

이단계 Latin Hypercube 추출법과 그 응용

임 미 정¹⁾, 권 우 주²⁾, 이 주 호³⁾

요 약

본 논문에서는 컴퓨터 모델을 이용하여 복잡한 시스템을 모형화할 때 결과값의 분포를 보다 정확히 추정하기 위한 입력변수의 추출방법으로서 McKay 등(1979)이 제안한 Latin Hypercube 추출법을 개선한 이단계 Latin Hypercube 추출법을 제시하고 모의실험을 통하여 새로운 표본추출법이 기존의 표본추출법들보다 더 효율적임을 보였다.

1. 서 론

구조가 복잡한 장치의 성능을 분석하거나 복잡한 공정을 연구하고자 할 경우에, 성능이나 공정을 나타내는 수학적 모형을 세우고 이 모형을 구현시키기 위해 컴퓨터 프로그램을 사용하게 되는데 이 수학적모형과 컴퓨터 프로그램을 컴퓨터 모델이라 부른다. 컴퓨터 모델을 실행시키므로써 입력변수들의 한 가지 조합으로부터 하나의 결과값을 얻을 수 있는데 대부분의 입력변수는 일정한 오차 범위내에서 변화하게 되므로 결과값의 오차범위를 예측하기 위해서는 입력변수 값들의 모든 가능한 조합에 대해 컴퓨터 모델을 실행시켜야 할 것이다. 그러나 하나의 입력변수 값들의 조합에 대한 결과값을 얻기 위한 1회의 모델 실행에 상당한 비용과 시간이 들기 때문에 입력변수 값들의 모든 조합에서 모델을 실행시킨다는 것은 거의 불가능하다. 따라서 입력변수를 미리 알려진 확률분포를 따르는 확률변수로 가정하고 모든 입력변수 값들의 조합으로 이루어진 모집단으로부터 비교적 작은 크기의 표본을 추출함으로써 결과값의 분포를 추정하는 통계적 기법이 사용되어 왔다.

컴퓨터 모델의 실행을 위한 통계적 기법은 크게 추출된 입력변수들의 표본에 대해 컴퓨터 모델을 실행시켜 직접 결과값의 분포 또는 그 특성치를 추정하는 방법과 컴퓨터 모델의 실행결과로부터 먼저 결과값과 입력변수들간의 함수관계를 추정하고 추정된 함수식을 이용하여 Monte Carlo 방법에 의해 결과값의 분포를 추정하는 방법으로 나눌 수 있다. 전자의 방법은 정확한 문제에 대한 근사적 해법임에 반해 후자의 방법은 근사적 문제에 대한 정확한 해법이라고 볼 수 있는 데 두 방법 모두 효율적인 표본추출방법의 사용이 필수적인 전제가 된다. 효율적으로 표본을 추출하기 위해서는 무작위성을 유지하면서 입력변수들의 값이 전 범위에서 가능한 한 고르게 추출되도록 함이 중요하다.

McKay 등(1979)이 제안한 Latin Hypercube 추출법은 이러한 점에서 기존의 랜덤추출법이나 층화추출법보다 효율적임이 Monte Carlo 모의실험을 통해 밝혀졌고 그 이후 컴퓨터 모델의 실

1) (449-800) 경기도 용인군 용인읍 운학리 산 46-1, 한국통신 용인위성관계소 위성분석실.

2) (305-353) 대전시 유성구 덕진동 150, 한국원자력연구소 핵연료품질관리실.

3) (305-764) 대전시 유성구 궁동 220, 충남대학교 통계학과.

행을 위한 표본추출방법으로 최근까지 가장 널리 사용되어 오고 있다. (Iman 등, 1981a, 1981b 를 참조.) 최근들어 최선영(1993)은 Latin Hypercube 추출법에 층화추출법의 장점을 보완한 층화 Latin Hypercube 추출법을 제안한 바 있는데 본 논문의 목적은 이를 보다 개선한 이단계 Latin Hypercube 추출법을 제시함에 있다.

