Mathematics and Computers in Simulation XXIII (1981) 21–33

STATISTICAL PERFORMANCE ANALYSIS OF NONLINEAR STOCHASTIC SYSTEMS BY THE MONTE CARLO METHOD

James H. TAYLOR
Department of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK 74078, U.S.A.

Until the recent advent of extended covariance analysis utilizing quasi-linearization techniques, the only approach for assessing the performance of a nonlinear system with random inputs and initial conditions has been the monte carlo method. This method involves direct simulation, i.e., determining the system response to a finite number of "typical" initial conditions and noise input functions which are generated according to their specified statistics, and averaging over the resulting ensemble of responses ("trials") to obtain estimated or sample statistics.

While the monte carlo method remains the most general trustworthy technique available for the estimation of nonlinear system performance statistics, sample statistics may be unreliable unless hundreds or perhaps thousands of trials are performed. Since computer budget constraints may not permit such extensive simulation of high-order systems, there is often a temptation to "make do" with sample statistics based on 20 to 25 trials. While this procedure might provide meaningful results in cases that are "nearly gaussian", it is dangerous to rely on limited sample statistics if deviations from normality are significant. This point is demonstrated in detail in this paper using a generalized confidence band concept (not Chi-square) as a measure of the reliability of monte carlo sample statistics for nongaussian random variables. Some compensatory approaches are discussed, including estimating higher-order moments and generating histograms (approximate cumulative distributions).

1. Introduction

Monte carlo methods provide a straightforward approach to the statistical analysis of the performance of a nonlinear system with random inputs, based on direct simulation. It entails determining the system response to a finite number of "typical" initial conditions and noise input functions which are generated according to their specified statistics. Thus, the information required for monte carlo analysis includes the system model, initial condition statistics, and random input statistics.

The system model can be given in the form of a state vector differential equation,

$$\dot{z} = g(z, y, t)\,, \tag{1}$$

where z is the vector of system states, y a vector of random inputs, and $g(z, y, t)$ represents the nonlinear time-varying dynamic relationships in the system. We assume at the outset that the elements of y are correlated random processes with deterministic components that may be nonzero; in this case, we can use a system model of the form

$$\dot{x} = f(x, t) + G(t)\, w(t)\,, \tag{2}$$

where x is an *augmented* state vector, $x^{\mathrm{T}} = [z \,\vdots\, y]^{\mathrm{T}}$, and w the sum of a vector of white noise processes and a deterministic vector which serves as the input to generate the random vector y. Such a model can generally be obtained that is equivalent to (1). Henceforth, we treat (2) as the basic system model. It is portrayed in block diagram notation in Fig. 1.

We will specify the initial condition of the state vector by assuming that the state variables are jointly normal. Thus, given an initial mean vector and covariance matrix [1],

$$\begin{aligned} &\mathrm{E}[x(0)] = m_0\,, \\ &\mathrm{E}[(x(0) - m_0)(x(0) - m_0)^{\mathrm{T}}] = P_0 \end{aligned} \tag{3}$$

the initial condition specification is complete. As stated above, the input vector w is assumed to be

[1] E[ ] denotes the expected value of the bracketed variable.

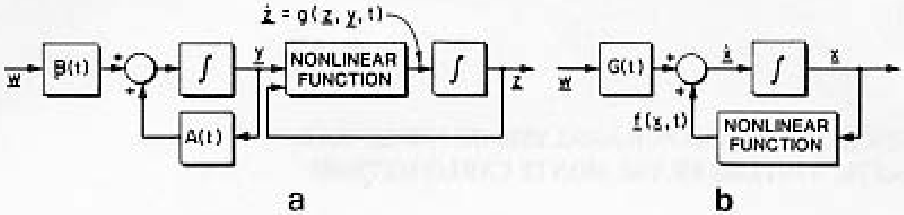

Fig. 1. Nonlinear stochastic system model; (a) example form of eq. (1); (b) form of eq. (2).

composed of elements that are white noise processes, plus an additive deterministic component or mean; thus

$$\mathrm{E}[w(t)] = b(t)\,, \qquad\qquad (4)$$
$$\mathrm{E}[(w(t)-b(t))(w(\tau)-b(\tau))^{\mathrm{T}}] = Q(t)\,\delta(t-\tau)\,,$$

where $Q(t)$ is the input spectral density matrix and the impulse function $\delta(t-\tau)$ indicates that the input vector random components have zero autocorrelation for $t \neq \tau$; i.e., the quantity $u(t) = w(t) - b(t)$ is 'white noise', as stated.

Given the above system model and statistical information, monte carlo analysis requires a large number, denoted q, of representative simulations of the system response. This involves performing the following procedure q times: First, an initial condition vector is chosen according to the statistics indicated above; i.e., a random number generator calculates the elements of a random vector $x(0)$ based on (3). Then a random initial input vector, $w(0)$, is generated, using the statistics given in (4) [2]. These vectors provide the data for evaluation of $\dot{x}(0)$ in (2) which in turn is used to propagate the solution from $t = 0$ to $t = h$ according to any standard technique for the digital integration of a state vector differential equation. Then, given $x(h)$, simulation continues by the generation of a new value of the input noise vector $w(h)$, evaluation of $\dot{x}(h)$, numerical integration to obtain $x(2h)$ and so on, to the specified terminal time t_f.

Performing q independent simulations yields an ensemble of state trajectories, each denoted $x^{(i)}(t; x^{(i)}(0), w^{(i)}(t))$ to stress the dependence of the trajectory on the random initial condition and noise input sample function:

$$\left.\begin{array}{l} x^{(1)}(t; x^{(1)}(0), w^{(1)}(t)) \\ x^{(2)}(t; x^{(2)}(0), w^{(2)}(t)) \\ \quad\vdots \\ x^{(q)}(t; x^{(q)}(0), w^{(q)}(t)) \end{array}\right| \; ; 0 \leqslant t \leqslant t_f\,. \qquad (5)$$

Each satisfies the state vector differential equation (2) to within the accuracy of the numerical integration method used, and the ensembles of initial conditions, $x^{(i)}(0)$, and random inputs, $w^{(i)}(t)$, obey the statistical conditions given in (3) and (4), subject to the limitations of the random number generator employed. The mean $m(t)$ and covariance $P(t)$ of the state vector are then estimated by averaging over the ensemble of trajectories using the relations

$$\hat{m}(t) \triangleq \frac{1}{q}\sum_{i=1}^{q} x^{(i)}(t) \simeq m(t)\,, \qquad (6)$$
$$\hat{P}(t) \triangleq \frac{1}{q-1}\sum_{i=1}^{q} (x^{(i)}(t) - \hat{m}(t))(x^{(i)}(t) - \hat{m}(t))^{\mathrm{T}} \simeq P(t)\,,$$

where $\hat{m}(t)$ and $\hat{P}(t)$ denote the estimated values [3].

