to 1. In practice the test is passed if and only if

$$\max\left(Q8, \frac{1}{Q8}\right) < \max(1.02, \min(1.5, C8)), \quad (12)$$

where the coefficient C8 is defined similarly to the coefficient C4 in Eq. (10) but with $VM = \max(VOV(js), VOV(n))$.

The 9th test consists in checking that the FOM has a nonmonotonic behavior (i.e. no up or down trend) as a function of NPH for the last half of the problem. This is done in a straightforward way by checking if the FOM keeps on decreasing or keeps on increasing from row js to row n . If this is the case, the test is failed, else it is passed.

The 10th test consists in checking that the final slope value (slope value of the n^{th} row in the TFC) is > 3 , thus guaranteeing that the quantities $E[x]$ and σ^2 are well-defined and finite.

For the sake of an example, Fig. 2 presents the MCS output table summarizing the results of the 10 tests associated with the TFC of Fig. 1.

2.4 Calculation of the PDF inverse power slope

The slope of the tally density function f is estimated as follows. For each time bin, energy bin and cell/mesh of a selected mesh/cell tally MCS keeps in memory throughout the simulation the 200 largest contributions x_i to the tally (“the high-score tail of the tally”). If fewer than 500 non-zero scores overall are recorded for the

tally, the slope is not calculated at all and is assigned a value of zero, thus failing the 10th test. In the opposite case, MCS first checks for the presence of an upper boundary $B > 0$ of the tally density function. In practice, if the 100 largest scores do not differ from each other within 1%, it is considered that the tally density function has an upper boundary. In this case, the integral $E[x^2] = \int_0^B x^2 \cdot f(x) dx$ is always finite and a value of 10 is assigned to the slope, thus passing the 10th test. In the absence of an upper boundary, MCS proceeds to the actual estimation of the slope. The n_f largest scores are used for the slope estimation, where n_f is given by $n_f = \min(200, 0.05 * N_{x>0})$ and $N_{x>0}$ is the number of non-zero scores for the tally. It is assumed that the n_f largest scores are sampled from a generalized Pareto PDF defined by the parameters a and k :

$$P_{a,k}(x) = \frac{1}{a} \left(1 + \frac{kx}{a}\right)^{-\left(1+\frac{1}{k}\right)}. \quad (13)$$

The Melder-Neard downhill simplex method [4] is applied to determine values of the parameters a and k that best fit the n_f largest scores. The slope is then calculated as $\text{slope} = 1 + 1/k$. Slope values greater than 10 (equivalent of a “perfect score”) are set equal to 10.

2.5 Empirical history score PDF

When the 10 statistical tests are turned on for a specific cell or mesh tally, MCS computes the empirical history score PDF f for every time bin, energy bin and cell/mesh of the mesh/cell tally. The empirical PDF is calculated on a logarithmically spaced array of size 600 spanning from 10^{-30} to 10^{+30} , representing 60 decades of possible score values. The discretized value of the empirical PDF f between the array boundaries b_i and $b_{i+1} = 10^{0.1} * b_i$ is calculated as

$$f(i) = \frac{N_i}{N * (b_{i+1} - b_i)}, \quad (14)$$

where N_i is the number of score values x_i such that $b(i) \leq x_i < b_{i+1}$ and N is the total number of simulated particle histories.

The empirical tally PDF is displayed on a log-log plot in MCS output. For the sake of an example, the plot associated to the TFC of Fig. 1 is shown in Fig. 3. This allows the user to check visually the shape of the empirical PDF, especially the shape of the high-score tail. For well-behaved tallies, the PDF high-score tail must be decreasing and should not contain unsampled score regions.

In addition, if the estimated slope of the tally density function is not equal to zero or ten, the Pareto fit is also displayed on the log-log plot of the empirical tally density function with “s” markers. The user can this way perform a visual control of the quality of the Pareto PDF fit applied on the high-score tail points of the tally density function. In the cases where the high-score tail points cannot be sampled by nature with a Pareto PDF, the Pareto PDF fit may not reflect the best estimate of the

slope of the tally density function. In those cases, the user can simply use a ruler and the log-log plot of the tally density function with the “s” markers to check the calculated value of the slope and determine a more accurate estimate of the slope of the high-score tail.

