JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 41, No. 6, November–December 2004

Engineering Notes

ENGINEERING NOTES are short manuscripts describing new developments or important results of a preliminary nature. These Notes cannot exceed six manuscript pages and three figures; a page of text may be substituted for a figure and vice versa. After informal review by the editors, they may be published within a few months of the date of receipt. Style requirements are the same as for regular contributions (see inside back cover).

Simplified Approach to Identifying Influential Uncertainties in Monte Carlo Analysis

Toshikazu Motoda*
Japan Aerospace Exploration Agency, Tokyo 181-0015, Japan

Nomenclature

n	=	number of uncertain parameters incorporated in Monte Carlo simulation
P_j	=	upper cumulative probability of $p_j(x)$
P_{Nj}	=	upper cumulative probability of $\phi(z)$
$p_j(x)$	=	probability of x failures occurring calculated by using a hypergeometric distribution when the null hypothesis is assumed
r	=	probability of taking each element of ε_F to create ε_F^*
$z(x)$	=	transformed random variable that has standard normal distribution
z'	=	continuity collection of z for an upper cumulative probability
α	=	level of significance
ε_F	=	uncertain parameter vector that yields unsatisfactory simulation result
ε_F^*	=	test vector created from ε_F
$\phi(z)$	=	probability density function of a standard normal distribution

Introduction

A SIMPLIFIED method is presented for identifying influential uncertain parameters in Monte Carlo simulation. When this method is used, the identification of influential uncertainties is more convenient and practical for engineers compared to the exact statistical test that is described in Ref. 1 because only a simple statistical test is required.

As computer power has increased, Monte Carlo simulation has become recognized as an effective and powerful tool for system evaluation in the development of aerospace vehicles.[2–6] The advantages of the Monte Carlo simulation are that 1) the nonlinear system can be evaluated directly, 2) the results reflect the influences of combinatorial effects of various uncertain parameters, and 3) the uncertain parameters can be determined as physical values. The primary objective is to maximize the probability of mission achievement obtained from Monte Carlo simulation, in other words, to minimize the probability of simulation outcomes in which the system fails to achieve its design objective. To achieve this, it is necessary to identify those uncertain parameters that have the greatest influence on such unsatisfactory simulation results. Once these influential parameters are identified, the most effective measures can be considered to improve the system to increase the probability of mission success, such as redesign of the control system, improvement of hardware such as sensors and actuators, or further ground testing. However, identifying those uncertain parameters is sometimes difficult because various uncertain parameters are incorporated in the simulation randomly and simultaneously.

An identification method by an exact statistical test that utilizes a hypergeometric distribution was presented in Ref. 1. However, the exact test may be practically inconvenient; because the hypergeometric distribution includes binomial coefficients that become huge as the number of samples increases, the computational ability of even a modern computer may be exceeded. For this reason, this Note presents a simplified statistical test utilizing a normal distribution and investigates and demonstrates its validity.

Identification of Influential Uncertainties

This section introduces the principle of the identification methodology. The identification procedure consists of two main steps. In the first step, Monte Carlo simulation for identification (MCSI) is performed by using test vectors that are generated from uncertain parameter vectors that cause unsatisfactory simulation results. In the second step, the influential uncertain parameters are identified by applying a statistical hypothesis test to the results of MCSI. The following discussion first briefly explains test vector generation and MCSI. Next, the identification method with the exact test which was introduced in Ref. 1 is described.

Test Vector Generation and MCSI

After a system is evaluated by Monte Carlo simulation, an uncertain parameter vector ε_F that yields an unsatisfactory simulation result is obtained. A test vector ε_F^* is created from ε_F as shown in Fig. 1. Elements of ε_F^* are randomly chosen from those of ε_F with a probability of $r = \frac{1}{2}$, and N_{MCSI} test vectors are generated. The purpose of generating the test vectors is to create various combinations of uncertain parameters using the elements of ε_F to investigate the influence of each uncertain parameter $\varepsilon_F(j)$.