[2] We simulate white noise with spectral density matrix $Q(t)$ by using a random number generator to obtain an independent sequence of random vectors $u(kh)$, $k = 0, 1, 2, \ldots$ satisfying

$$\mathrm{E}[u(kh)] = \mathbf{0}\,, \qquad \mathrm{E}[u(kh)\,u^{\mathrm{T}}(kh)] = \frac{1}{h}Q(kh)\,.$$

Then we define $u(t)$ by

$$u(t) = u(kh)\,, \qquad kh \leqslant t < (k+1)\,h\,,$$

where h is a small time increment. For h small ($1/h$ much larger than the bandwidth of the system in question), $u(t)$ is an accurate approximation to a white noise process.

[3] In estimating P, we observe that it is necessary to divide by $(q-1)$, since the sample variance,

$$P_s \triangleq \frac{1}{q}\Sigma_{i=1}^{q}\,(x^{(i)} - \hat{m})(x^{(i)} - \hat{m})^{\mathrm{T}}$$

is biased [5], i.e.,

$$\mathrm{E}[P_s] = \frac{q-1}{q}\,P\,.$$

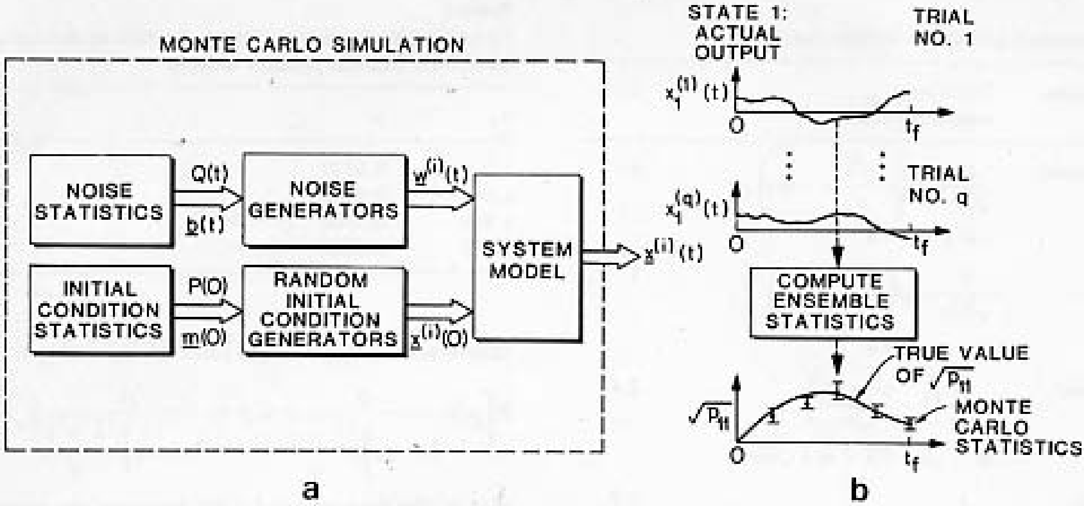

Fig. 2. Schematic characterization of the monte carlo technique; (a) random ensemble generator; (b) ensemble statistical analysis.

The essence of the monte carlo technique is illustrated in Fig. 2.

2. Assessment of accuracy – confidence intervals

In order to assess the accuracy of the approximate statistics given in (6), it is necessary to consider the statistical properties of the estimates $\hat{m}(t)$ and $\hat{P}(t)$. To simplify the notation, consider a scalar random variable y (e.g., the value of some system state variable at some time of interest), and let m and p represent the true values of the mean and variance of y,

$$m = \mathbf{E}[y] , \quad p = \mathbf{E}[(y - m)^2] . \tag{7}$$

By performing one set of q monte carlo trials, we obtain a single estimate of m and p, which we denote $\hat{m}$ and $\hat{p}$. These estimates are also random variables; that is, if another set of q monte carlo trials were performed independently of the first set, but with the same statistics for the initial conditions and noise inputs, then a different ensemble of simulations results, and different estimates for the mean and variance would be obtained. If q is sufficiently large, then we can invoke the central limit theorem to justify the assumption that the random variables $\hat{m}$ and $\hat{p}$ are gaussian [4], and thus that their distributions are asymptotically specified by the following statistics for large q [2]:

$$\mathbf{E}[\hat{m}] = m , \quad \sigma_{\hat{m}}^2 \triangleq \mathbf{E}[(\hat{m} - m)^2] = \frac{p}{q} ,$$
$$\mathbf{E}[\hat{p}] = p , \quad \sigma_{\hat{p}}^2 \triangleq \mathbf{E}[(\hat{p} - p)^2] = \frac{\mu_4 - p^2}{q} , \tag{8}$$

where μ_4 is the fourth central moment,

$$\mu_4 = \mathbf{E}[(y - m)^4] . \tag{9}$$

For many common probability density functions (pdf's), a constant λ exists such that

$$\mu_4 = \lambda p^2 . \tag{10}$$

Table 1 gives a summary of values of λ, known as the kurtosis or excess of the density, for some common pdf's. For pdf's of this type, we can express both of the standard deviations of the estimated statistics

[4] For $q < 20$, it is necessary to assume that $\hat{p}$ has the chi square distribution if y is a gaussian variable [2]. If y is significantly nongaussian, the validity of the Gaussian assumption for $\hat{m}$ and $\hat{p}$ may require considerably more than twenty trials.

Table 1
Some common probability density functions

Designation	Functional representation [a]	λ
Exponential	$\frac{1}{\sqrt{2}\sigma}\exp\left(-\frac{\sqrt{2}}{\sigma}\lvert x-m\rvert\right)$, $-\infty < x < +\infty$	6
Normal	$\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right)$, $-\infty < x < +\infty$	3
Triangular	$\frac{1}{\sqrt{6}\sigma}\left(1-\frac{\lvert x-m\rvert}{\sqrt{6}\sigma}\right)$, $m-\sqrt{6}\sigma \leqslant x \leqslant m+\sqrt{6}\sigma$	2.4
Uniform	$\frac{1}{\sqrt{12}\sigma}$, $m-\sqrt{3}\sigma \leqslant x \leqslant m+\sqrt{3}\sigma$	1.8
Bipolar (Discrete)	$\frac{1}{2}\delta(x-m-\sigma) + \frac{1}{2}\delta(x-m+\sigma)$	1.0

[a] Formulated to have mean m and standard deviation σ.