2. Latin Hypercube 추출법과 층화 Latin Hypercube 추출법

여러 표본추출방법을 비교하기 위해 $X=(X_1, \dots, X_K)$ 는 $S = S_1 \times \dots \times S_K$ 를 support로 갖는 서로 독립인 입력변수들의 집합이고 X 의 미지의 함수로 표시되는 $Y = h(X)$ 는 결과변수이며 관측가능하다고 가정하자. 이제 임의의 함수 g 에 대해 크기 N 인 표본으로부터 구한 $E[g(Y)] = \tau$ 의 추정량 T 를 다음과 같이 정의한다.

$$T(Y_1, \dots, Y_N) = N^{-1} \sum_{i=1}^N g(Y_i)$$

예를 들어, $g(Y) = Y$ 이면 T 는 $E(Y)$ 를 추정하기 위한 표본평균이다.

2.1 Latin Hypercube 추출법

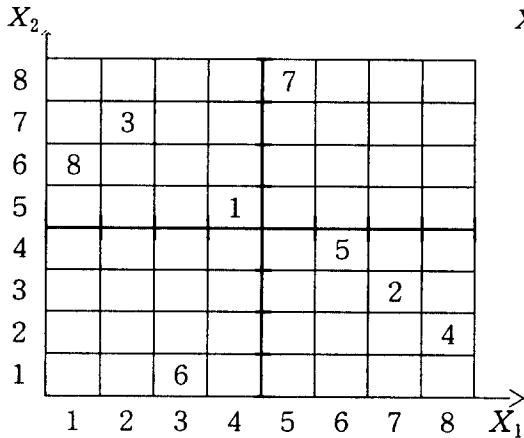
McKay 등(1979)이 제안한 Latin Hypercube 추출법(LHS)은 랜덤 추출법보다 S 전체에서 N 개의 점이 골고루 추출되도록 고안된 방법이다. 즉, 각 S_t , $t = 1, \dots, K$, 를 $1/N$ 의 확률을 갖는 N 개의 등확률구간으로 나누어 S 전체를 N^K 개의 방(cell)으로 만든 다음, 이중 서로 다른 N 개의 방에서 한 점씩 추출하되 뽑힌 N 개의 점이 S_t 의 각 구간에 하나씩 들어가도록 추출하는 방법이다. 이제 각 S_t 의 N 개의 구간에 1에서 N 까지의 색인을 붙이면 임의의 방의 위치는 벡터 (e_1, \dots, e_K) , $e_t \in \{1, \dots, N\}$; $t = 1, \dots, K$, 에 의해 표시할 수 있다. LHS에 의해 뽑힌 N 개의 방의 위치는 K 개의 집합 $\{1, \dots, N\}$ 을 각각 랜덤 순열화한 후, 동일한 위치의 원소들을 순서대로 짝지어 N 개의 K 차원 벡터를 형성하므로써 간단히 결정될 수 있다. 예를들어 두 개의 입력변수가 0과 1사이의 균일분포를 따를 때 크기 8인 표본을 LHS로 추출할 경우, 두 개의 집합 $\{1, \dots, 8\}$ 을 랜덤 순열화하여 $\{4, 7, 2, 8, 6, 3, 5, 1\}$ 과 $\{5, 3, 7, 2, 4, 1, 8, 6\}$ 을 얻고, 두 집합의 원소들을 순서대로 짝지어 추출될 여덟 개의 방의 위치를 결정하면 <그림 1>과 같은 결과를 얻는다. <그림 1>에서 뽑힌 방에 적힌 숫자는 표본이 추출된 순서를 나타내는데 각 행과 열에 해당하는 구간에서는 정확히 하나의 점만이 뽑혔음을 알 수 있다. 이는 입력변수의 값들이 각 주변분포의 전 범위에서 같은 비율로 뽑혔다는 것을 의미한다.

McKay 등(1979)은 LHS에 의해 추출된 표본으로부터 구한 τ 의 추정량 T_L 이 불편추정량임

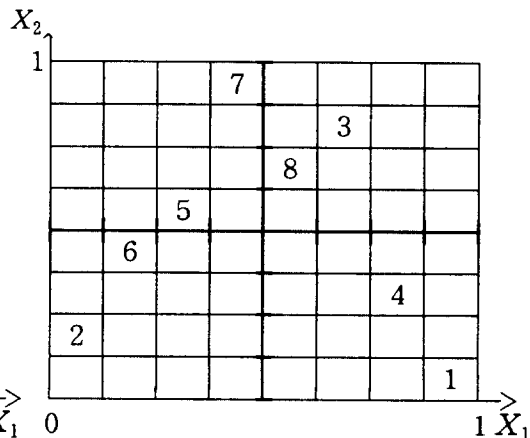
을 보였고, Stein(1987)은 랜덤 추출법에 의해 추출된 표본으로부터 구한 r 의 추정량 T_R 에 대한 T_L 의 점근상대효율이 1 이상이 됨을 보였다.