Next, N_{MCSI} MCSI runs are performed by using the N_{MCSI} test vectors. From the results of MCSI, N_F unsatisfactory simulation results occur among the N_{MCSI} simulations. The number of simulation cases that have a nonzero value of the jth uncertain parameter is M_j, and M_{Fj} unsatisfactory cases occur among M_j simulations. The relationship among N_{MCSI}, N_F, M_j, and M_{Fj} is shown in Ref. 1. If M_{Fj} is sufficiently large, the jth uncertain parameter is likely to be influential. Influential uncertain parameters are then identified by a statistical hypothesis test.

Exact Hypothesis Testing

In hypothesis testing, a null hypothesis is assumed to be tested. If the data in the sample strongly disagree with the null hypothesis, the null hypothesis is rejected and the conclusion is the negation of the null hypothesis. For identifying influential uncertain

Received 20 April 2004; revision received 19 August 2004; accepted for publication 29 July 2004.

*Senior Researcher, Advanced Control Research Group, Institute of Space Technology and Aeronautics, Aerospace Research Center Aerodrome Branch, 6-13-1 Osawa, Mitaka-City; motoda.toshikazu@jaxa.jp.

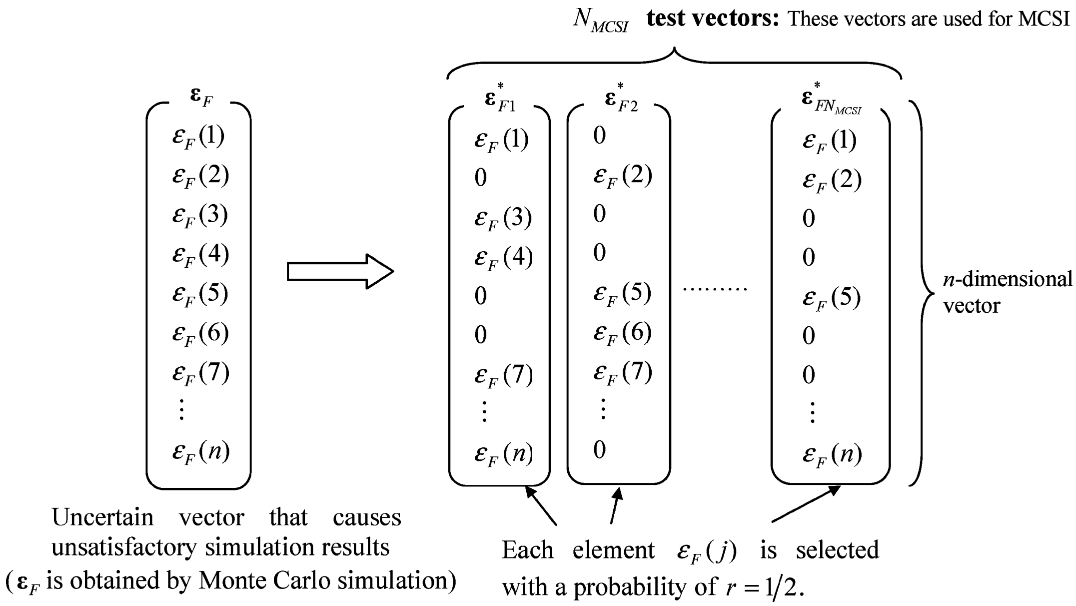


Fig. 1 Generation of test vectors.

parameters, the null hypothesis is "Each uncertain parameter has no influence on whether or not the result of MCSI is satisfactory."