given in (8) in terms of the true variance, p, to obtain

$$\sigma_{\hat{m}} = \sqrt{\frac{p}{q}}, \qquad \sigma_{\hat{p}} = \sqrt{\frac{\lambda-1}{q}}\,p\,. \tag{11}$$

The above discussion of the statistics of the Gaussian random variable $\hat{p}$ provides the basis for determining a range in the vicinity of $\hat{p}$ such that the true value of p is guaranteed to lie within that range with a specified probability, ψ. This is done by determining the number, n_σ, of standard deviations, $\sigma_{\hat{p}}$, such that

$$\mathbf{P}[0 \leqslant |p-\hat{p}| \leqslant n_\sigma \sigma_{\hat{p}}] = \psi\,. \tag{12}$$

Since $\hat{p}$ is approximately gaussian, n_σ is the solution to

$$\frac{1}{\sqrt{2\pi}}\int_{-n_\sigma}^{n_\sigma} \exp(-\tfrac{1}{2}\zeta^2)\,\mathrm{d}\zeta = \psi\,. \tag{13}$$

For example, if the desired probability is 0.95, (13) yields $n_\sigma = 1.96$. Other values of n_σ corresponding to different values of ψ can be obtained from probability integral tables [4]; several representative values are given in Table 2.

To reformulate (12) into an inequality for p, we

Table 2
Cumulative probability within n_σ standard deviations of the mean for a Gaussian random variable

n_σ	ψ
1.0	0.6827
1.645	0.9000
1.960	0.9500
2.576	0.9900

substitute for $\sigma_{\hat{p}}$ from (11) into (12) to obtain

$$\mathbf{P}\left[\underline{p} \triangleq \frac{\hat{p}}{1+n_\sigma\sqrt{\frac{\lambda-1}{q}}} \leqslant p \leqslant \frac{\hat{p}}{1-n_\sigma\sqrt{\frac{\lambda-1}{q}}} \triangleq \overline{p}\right] = \psi \tag{14}$$

that is, the true value of p lies between the values $\underline{p}$ and $\overline{p}$ indicated in (14) with probability ψ. Alternatively, in terms of the estimated rms value of the variable, $\hat{\sigma}$, we have the comparable result

$$\mathbf{P}[\underline{\sigma} \leqslant \sigma \leqslant \overline{\sigma}] = \psi\,,$$

where $\underline{\sigma}$ and $\overline{\sigma}$ are given by

$$\underline{\sigma} \triangleq \sqrt{\underline{p}} = \frac{\hat{\sigma}}{\left[1+n_\sigma\sqrt{\frac{\lambda-1}{q}}\right]^{1/2}} \triangleq \underline{\rho}\hat{\sigma}\,, \tag{15}$$

$$\overline{\sigma} \triangleq \sqrt{\overline{p}} = \frac{\hat{\sigma}}{\left[1-n_\sigma\sqrt{\frac{\lambda-1}{q}}\right]^{1/2}} \triangleq \overline{\rho}\hat{\sigma}\,. \tag{15}$$

The quantities $\underline{\sigma}$ and $\overline{\sigma}$ are referred to as *lower and upper confidence limits*; the value of ψ expressed as a percent is the *degree of confidence*. Eq. (15) demonstrates that the standard deviation confidence limits can be obtained from $\hat{\sigma}$ simply by using the multipliers $\underline{\rho}$ and $\overline{\rho}$. The latter are functions only of the kurtosis, λ, the number of monte carlo trials, q, and the number of standard deviations, n_σ, required to achieve the desired degree of confidence.

The problem of making a reasonable choice of λ, which depends upon the statistics of the random variable y, must be faced before the confidence limit multipliers can be calculated. One option is to determine an approximate value of λ by estimating the fourth central moment using the q sample values of the variable y, and calculating

$$\lambda \simeq \hat{\mu}_4/\hat{p}^2 \triangleq \hat{\lambda}\,.$$

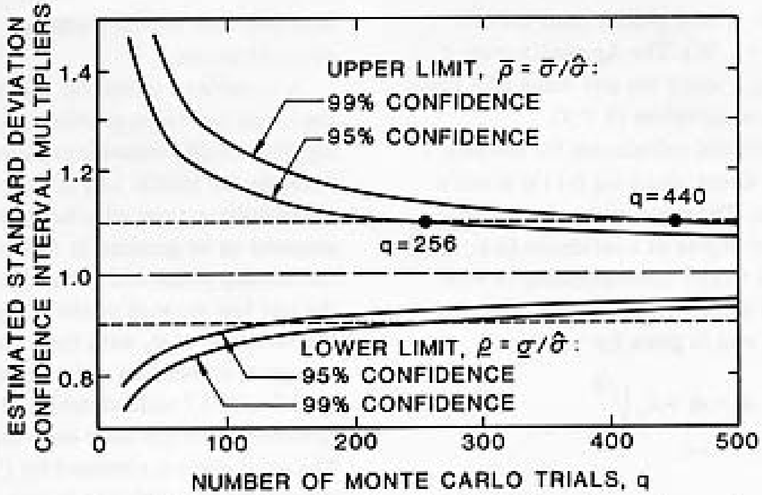

Fig. 3. Typical confidence interval multipliers for the estimated standard deviation of a gaussian random variable (λ = 3).

The value of λ need not be known exactly, since the confidence limits $\underline{\sigma}$ and $\bar{\sigma}$ are not extremely sensitive to errors in this parameter. Unfortunately, as we note in a subsequent example, a meaningful estimate of λ can often require several hundred trials. In the absence of reliable information about the higher central moments, it is frequently assumed for the confidence limit calculation that y is Gaussian; i.e., that λ = 3. However, if there is any reason to believe that the pdf for y has abnormally heavily weighted tails – as in the case of the exponential distribution in Table 1, for example – then a larger value of λ may be required in order to arrive at a realistic assessment of the accuracy of an estimated rms value obtained via the monte carlo technique.

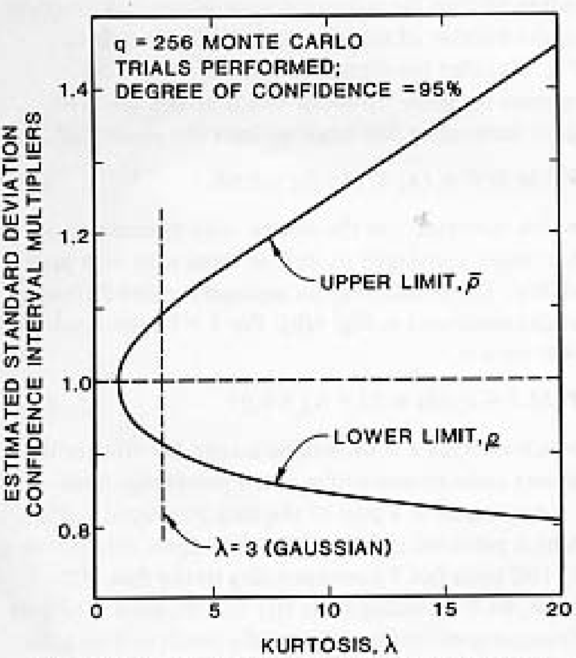

Fig. 4. Effect of kurtosis on confidence interval limits.