2.6 Analysis of an example

Fig. 1, Fig. 2 and Fig. 3 are successively shown after each other in MCS output. Fig. 1 reveals that the information provided is about a cell tally (“ct3”) and specifically about the quantity tallied in the 1st time bin (“TIME=1”) in the 3rd cell (“CELL=3”) in the 2nd energy bin (“ENERGY = 2”) of said cell tally. According to Fig. 2, this tally did not pass one of the 10 statistical tests: the slope estimated with a Pareto fit of the 199 largest scores x_i of the tally equals 2.1368, which is smaller than 3. Looking at Fig. 3, the user can check with a ruler the Pareto fit estimate of the slope. The red line on the right of Fig. 4 covers the “s” markers printed by MCS. Another red line, parallel to the red line covering the “s” markers, is drawn in Fig. 4 for clarity. One can see that this latter red line decreases by about 2 decades (vertical lines indicated by a “d”) when the score bin goes up by one decade from $1.58E+01$ to $1.58E+02$. This decrease of about 2 decades per decade is consistent with the slope estimate of 2.1368 calculated by MCS. However, one can notice that the “s” markers do not represent accurately the decrease of the high-score tail of the tally density function. Most likely, the discrete distribution of the 200 largest scores is such that by nature, it cannot be sampled with a Pareto PDF. In that case, the user can position the ruler according to the blue line in Fig. 4 to get a more accurate estimate of the slope. The blue line decreases by more than 5 decades when the score bin goes up from $1.00E+02$ to $1.58E+02$, that is, roughly 0.3 decade. Thus, despite the MCS warning, the user can be sure that the slope of the tally density function is at least 15 and that the true mean and variance are therefore finite and well-defined. The tally can therefore be accepted by the user even if the 10th test seemingly failed.

3. Conclusions

Tally convergence diagnosis tools inspired by the 10 statistical tests of MCNP have been implemented in the Monte Carlo neutron-photon transport code MCS developed by Ulsan National Institute of Science and Technology. Empirical history score probability density functions, tally fluctuation charts and results of statistical tests are now available in MCS. These new features are designed to help the user assess more efficiently the convergence of MCS calculation results, thus increasing the user-friendliness of MCS. In the future, these new features will be especially useful in simultaneous use with weight-window-based variance reduction techniques currently being implemented in MCS.

Acknowledgments

This research was supported by the project(L18S015000) by Korea Hydro & Nuclear Power Co. Ltd..

REFERENCES

[1] Rouaud M, "Probability, Statistics and Estimation," [online PDF] retrieved 2018, July 4, link: <http://www.incertitudes.fr/book.pdf>, 2013.

[2] Jang J et al., "Validation of UNIST Monte Carlo Code MCS for Criticality Safety Analysis of PWR Spent Fuel Pool and Storage Cask," Annals of Nuclear Energy 114: 495-509, 2018.
[3] X-5 Monte Carlo team, "MCNP — A General Monte Carlo N-Particle Transport Code, Version 5. Volume I: Overview and Theory," Los Alamos National Laboratory, LA-UR-03-1987, 2005.
[4] Nelder JA, Mead R, "A simplex method for function minimization," Computer Journal. 7: 308-313, 1965.

```
tfc bin --mean-- -----relative error----- ----variance of the variance---- --figure of merit-- -pdf-
      |behavior value decrease decrease rate value decrease decrease rate value behavior slope
desired random <0.10 yes 1/sqrt(nps) <0.10 yes 1/nps constant random >3.00
observed random 0.00 yes yes 0.00 yes yes constant random 2.14
passed? yes yes yes yes yes yes yes yes yes no
TESTS FAILED: 1 out of 10
```

of histories with non-zero contributions to the tally: 1000000
of histories with contributions in the range from 1.0E-30 to 1.0E+30 : 1000000
The estimated inverse power slope of the 199 largest scores starting at 1.2043E+02 is 2.1368

!Warning: at least one statistical test failed. Try increasing the number of histories and/or active cycles.

Fig. 2. Example of MCS output table summarizing the results of the 10 tests applied on the TFC of Fig. 1.