Under the null hypothesis, the probability of x failures occurring among M_j simulations, $p_j(x)$, is given by the hyper-geometric distribution[7]:

$$p_j(x) = \left[\left({}_{N_F}C_x\right)\left({}_{N_{\mathrm{MCSI}}-N_F}C_{M_j-x}\right)\right]\Big/{}_{N_{\mathrm{MCSI}}}C_{M_j}$$
$$x = 0, \ldots, x_{\max}, \qquad x_{\max} = \min\{M_j, N_F\} \tag{1}$$

Because x is a discrete value, $p_j(x)$ also has a discrete type of distribution. The probability that the failure x occurs the same number of times as, or more often than, the experimental result M_{Fj} is given by

$$P_j = \Pr[x \geq M_{Fj}] = \sum_{x=M_{Fj}}^{x_{\max}} p_j(x) \tag{2}$$

where Pr[ ] indicates a probability. The upper cumulative probability P_j is called the p value, that is, the smallest significance level at which the null hypothesis would be rejected for the obtained M_{Fj} (Ref. 7). From Eq. (2), as M_{Fj} becomes larger, the p value becomes smaller and the null hypothesis is likely to be rejected.

The p value is calculated for each uncertain parameter $\varepsilon(j)$, $j = 1, \ldots, n$. It is necessary to find those uncertain parameters that have a sufficiently small corresponding p value. If P_j is smaller than the level of significance α, the null hypothesis is rejected and the jth uncertainty is determined to be an influential uncertain parameter. The value of α is determined by the practitioner, and Ref. 1 gives guidelines as to its value.

Simplified Hypothesis Testing

The simplified hypothesis test utilizes a normal distribution instead of a hyper-geometric distribution. Though this is an approximation, it avoids computational inconvenience. The principle of the simplified hypothesis test is described next and its accuracy is examined in the next section.

Under the null hypothesis, the exact probability distribution is given by Eq. (1), which includes binomial coefficients. If the binomial coefficients are calculated directly, their values become extremely large and computational burden becomes excessive as the number of simulations N_{MCSI} becomes large. Although a large value of N_{MCSI} is desirable to obtain reliable results in hypothesis testing, it may be impossible to carry out the binomial coefficient calculations even if a modern computer is used. Moreover, specialized statistical software may be necessary to perform the identification, but it would be more convenient practically if the Monte Carlo simulation and the identification could be easily performed by the same software package. The simplified hypothesis test does not require specialized statistical software and so makes this feasible. The advantages of the simplified hypothesis test are, therefore, that 1) both the Monte Carlo simulation and the identification can be easily performed automatically by the same software, 2) software and human resources costs are reduced, and 3) computational time is reduced.

The simplified test uses a normal distribution to avoid the inconvenience of the hypergeometric distribution. The number of failures x in Eq. (1) is transformed by a large-sample approximation,[7]

$$z(x) = [x - (M_j \cdot N_F)/N_{\mathrm{MCSI}}] \Bigg/ \sqrt{\frac{M_j \cdot N_F \cdot (N_{\mathrm{MCSI}} - M_j) \cdot (N_{\mathrm{MCSI}} - N_F)}{N_{\mathrm{MCSI}}^2 \cdot (N_{\mathrm{MCSI}} - 1)}} \tag{3}$$

Random variable z is a normalized number of failures x and has a standard normal distribution. This large-sample approximation is improved by using a continuity collection. For upper cumulative probabilities, $\frac{1}{2}$ is subtracted from the numerator,[7] namely,

$$z_j' = z\left(M_{Fj} - \tfrac{1}{2}\right) = \left[\left(M_{Fj} - \tfrac{1}{2}\right) - (M_j \cdot N_F)/N_{\mathrm{MCSI}}\right] \Bigg/ \sqrt{\frac{M_j \cdot N_F \cdot (N_{\mathrm{MCSI}} - M_j) \cdot (N_{\mathrm{MCSI}} - N_F)}{N_{\mathrm{MCSI}}^2 \cdot (N_{\mathrm{MCSI}} - 1)}} \tag{4}$$

For each uncertain parameter j, the approximated p value is obtained by

$$P_{Nj} = \Pr[z \geq z_j'] = \int_{z_j'}^{\infty} \phi(z)\,\mathrm{d}z \tag{5}$$

P_{Nj} corresponds to the exact value of P_j and is obtained using standard statistical tables by looking up values corresponding to z_j'. As shown in Eq. (4), z_j' is easily calculated even when N_{MCSI} becomes large because it does not include binomial coefficients.