Values of $\underline{\rho}$ and $\bar{\rho}$ for λ = 3 are indicated as functions of the number of monte carlo trials in Fig. 3, for two typical values of confidence. As an example of the significance of the confidence interval, if we desire to have 99% certainty that σ is within 10% of the estimated value, $\hat{\sigma}$; i.e.,

$$P[0.90\hat{\sigma} \leqslant \sigma \leqslant 1.1\hat{\sigma}] = 0.99 \,, \qquad (16)$$

then Fig. 3 demonstrates that it is necessary to perform 440 trials; 256 trials suffice for 95% confidence [5]. Additional cases are available in the confidence multiplier tables provided in the Appendix.

Fig. 4 shows the deterioration that occurs in the accuracy of the monte carlo estimated standard deviation, for a given level of confidence, if the kurtosis of the random variable is greater than 3 due to y being nongaussian. We discuss an instance where $\lambda \simeq 15$ in Section 3; in this case, even for 256 trials, the upper

[5] Note that the bounds, $\underline{\rho}$ and $\bar{\rho}$ are not symmetric with respect to one; thus the point at which $\bar{\rho}$ crosses 1.1 determines the value of q for which (16) is satisfied.

95% confidence limit is 36% greater than the estimated value of σ ($\bar{\rho} = 1.36$). The Appendix gives a method of evaluating $\underline{\rho}$ and $\bar{\rho}$ for any value of λ using the corresponding normal values ($\lambda = 3$).

The confidence interval calculation for the estimated mean is quite direct, since $\sigma_{\hat{m}}$ ((11)) is not a function of the mean. The same value of n_σ is obtained for the desired degree of confidence (e.g., from (13), $n_\sigma = 1.96$ for $\psi = 0.95$ corresponding to 95% confidence), and the value [6] of $\hat{p}$ is used in deriving the result that for $\underline{m}$ and $\bar{m}$ given by

$$\underline{m} = \hat{m} - n_\sigma \sqrt{\frac{\hat{p}}{q}}, \qquad \bar{m} = \hat{m} + n_\sigma \sqrt{\frac{\hat{p}}{q}} \tag{17}$$

one can assert that

$$P[\underline{m} \leqslant m \leqslant \bar{m}] = \psi \,. \tag{18}$$

Here, we see that $\underline{m}$ and $\bar{m}$ cannot be readily expressed in terms of a multiple of $\hat{m}$.

The confidence limit concept developed above provides a statistical measure of the accuracy of the estimated mean and standard deviation of a random variable obtained by using the monte carlo method. It is only possible to assess the accuracy of such estimates in a probabilistic sense; e.g., for 256 trials, we can assert, for example, that an estimated standard deviation (rms value) of a gaussian random variable is within 10% of the true value, with probability 0.95 (with 95% confidence). We note below that even this assessment may be open to question if kurtosis is not known at least approximately, however.

3. Illustrative example and the basic dilemma

Considerable practical experience has been gained in applying the monte carlo method in studies undertaken to validate the use of a more recently-developed describing function method called CADET to provide accurate and efficient performance evaluations for tactical missile guidance systems [3,4,6]. The significance of the confidence interval concept and the important role played by kurtosis have been graphically demonstrated by the results obtained, as the following example shows.

A variable of particular interest in the planar missile-target intercept problem during the terminal homing phase is the cross-range (lateral) separation between the missile and target, denoted y. In a typical analysis, y (and all other system variables) is assumed to be gaussian at the initiation of the terminal homing phase and y remains quite gaussian until the last few seconds of the engagement. Fig. 5 shows the variation of σ_y with time during a six-second engagement, where a quite highly nonlinear system model with 17 state variables, 9 nonlinearities and 5 random inputs has been used for simulation purposes. The solid curve is obtained by CADET [3,4,6], and the results of a 500-trial monte carlo study are indicated with circled data points to indicate $\hat{\sigma}_y$ and vertical I-bars to indicate the 95% confidence interval. The estimated value of kurtosis is also indicated near each data point; as observed above, $\hat{\lambda}$ is nearly 3 until the last second, while at the final time, $t = 6$ sec, $\hat{\lambda}$ is 15, which is indicative of the quite highly nongaussian character of the final lateral separation (miss distance).

Fig. 6 gives a more detailed view of the CADET and monte carlo analysis depicted in Fig. 5; for two values of time the estimated σ_y is shown as a function of the number of trials performed, q. We note in Fig. 6(a) that the estimated value of σ_y at $t = 4$ appears to 'settle' to about 145 ft after a few hundred trials; after 500 trials we have the result that

$$P[138 \text{ ft} \leqslant \sigma_y(4) \leqslant 156 \text{ ft}] = 0.95 \tag{19}$$

which indicates that the monte carlo estimate of σ_y has nearly converged to its true value with high probability. The situation at six seconds is quite different, as demonstrated in Fig. 6(b). For $\hat{\lambda} = 15$ the result of 500 trials is

$$P[24.7 \leqslant \sigma_y(6) \leqslant 33.9 \text{ ft}] = 0.95 \tag{20}$$

which indicate a considerable margin for error in the monte carlo estimate of σ_y, on a percentage basis.

A synopsis of a part of the data portrayed in Fig. 6(b) is provided in Table 3, broken down into five sets of 100 trials (set 1 corresponding to the first 100 trials, set 2 including trials 101 to 200, etc.). The data demonstrates that in this case the result of 100 trials is highly random – with $\hat{\sigma}_y(6)$ varying between

[6] While $\sigma_{\hat{m}}$ is given by $\sqrt{p/q}$ in (1), the true value of p is unknown. Since $\hat{m}$ and $\hat{p}$ are independent [4], it can be shown that $\hat{p}$ can be used instead.