2.2 층화 Latin Hypercube 추출법

LHS는 S 전체에서 골고루 표본이 추출되도록 고안된 방법이나 경우에 따라서는 표본이 한 편으로 치우쳐 뽑힐 가능성이 있다. 예를 들어 <그림 1>에서 X_1 과 X_2 의 support를 점 (1/2, 1/2)을 중심으로 네 개의 등확률 정사각형 형태의 층으로 나누어 볼 때 두 개의 층에서는 세 점씩 뽑혔으나 나머지 두 개의 층에서는 한 점씩만 뽑혔음을 알 수 있다. 층화 Latin Hypercube 추출법(SLHS)은 이러한 단점을 보완하기 위해 LHS로 표본을 추출하되, S 를 여러 개의 층으로 나누어 각 층에서 같은 수의 점이 뽑히도록 설계된 추출법이다. 즉, 각 S_i , $i = 1, \dots, K$, 를 $1/M$ 의 확률을 갖는 M 개의 등확률구간으로 나누어 S 전체를 M^K 개의 층으로 만든 후, 각 층마다 n 개씩 추출하되 전체 $N = n \cdot M^K$ 개의 점이 LHS에 의한 표본의 성질을 만족시키도록 추출하는 방법이다. <그림 2>는 각각 0과 1사이의 균일분포를 따르는 두 개의 입력변수의 support를 두 부분으로 나누어 네 개의 층을 만들고 각 층에서 두 개씩 모두 8개의 점을 SLHS로 뽑은 예이다. 최선영(1993)은 SLHS에 의해 추출한 표본으로부터 구한 r 의 추정량은 역시 불편추정량이 됨을 보였다.



<그림 1> LHS의 한 예



<그림 2> SLHS의 한 예

3. 이단계 Latin Hypercube 추출법

SLHS는 LHS의 기본 골격을 유지하되 전체 S 를 여러 개의 층으로 나누고 각 층에서 같은 수의 점을 뽑으므로써 LHS보다 더 고른 표본이 추출되도록 하였으나 <그림 2>에서 보는 바와

같이 각 층에서는 점들이 반드시 골고루 선택되지 않을 수도 있음을 알 수 있다. 이단계 Latin Hypercube 추출법(TLHS)은 이러한 단점을 보완하여 점들이 S 전체에서 보다 고르게 뿔히도록 S 를 여러 개의 층으로 나누는 뒤 각 층을 다시 같은 수의 블록으로 나누어 표본을 추출하는 방법이다. 이를 위해 각 S_t , $t = 1, \dots, K$, 를 $1/M$ 의 확률을 갖는 M 개의 등확률구간 S_{t1}, \dots, S_{tM} 으로 나누어 S 전체를 M^K 개의 층으로 만든 후, 다시 각 S_{tu} , $t = 1, \dots, K$; $u = 1, \dots, M$, 를 n 개의 등확률구간 $S_{t u 1}, \dots, S_{t u n}$ 으로 나누어 한 층을 n^K 개의 블록으로 만든다. 그 다음 각 층에서 LHS에 의해 n 개의 블록을 선택하고, 선택된 블록내에 있는 총 $(N/nM)^K$ 개의 방들에서 랜덤하게 한 방을 선택하여 한 점을 추출하되 전체 $N = n \cdot M^K$ 개의 점이 LHS에 의한 표본의 성질을 만족시키도록 추출하게 된다. 이 절차를 보다 상세히 설명하기로 한다.