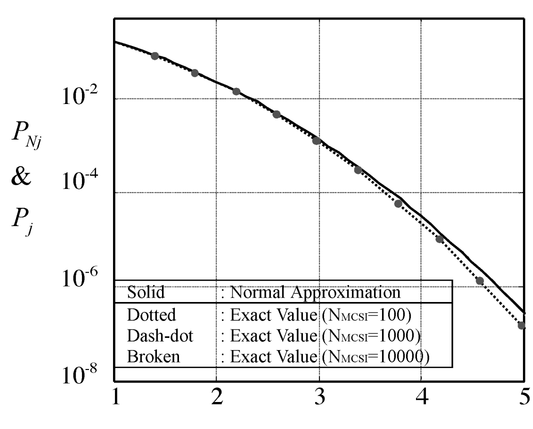
A further important feature of the simplified test is that the ascending order of z'_j, $j = 1, \ldots, n$, corresponds to the descending order of P_{Nj}. This means that the order of P_{Nj} is found when z'_j is obtained without calculating P_{Nj}. In parameter identification, an uncertain parameter is likely to be influential when the corresponding p value is small, and so it is necessary to find uncertain parameters that have small p values. Although in the exact test the order of P_j is unknown until the values of P_j is calculated by using Eqs. (1) and (2), which include binomial coefficients, in the simplified test candidates of the influential uncertain parameter are found just after z'_j is calculated. The simplified test is, therefore, more convenient for practitioners. In the next section, the difference between the identified results of the exact test and the simplified test is discussed.

Approximation Error of the Simplified Test

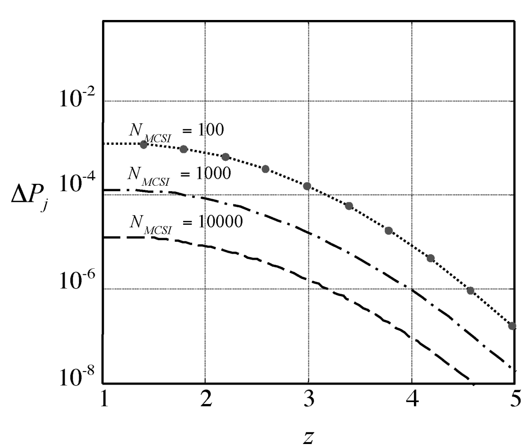
The accuracy of the simplified test is now investigated. As influential uncertain parameters are detected by p value, for the investigation the approximated value P_{Nj} should be compared with the exact value P_j. The simplified p value, $P_{Nj}(z)$, is obtained by Eq. (5), which is an integral of standard normal distribution. However, because the exact p value $P_j(x)$ is a function of x given by Eq. (2), the corresponding value of z is necessary to compare with P_{Nj}.

When a continuity collection is considered, the corresponding value of x is z'_j, which is given by Eq. (4). Because each uncertain parameter is included in a test vector with a probability of $r = \frac{1}{2}$ as shown in Fig. 1,

$$M_j \approx N_{\mathrm{MCSI}}/2 \tag{6}$$



a) Comparison of p values



b) Approximation error

Fig. 2 Comparison between exact and approximated p values: $R = 0.5$.