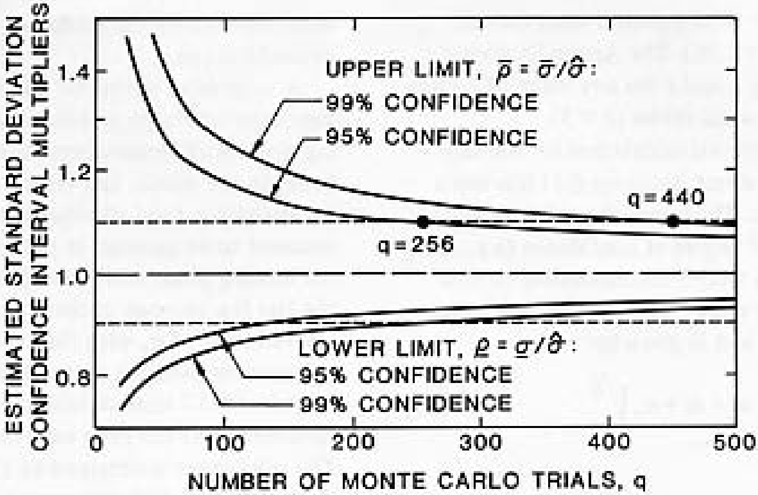

Fig. 3. Typical confidence interval multipliers for the estimated standard deviation of a gaussian random variable ($\lambda = 3$).

The value of λ need not be known exactly, since the confidence limits $\underline{\sigma}$ and $\bar{\sigma}$ are not extremely sensitive to errors in this parameter. Unfortunately, as we note in a subsequent example, a meaningful estimate of λ can often require several hundred trials. In the absence of reliable information about the higher central moments, it is frequently assumed for the confidence limit calculation that y is Gaussian; i.e., that $\lambda = 3$. However, if there is any reason to believe that the pdf for y has abnormally heavily weighted tails – as in the case of the exponential distribution in Table 1, for example – then a larger value of λ may be required in order to arrive at a realistic assessment of the accuracy of an estimated rms value obtained via the monte carlo technique.

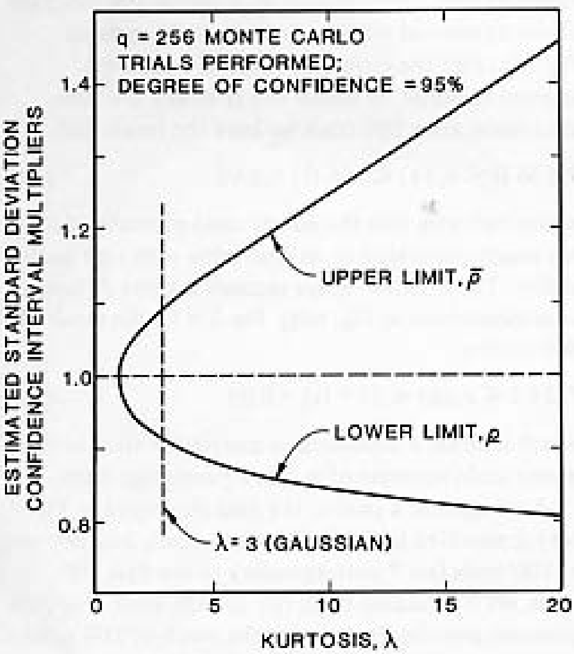

Fig. 4. Effect of kurtosis on confidence interval limits.

Values of $\underline{\rho}$ and $\bar{\rho}$ for $\lambda = 3$ are indicated as functions of the number of monte carlo trials in Fig. 3, for two typical values of confidence. As an example of the significance of the confidence interval, if we desire to have 99% certainty that σ is within 10% of the estimated value, $\hat{\sigma}$; i.e.,

$$P[0.90\hat{\sigma} \leqslant \sigma \leqslant 1.1\hat{\sigma}] = 0.99\,, \tag{16}$$

then Fig. 3 demonstrates that it is necessary to perform 440 trials; 256 trials suffice for 95% confidence [5]. Additional cases are available in the confidence multiplier tables provided in the Appendix.

Fig. 4 shows the deterioration that occurs in the accuracy of the monte carlo estimated standard deviation, for a given level of confidence, if the kurtosis of the random variable is greater than 3 due to y being nongaussian. We discuss an instance where $\lambda \simeq 15$ in Section 3; in this case, even for 256 trials, the upper

[5] Note that the bounds, $\underline{\rho}$ and $\bar{\rho}$ are not symmetric with respect to one; thus the point at which $\bar{\rho}$ crosses 1.1 determines the value of q for which (16) is satisfied.

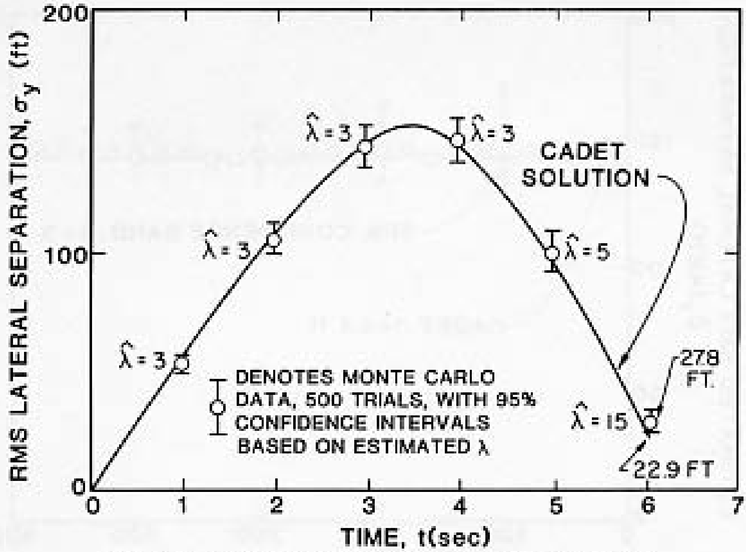

Fig. 5. Time history of rms missile-target lateral separation.

19.72 ft and 35.88 ft; the variation exhibited by $\hat{\lambda}$ is even more dramatic. We also observe that there exists a clear relation between $\hat{\lambda}$ and $\hat{\sigma}_y$; $\hat{\sigma}_y$ is small if $\hat{\lambda}$ is small and $\hat{\sigma}_y$ is large if $\hat{\lambda}$ is large. This phenomenon is a direct result of the basic significance of kurtosis: if λ is appreciably larger than 3, then the 'tails' of the density function are abnormally heavily weighted – implying that there is a high probability of the occurrence of very large values of the random variable in comparison with a gaussian random variable having the same standard deviation. (To cite an example, given two random variables with unity variance, y_1 normally distributed and y_2 exponentially distributed ($\lambda = 6$; Table 1), the probability that $|y_1| \geqslant 3$ is only 0.0027, compared with the probability of 0.0144 that $|y_2| \geqslant 3$.) Thus the incidence of several large values of $|y|$ in the space of a few trials results in a sudden jump in the estimated σ_y, as evident in the vicinity of 160 and 440 trials in Fig. 6(b), while it is probable that the "settling" observed during the third and fourth sets of trials is due to the untypically benign character of these trials (an abnormally small number of trials occurred in which $|y|$ is large).