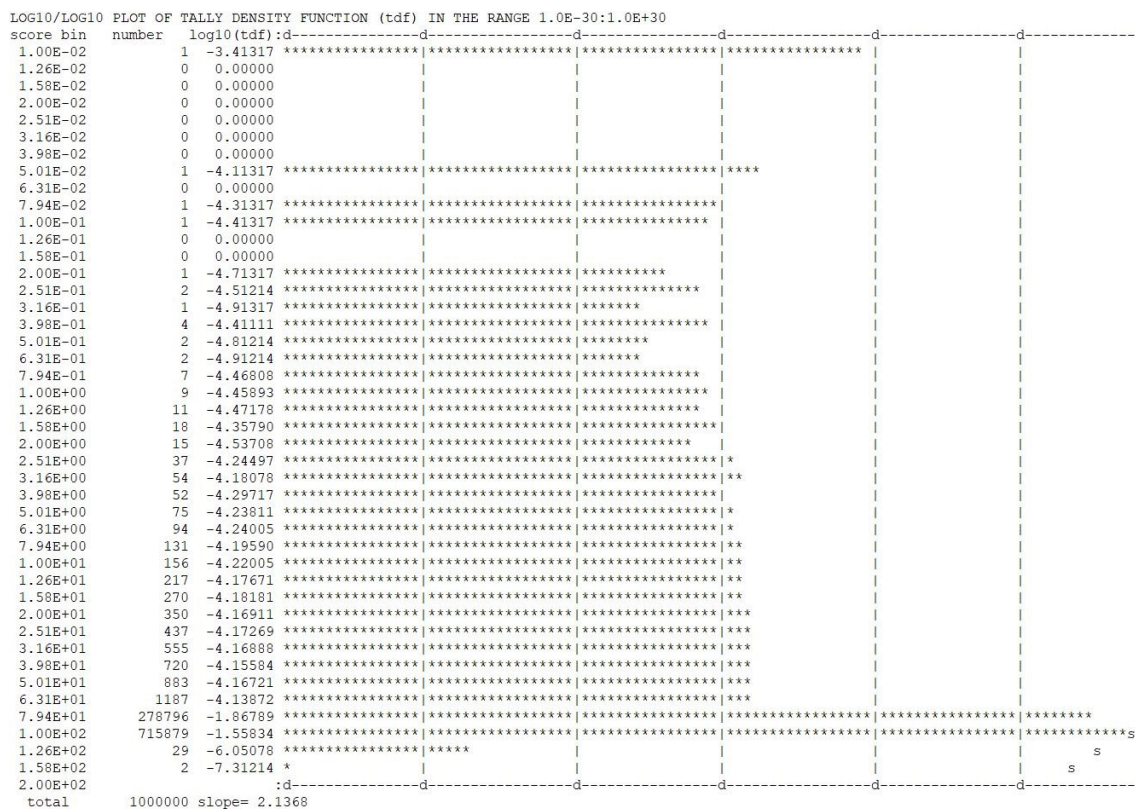


Fig. 3. Example of MCS tally density log-log plot associated with the TFC of Fig. 1.

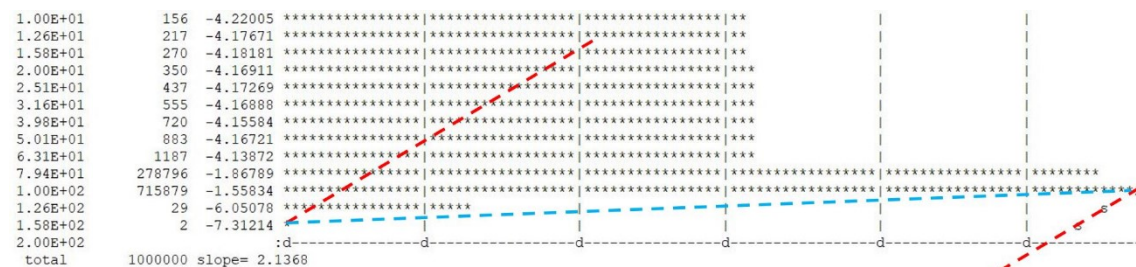


Fig. 4. Zoom on the high-score tail of the tally density. The Pareto fit slope estimate is highlighted with a red line over the "s" markers. The two red lines are parallel to each other. The slope measured by the user with a ruler is shown with a blue line.