Thus, Eq. (4) becomes

$$z'_j(M_{Fj}) \approx \left[\left(M_{Fj} - \tfrac{1}{2}\right) - R \cdot N_{\mathrm{MCSI}}/2\right] \Bigg/ \sqrt{\frac{N_{\mathrm{MCSI}}^2 \cdot R \cdot (1-R)}{4 \cdot (N_{\mathrm{MCSI}} - 1)}} \tag{7}$$

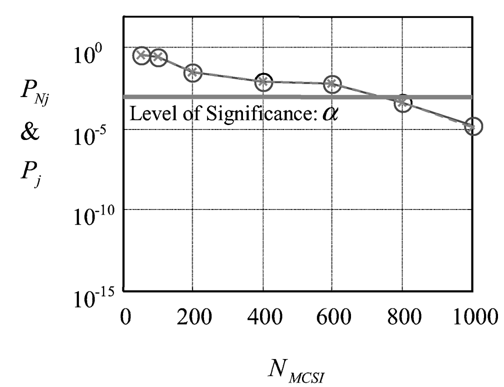
where

$$R = N_F / N_{\mathrm{MCSI}}$$

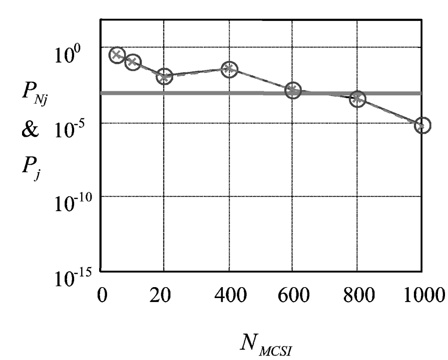
The corresponding exact p value $P_j(M_{Fj})$ is obtained from Eqs. (1) and (2):

$$P_j(M_{Fj}) = \sum_{x = M_{Fj}}^{x_{\max}} \frac{\left\{{}_{R \cdot N_{\mathrm{MCSI}}}C_x\right\} \cdot \left\{{}_{(1-R) \cdot N_{\mathrm{MCSI}}}C_{N_{\mathrm{MCSI}}/2 - x}\right\}}{{}_{N_{\mathrm{MCSI}}}C_{N_{\mathrm{MCSI}}/2}} \tag{8}$$

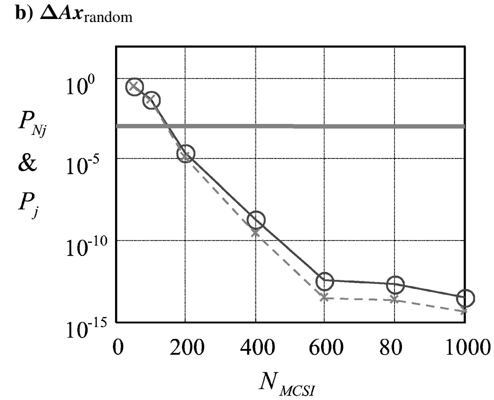
For fixed value of R, the relationship between z'_j and P_j is obtained by Eqs. (7) and (8) and depends on the number of simulations N_{MCSI}, whereas the relationship between z'_j and P_{Nj} is independent of N_{MCSI}.



a) $\Delta Az_{\mathrm{random}}$



b) $\Delta Ax_{\mathrm{random}}$



c) ΔCm_α

Fig. 3 Example of p value comparison for three identified uncertain parameters: ×, exact value P_j and ○, approximated value P_{Nj}.

As an example, Fig. 2 shows exact p values for $N_{\mathrm{MCSI}} = 100$, 1000, and 10,000 when $R = 0.5$. The approximated value P_{Nj} is also shown as a solid line. The vertical axis of the graph has a log scale. As shown in Fig. 2a, the approximated value P_{Nj} is almost coincident with the exact values P_j when N_{MCSI} is more than 100. Approximation errors, ΔP_j, are in Fig. 2b. ΔP_j is determined as

$$\Delta P_j = P_{Nj} - P_j \tag{9}$$

The approximation errors become small as z becomes large. In other words, the approximated value P_{Nj} approaches the exact value P_j as the p value becomes small. Also note that the approximated values P_{Nj} are always greater than the exact values P_j because ΔP_j is always positive. This means that whenever an uncertain parameter is classified as influential by the simplified test, it will also be classified as influential by the exact test. Although Fig. 2 only shows the result for $R = 0.5$, similar results can be obtained for other values of R. From this investigation of approximation error, the simplified test is shown to be valid and convenient for practical use.