Table 3 thus demonstrates a fundamental dilemma with the monte carlo method applied to nonlinear systems: *Analysis based on a modest but seemingly reasonable number of trials (e.g., q = 100) may be quite inconclusive unless the value of λ is known quite accurately in advance.* Thus the analyst should be extremely cautious in assessing the reliability of monte carlo estimated statistics, even if the estimated kurtosis is monitored. In the preceding example, the importance of a few large values of miss distance that occur in a set of trials in characterizing the tails of the pdf, and thus in determining the kurtosis of a non-gaussian random variable, also demonstrates that the common practice of "discarding the pathological trials" can lead to very misleading results.

Table 3
Estimated standard deviation and kurtosis for lateral separation, $t = 6$ sec

100-trial set number	$\hat{\sigma}_y$ (ft)	$\hat{\lambda}$
1	19.72	4
2	32.08	15
3	22.25	6
4	25.67	4
5	35.88	23
Aggregate [a] (500 trials)	27.78	15

[a] To obtain aggregated values for $\bar{\sigma}_y$ and $\bar{\lambda}$, it is necessary to average the corresponding values of variance and fourth central moment (Eqs. (8) and (9)).

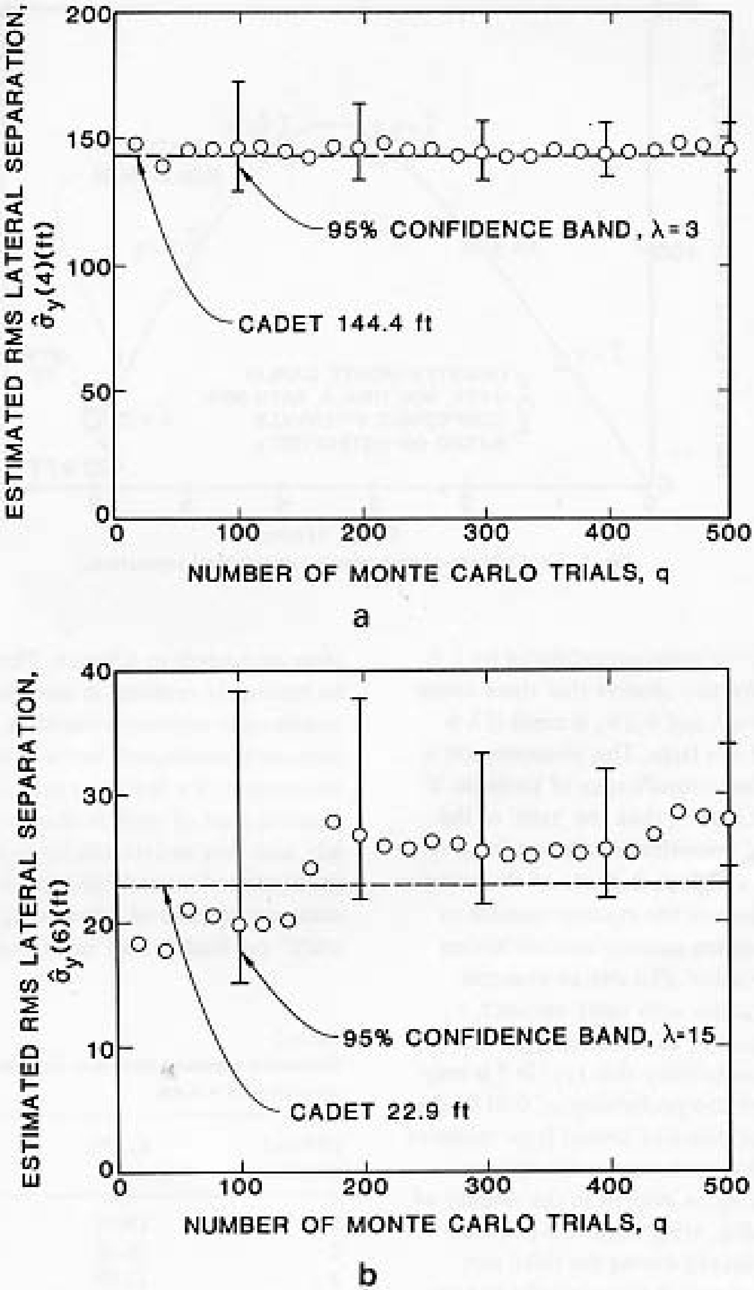

Fig. 6. Comparison of CADET and monte carlo rms lateral separation; (a) $t = 4$ sec; (b) $t = 6$ sec.

4. Resolution of the dilemma

The problem identified in the preceding section can be avoided by abandoning any attempt to estimate the statistics of nongaussian (especially high-kurtosis) variables. A better course of action is to develop approximate cumulative distribution functions or histograms from the simulation ensemble. A

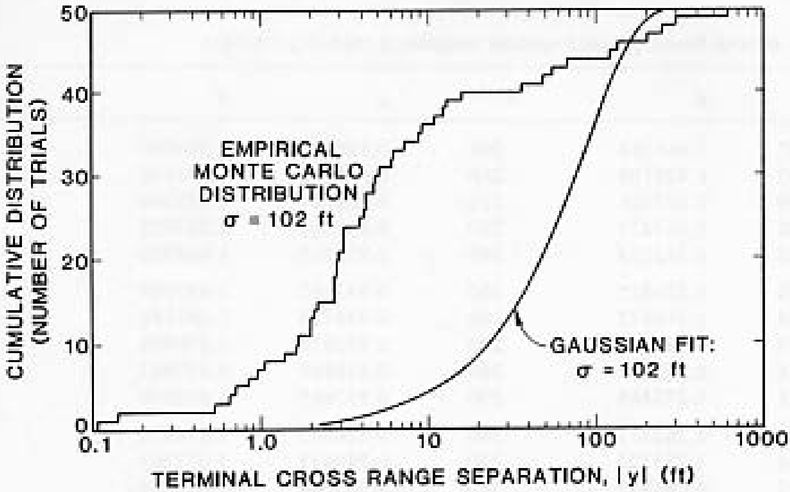

Fig. 7. Empirical distribution of miss distance.

typical empirical miss distance distribution [7] is shown in Fig. 7.

