We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Latin Hypercube Sampling

Jaroslav Menčík

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/62370>

Abstract

The simultaneous influence of several random quantities can be studied by the Latin hypercube sampling method (LHS). The values of distribution functions of each quantity are distributed uniformly in the interval (0; 1) and these values of all variables are randomly combined. This method yields statistical characteristics with less simulation experiments than the Monte Carlo method. In this chapter, the creation of the randomized input values is explained.

Keywords: Probability, Monte Carlo method, Latin Hypercube Sampling, probabilistic transformation, randomization

The Monte Carlo method has two disadvantages. First, it usually needs a very high number of simulations. If the output quantity must be obtained by time-consuming numerical computations, the simulations can last a very long time, and the response surface method is not always usable. Second, it can happen that the generated random numbers of distribution function F (which serves for the creation of random numbers with nonstandard distributions) are not distributed sufficiently and regularly in the definition interval (0; 1). Sometimes, more numbers are generated in one region than in others, and the generated quantity has thus somewhat different distribution than demanded. This problem can appear especially if the output function depends on many input variables.

A method called Latin Hypercube Sampling (LHS) removes this drawback. The basic idea of LHS is similar to the generation of random numbers via the inverse probabilistic transformation (3) and Figure 2 shown in Chapter 15 [1, 2]. The difference is that LHS creates the values of F not by generating random numbers dispersed in chaotic way in the interval (0; 1), but by assigning them certain fix values. The interval (0; 1) is divided into several layers of the same width, and the x values are calculated via the inverse transformation (F^{-1}) for the F values

corresponding to the center of each layer. With reasonably high number of layers (tens or hundreds), the created quantity x will have the proper probability distribution. This approach is called stratified sampling. If the output variable y depends on several input quantities, $x_1, x_2, \dots, x_m$, it is necessary that each quantity is assigned values of all layers and that the quantities and layers of individual variables are randomly combined. This is done by random assigning the order numbers of layers to the individual input quantities.

Figure 1. Principle of the LHS method.

The procedure is as follows. The definition interval of the distribution function F of each of m variables is divided into N layers. N , the same for all variables, also corresponds to the number of trials (= simulation experiments). In each trial, the order numbers of layers are assigned randomly to the individual variables ($X_1, X_2, \dots, X_m$). In this way, various layers of the individual variables are always randomly combined. In practice, this is achieved by means of random numbers and their rank-ordering. Then, each input variable is assigned the value corresponding to the center of the pertinent layer of its distribution function.

The application is illustrated on a case with four random quantities (X_1, X_2, X_3 , and X_4) and the definition interval of F divided into five layers (Fig. 1). Only five layers are used here for simplicity; usually, several tens of layers are used. In our case, Y will be calculated for five combinations of the four input quantities. Thus, $5 \times 4 = 20$ random numbers with uniform distribution in interval $(0; 1)$ are generated (see the table in the left part of Fig. 2). Then, the layer numbers for variable X_1 (for example) for individual trials are assigned with respect to the order of random values (for X_1) ranked by size from the maximum to minimum. Here, layer no. 3 (with the highest number 0.83) for the first trial, layer no. 1 for the second, no. 5 for the third, no. 2 for the fourth, and no. 4 for the fifth, corresponding to the numbers 0.56 - 0.25 - 0.83 - 0.17 and 0.30 in the column for X_1 . Similar operations are done for each variable. Thus, in the first trial, variables X_1, X_2, X_3 , and X_4 are assigned the values corresponding to the second,

fourth, second, and fifth layers of their distribution functions, respectively. Inverse probabilistic transformation F^{-1} is then used for the determination X_1 from $F_{1,1}$, etc.; see the table on the right. Now, the investigated quantity $Y = Y(X_1, X_2, X_3, X_4)$ is calculated five times. The obtained values Y_1, Y_2, Y_3, Y_4 , and Y_5 can be used for the determination of statistical characteristics (mean, standard deviation,...).

Random numbers (RN)					Layer numbers for individual trials				
Variable	X_1	X_2	X_3	X_4	Variable	X_1	X_2	X_3	X_4
Trial	RN	RN	RN	RN	Trial	layer	layer	layer	layer
1	0.56	0.12	0.72	0.13	1	2	4	2	5
2	0.25	0.40	0.84	0.60	2	4	3	1	3
3	0.83	0.05	0.21	0.55	3	1	5	4	4
4	0.17	0.62	0.03	0.99	4	5	2	5	1
5	0.30	0.70	0.61	0.73	5	3	1	3	2

Figure 2. LHS method: assignment of layers to individual variables and trials.

Usually, several tens or hundreds of trials are made, which enable the construction of distribution function $F(Y)$ and determination of the mean value, standard deviation, various quantiles, and other characteristics.

Author details

Jaroslav Menčík

Address all correspondence to: jaroslav.mencik@upce.cz

Department of Mechanics, Materials and Machine Parts, Jan Perner Transport Faculty,
 University of Pardubice, Czech Republic

References

- [1] Florian A. An efficient sampling scheme: Updated Latin Hypercube Sampling. Probabilistic Engineering Mechanics, 7 (1992), issue 2, 123 – 130.
- [2] Olsson A, Sandberg G, Dahlblom G. On Latin hypercube sampling for structural reliability analysis. Probabilistic Engineering Mechanics. 2002; 25: issue 1, 47 – 68.