Example

To demonstrate the validity of the simplified test, its results are compared with those obtained by the exact test of Ref. 1. In this example, the system is the experimental vehicle for an automatic landing flight experiment,[6,8] which was conducted to establish automatic landing technology for a future Japanese unmanned reentry vehicle.

In Ref. 1, after the system was evaluated by Monte Carlo simulation, influential uncertain parameters on the touchdown sink rate were identified by the exact test. As a result, three influential uncertain parameters were found: an aerodynamic model error ΔCm_α and measurement random noises of accelerometer $\Delta Az_{\mathrm{random}}$ and $\Delta Ax_{\mathrm{random}}$. The p values of these identified parameters vs N_{MCSI} are indicated in Fig. 3 with a symbol $\times$. The corresponding p values obtained by the simplified test are indicated in Fig. 3 with a symbol $\circ$. The horizontal thick solid line indicates the level of significance $\alpha (= 0.001$, namely, $0.1\%)$, and an uncertain parameter is determined as influential when its p value is less than α. In this example, the value of R is about 0.13.

The results show that the approximated p values are almost coincident with the corresponding exact values. In the graph of ΔCm_α, although the difference between the exact and the approximated p values appears to become large as N_{MCSI} becomes large, in fact the difference becomes small as N_{MCSI} becomes large because the vertical axis is shown on a log scale. Furthermore, whenever an uncertain parameter is determined as influential by the simplified test, the parameter is always determined as influential by the exact test as already discussed. This example, therefore, demonstrates that influential uncertain parameters can be effectively identified by the simplified hypothesis test.

Conclusions

A simplified approach of identifying influential uncertain parameters in Monte Carlo analysis was presented, and its validity was investigated. Because computational inconvenience is avoided, the simplified test can be easily applied even when a large number of simulations is required for the identification. The accuracy of the simplified test was investigated, and the validity of the simplified approach was demonstrated in an example. The simplified test is found to be practically efficient, and by its application, Monte Carlo simulation will become a still more efficient tool for system evaluation in the development of flight vehicles.

References

[1] Motoda, T., and Miyazawa, Y., "Identification of Influential Uncertainties in Monte Carlo Analysis," *Journal of Spacecraft and Rockets*, Vol. 39, No. 4, 2002, pp. 615–623.

[2] Shakarian, A., "Application of Monte-Carlo Techniques to the 757/767 Autoland Dispersion Analysis by Simulation," AIAA Paper 83-2193, Aug. 1983.

[3] Ray, L. R., and Stengel, R. F., "Application of Stochastic Robustness to Aircraft Control Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 14, No. 6, 1991, pp. 1251–1259.

[4] Hanson, J. M., and Coughlin, D. J., "Ascent, Transition, Entry, and Abort Guidance Algorithm Design for the X-33 Vehicle," AIAA Paper 98-4409, Aug. 1998.

[5] Way, D. W., Desai, P. N., Engelund, W. C., Cruz, J. R., and Hughes, S. J., "Design and Analysis of the Drop Test Vehicle for the Mars Exploration Rover Parachute Structural Tests," AIAA Paper 2003-2128, May 2003.

[6] Motoda, T., Miyazawa, Y., Ishikawa, K., and Izumi, T., "Automatic Landing Flight Experiment Flight Simulation Analysis and Flight Testing," *Journal of Spacecraft and Rockets*, Vol. 36, No. 4, 1999, pp. 554–560.

[7] Conover, W. J., *Practical Nonparametric Statistics*, 3rd ed., Wiley, New York, 1999, pp. 30, 31, 95–104, 188–190.

[8] Miyazawa, Y., Nagayasu, M., and Nakayasu, H., "Flight Testing of ALFLEX Guidance, Navigation and Control System," International Council of the Aeronautical Sciences, ICAS Paper 98-1.1.3, Sept. 1998.

J. Korte
Associate Editor