Two advantages to the recommended procedure (generating approximate distribution functions) are readily apparent in Fig. 7. First, it is estimated that the kurtosis of the distribution is $\lambda \simeq 24$; thus for the 50 trials shown, the 95% confidence interval multipliers are undefined (from (25), the equivalent number of trials for the gaussian case is $q_{eq} = 4$ for which (15) is meaningless). Inverting the relations given in the Appendix shows that many more trials ($q = 230$) would be needed to obtain the usual 20-trial 95% confidence limits for a gaussian variable ($\underline{\rho} = 0.78$, $\overline{\rho} = 1.66$). The second advantage is that even if the estimated rms miss distance of 102 ft is nearly correct, this statistic alone does not characterize the missile performance meaningfully unless the underlying distribution is known. To illustrate this point, a gaussian distribution with the same standard deviation is also portrayed in Fig. 7. The latter distribution indicates very poor performance (in one half of the trials miss would exceed 60 ft) while the empirical distribution is quite good (in one half of the trials, the miss is less than 4 ft). Thus, in this case the rms value alone could be quite misleading.

[7] This is not the same case as that shown in Fig. 5.

5. Summary and conclusions

The basic principles of monte carlo analysis for assessing the performance of nonlinear stochastic systems have been outlined, and a method for determining the statistical accuracy of the monte carlo-generated estimates of system variable means and standard deviations has been given. A fundamental difficulty has been illustrated: unless the analyst knows in advance what types of distributions he is dealing with, or can afford to perform a very large ensemble of simulation experiments (perhaps several *thousand* trials), it may be very difficult to quantify the quality of the resulting estimated system variable statistics. The recommended solution to this problem is to use the simulation ensemble to create approximate distribution plots, and base the performance evaluation of the nonlinear stochastic system on this information.

Appendix. Confidence interval limit calculation

The confidence interval limits of an estimated standard deviation $\hat{\sigma}$ can be expressed as multiples of $\hat{\sigma}$, viz.,

$$\underline{\sigma} = \underline{\rho}\,\hat{\sigma}\,, \qquad \overline{\sigma} = \overline{\rho}\,\hat{\sigma}\,, \tag{21}$$

where $\underline{\rho}$ and $\overline{\rho}$ are determined only by the desired degree of confidence, the kurtosis of the random vari-

Table 4
90 percent confidence interval limits, gaussian random variables, q trials ($n_\sigma = 1.645$)

q	$\underline{\rho}$	$\overline{\rho}$	q	$\underline{\rho}$	$\overline{\rho}$
20	0.807487	1.464364	200	0.926523	1.094285
22	0.814433	1.425104	210	0.928115	1.091678
24	0.820599	1.393521	220	0.929604	1.089268
26	0.826126	1.367472	230	0.931001	1.087032
28	0.831122	1.345555	240	0.932315	1.084951
30	0.835670	1.326815	250	0.933555	1.083007
32	0.839834	1.310573	260	0.934726	1.081186
34	0.843669	1.296336	270	0.935835	1.079476
36	0.847216	1.283735	280	0.936888	1.077867
38	0.850511	1.272488	290	0.937889	1.076348
40	0.853584	1.262375	300	0.938843	1.074912
42	0.856458	1.253223	320	0.940621	1.072261
44	0.859157	1.244893	340	0.942248	1.069865
46	0.861696	1.237273	360	0.943745	1.067686
48	0.864092	1.230269	380	0.945127	1.065693
50	0.866358	1.223806	400	0.946410	1.063863
55	0.871527	1.209615	420	0.947605	1.062173
60	0.876101	1.197662	440	0.948721	1.060607
65	0.880188	1.187427	460	0.949767	1.059151
70	0.883870	1.178543	480	0.950750	1.057793
75	0.887213	1.170742	500	0.951676	1.056522
80	0.890265	1.163826	520	0.952550	1.055329
85	0.893068	1.157642	540	0.953378	1.054207
90	0.895655	1.152072	560	0.954163	1.053149
95	0.898053	1.147023	580	0.954908	1.052149
100	0.900284	1.142420	600	0.955617	1.051203
105	0.902367	1.138202	650	0.957251	1.049040
110	0.904318	1.134320	700	0.958713	1.047126
115	0.906150	1.130731	750	0.960032	1.045415
120	0.907876	1.127402	800	0.961230	1.043874
125	0.909506	1.124303	850	0.962324	1.042478
130	0.911048	1.121409	900	0.963329	1.041205
135	0.912509	1.118699	950	0.964257	1.040038
140	0.913898	1.116154	1000	0.965116	1.038963
145	0.915220	1.113760	1100	0.966660	1.037046
150	0.916480	1.111501	1200	0.968013	1.035382
160	0.918833	1.107343	1300	0.969212	1.033922
170	0.920990	1.103600	1400	0.970283	1.032626
180	0.922977	1.100208	1500	0.971248	1.031466
190	0.924815	1.097117	2000	0.974959	1.027075

able, λ, and the number of trials performed, q. These multiples, $\underline{\rho}$ and $\overline{\rho}$, have the form

$$\underline{\rho} = \frac{1}{\left[1 + n_\sigma \sqrt{\frac{\lambda - 1}{q}}\right]^{1/2}}, \qquad \overline{\rho} = \frac{1}{\left[1 - n_\sigma \sqrt{\frac{\lambda - 1}{q}}\right]^{1/2}}, \tag{22}$$

where n_σ is determined by the confidence ψ expressed as a decimal fraction,

$$\frac{1}{\sqrt{2\pi}} \int_{-n_\sigma}^{n_\sigma} \exp(-\tfrac{1}{2}\zeta^2)\, d\zeta = \psi \qquad (23)$$

(see Table 2). This formulation, (21), makes it particularly convenient to present the confidence interval multipliers in tabular form. Thus we include Tables 4 to 6 for easy reference, giving confidence interval multipliers $\underline{\rho}$ and $\overline{\rho}$ for 90% confidence ($\psi = 0.90$), 95% confidence and 99% confidence. These data are directly applicable to gaussian variables, or any other case where $\lambda = 3$.

For other values of kurtosis, the confidence inter-

Table 5
95 percent confidence interval limits, gaussian random variables, q trials ($n_\sigma = 1.960$)

q	$\underline{\rho}$	$\overline{\rho}$	q	$\underline{\rho}$	$\overline{\rho}$
20	0.781848	1.657247	200	0.914210	1.115588
22	0.789372	1.590828	210	0.916033	1.112299
24	0.796071	1.539303	220	0.917740	1.109265
26	0.802092	1.497989	230	0.919342	1.106453
28	0.807546	1.464009	240	0.920850	1.103840
30	0.812522	1.435490	250	0.922274	1.101402
32	0.817087	1.411156	260	0.923620	1.099121
34	0.821299	1.390109	270	0.924896	1.096982
36	0.825201	1.371693	280	0.926107	1.094970
38	0.828832	1.355420	290	0.927259	1.093074
40	0.832223	1.340917	300	0.928358	1.091283
42	0.835400	1.327695	320	0.930407	1.087980
44	0.838386	1.316127	340	0.932284	1.084999
46	0.841109	1.305430	360	0.934012	1.082293
48	0.843856	1.295656	380	0.935610	1.079822
50	0.846373	1.286683	400	0.937094	1.077554
55	0.852123	1.267138	420	0.938476	1.075463
60	0.857222	1.250839	440	0.939769	1.073527
65	0.861789	1.236999	460	0.940980	1.071729
70	0.865911	1.225069	480	0.942120	1.070053
75	0.869659	1.214659	500	0.943193	1.068485
80	0.873087	1.205478	520	0.944208	1.067016
85	0.876240	1.197308	540	0.945169	1.065634
90	0.879153	1.189980	560	0.946080	1.064333
95	0.881857	1.183362	580	0.946946	1.063104
100	0.884375	1.177348	600	0.947770	1.061940
105	0.886729	1.171856	650	0.949670	1.059286
110	0.888936	1.166813	700	0.951372	1.056938
115	0.891011	1.162165	750	0.952908	1.054843
120	0.892967	1.157863	800	0.954305	1.052958
125	0.894815	1.153867	850	0.955581	1.051251
130	0.896566	1.150143	900	0.956754	1.049696
135	0.898227	1.146662	950	0.957837	1.048271
140	0.899806	1.143400	1000	0.958840	1.046960
145	0.901309	1.140335	1100	0.960646	1.044624
150	0.902744	1.137448	1200	0.962229	1.042600
160	0.905425	1.132145	1300	0.963632	1.040823
170	0.907885	1.127385	1400	0.964887	1.039248
180	0.910154	1.123081	1500	0.966018	1.037840
190	0.912256	1.119167	2000	0.970373	1.032517

Table 6
99 percent confidence interval limits, gaussian random variables, q trials ($n_\sigma = 2.576$)

q	$\underline{\rho}$	$\overline{\rho}$	q	$\underline{\rho}$	$\overline{\rho}$
20	0.738071	2.467137	200	0.891497	1.161088
22	0.746410	2.208189	210	0.893721	1.156225
24	0.753870	2.039417	220	0.895806	1.151753
26	0.760604	1.919369	230	0.897765	1.147623
28	0.766728	1.828951	240	0.899613	1.143796
30	0.772334	1.758031	250	0.901358	1.140236
32	0.777496	1.700693	260	0.903010	1.136915
34	0.782271	1.653232	270	0.904578	1.133807
36	0.786709	1.613199	280	0.906068	1.130891
38	0.790849	1.578906	290	0.907487	1.128149
40	0.794725	1.549151	300	0.908840	1.125563
42	0.798365	1.523048	320	0.911370	1.120809
44	0.801794	1.499936	340	0.913690	1.116533
46	0.805031	1.479303	360	0.915830	1.112663
48	0.808095	1.460754	380	0.917811	1.109139
50	0.811001	1.443971	400	0.919653	1.105912
55	0.817664	1.408197	420	0.921371	1.102945
60	0.823597	1.379144	440	0.922979	1.100204
65	0.828928	1.355002	460	0.924488	1.097662
70	0.833757	1.334565	480	0.925909	1.095298
75	0.838160	1.317003	500	0.927249	1.093091
80	0.842199	1.301717	520	0.928517	1.091024
85	0.845922	1.288269	540	0.929718	1.089085
90	0.849371	1.276330	560	0.930858	1.087260
95	0.852579	1.265644	580	0.931942	1.085539
100	0.855572	1.256012	600	0.932976	1.083913
105	0.858376	1.247278	650	0.935359	1.080208
110	0.861009	1.239313	700	0.937498	1.076940
115	0.863489	1.232014	750	0.939431	1.074030
120	0.865831	1.225295	800	0.941191	1.071418
125	0.868047	1.219087	850	0.942801	1.069057
130	0.870148	1.213328	900	0.944282	1.066909
135	0.872145	1.207969	950	0.945651	1.064945
140	0.874046	1.202966	1000	0.946920	1.063139
145	0.875858	1.198283	1100	0.949208	1.059929
150	0.877589	1.193888	1200	0.951216	1.057152
160	0.880830	1.185856	1300	0.952999	1.054720
170	0.883810	1.178686	1400	0.954595	1.052569
180	0.886563	1.172238	1500	0.956035	1.050647
190	0.889118	1.166402	2000	0.961594	1.043409

val can be determined by use of the *gaussian equivalent number of trials,* derived as follows: for a specified degree of confidence (or, equivalently, a given value of n_σ) the multipliers $\underline{\rho}$ and $\overline{\rho}$ (22) are determined solely by the ratio $(\overline{\lambda} - 1)/q$. Thus given a set of monte carlo trials typified by the parameters (q, λ), the confidence interval multipliers are identical to those for $(q_{eq}, 3)$, where q_{eq} is chosen to satisfy

$$\frac{3-1}{q_{eq}} = \frac{\lambda - 1}{q} \tag{24}$$

or

$$q_{eq} = \frac{2q}{\lambda - 1}\,. \tag{25}$$

The desired multipliers $\underline{\rho}$ and $\overline{\rho}$ may then be obtained from the appropriate table of confidence interval multipliers for gaussian random variables under q_{eq}.

Example. In the preceding section we discussed a study of 500 trials, where $\lambda \simeq 15$; to obtain $\underline{\rho}$ and $\overline{\rho}$ use

$$q_{\text{eq}} = \frac{1000}{14} \simeq 70$$

as given in (25). From Table 5, under the entry for 70 trials, we see that the 95% confidence interval limit multipliers are $\underline{\rho} \simeq 0.866$, $\overline{\rho} \simeq 1.225$.

References

[1] A.H. Bowker and G.J. Lieberman, Engineering Statistics (Prentice-Hall, Englewood Cliffs, NJ, 1959).

[2] R.S. Bucy, C. Hecht and K.D. Senne, An engineer's guide to building nonlinear filters, Seiler Research Lab. Report No. SRL-TR-72-0004 (USAF Systems Command) (1972).

[3] A. Gelb, Editor, Applied Optimal Estimation (MIT Press, Cambridge, MA, 1974).

[4] A. Gelb and R.S. Warren, Direct statistical analysis of nonlinear systems: CADET, AIAA J. 11 (5) (1973) 689–694.

[5] A. Papoulis, Probability, Random Variables and Stochastic Processes (McGraw-Hill, New York, 1965).

[6] J.H. Taylor, C.F. Price, J. Siegel and A. Gelb, Statistical analysis of nonlinear stochastic systems via CADET, in: R.V. Ramnath, H.M. Paynter and J.K. Hedrick, Eds., Nonlinear System Analysis and Synthesis, Vol. 2 – Techniques and Applications (ASME Press, 1981).