먼저 층, 블록, 방에 대한 색인 시스템을 정의하기 위해 구간 S_{tu} , $u = 1, \dots, M$, 에 1부터 M 까지의 색인을 붙이고, S_{tu} 를 n 개로 나눈 $S_{t u v}$, $v = 1, \dots, n$, 에 1부터 n 까지의 색인을 붙인 다음, 각 $S_{t u v}$ 를 다시 N/nM 개의 등확률구간 $S_{t u v w}$, $w = 1, \dots, N/nM$,으로 나누어 1부터 N/nM 까지의 색인을 붙인다. 예를 들어 $N = 8$, $M = 2$, $K = 2$, $n = 2$ 일 때의 색인 시스템은 <그림 3>과 같다. 이와같은 색인 시스템하에서 임의의 방 $S_{1 u_1 v_1 w_1} \times \dots \times S_{K u_K v_K w_K}$ 는 색인 벡터 $(u_1, \dots, u_K; v_1, \dots, v_K; w_1, \dots, w_K)$ 로 표시할 수 있다. 여기서 $U = (u_1, \dots, u_K)$ 는 층 벡터, $V = (v_1, \dots, v_K)$ 는 블록 벡터, $W = (w_1, \dots, w_K)$ 는 방 벡터를 나타낸다. 한편 층의 위치를 컴퓨터에 기억시키기 위해서는 층 벡터를 다음과 같이 스칼라로 변환하는 것이 편리하다.

$$p = \sum_{i=1}^K (u_i - 1)M^{K-i} + 1 \quad (3.1)$$

예를 들어 <그림 3>의 경우는 네개의 층 벡터 $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$ 가 각각 1, 2, 3, 4로 변환된다. 이와같은 정의하에서 TLHS 절차를 기술하면 다음과 같다.

- 1) P 를 1부터 M^K 까지의 숫자가 각각 n 개 만큼씩 중복된 집합으로 초기화한다. 또한 집합 R_{lt} , $l = 1, \dots, M^K$; $t = 1, \dots, K$, 를 각각 $\{1, \dots, n\}$ 으로, 집합 $Q_{t u v}$, $t = 1, \dots, K$; $u = 1, \dots, M$; $v = 1, \dots, n$, 를 각각 $\{1, \dots, N/nM\}$ 으로 초기화한다.
- 2) P 에서 한 원소 p 를 랜덤하게 추출하여 식 (3.1)을 만족시키는 층 벡터 $U = (u_1, \dots, u_K)$ 를 구하고 $P = P - \{p\}$ 로 갱신한다.
- 3) $R_{p,t}$, $t = 1, \dots, K$, 에서 각각 한 원소 v_t 를 랜덤하게 추출하여 블록 벡터 $V = (v_1, \dots, v_K)$ 를 구하고 $R_{p,t} = R_{p,t} - \{v_t\}$ 로 갱신한다.

- 4) $Q_{tu,v}, t = 1, \dots, K$, 에서 각각 한 원소 w_t 를 랜덤하게 추출하여 방 벡터 $W = (w_1, \dots, w_K)$ 를 구하고 $Q_{tu,v_i} = Q_{tu,v_i} - \{w_j\}$ 로 갱신한다.
- 5) 선택된 방 U, V, W 에서 랜덤하게 한 점을 선택한다.
- 6) 2 - 5의 절차를 N 회 반복한다.

블럭	1		2		1		2	
방	1	2	1	2	1	2	1	2
층	1				2			

<그림 3> 층, 블럭, 방의 분할법

이상의 절차를 간단한 예를 통하여 살펴 보기로 한다.

예 $N = 8, M = 2, K = 2, n = 2$ 일 경우의 TLHS 절차

- 1) $P = \{1, 1, 2, 2, 3, 3, 4, 4\}$ 가 되고, $R_{it} = \{1, 2\}, i = 1, 2, 3, 4; t = 1, 2$, 그리고 $Q_{tuv} = \{1, 2\}, t = 1, 2; u = 1, 2; v = 1, 2$, 가 된다.
- 2) P 에서 $p = 3$ 이 첫번째로 추출되었다면 $3 = 1 \cdot 2^1 + 0 \cdot 2^0 + 1$ 로 나타낼 수 있으므로 이에 대응되는 층 벡터는 $U = (2, 1)$ 이고 $P = \{1, 1, 2, 2, 3, 4, 4\}$ 로 갱신된다.
- 3) R_{31} 과 R_{32} 에서 모두 2가 추출되었다면 이에 대응되는 블럭 벡터는 $V = (2, 2)$ 가 되고 $R_{31} = \{1\}, R_{32} = \{1\}$ 로 갱신된다.
- 4) Q_{122} 에서 1, Q_{212} 에서 2가 추출되었다면 이에 대응되는 방 벡터는 $W = (1, 2)$ 가 되고 $Q_{122} = \{2\}, Q_{212} = \{1\}$ 로 갱신된다.

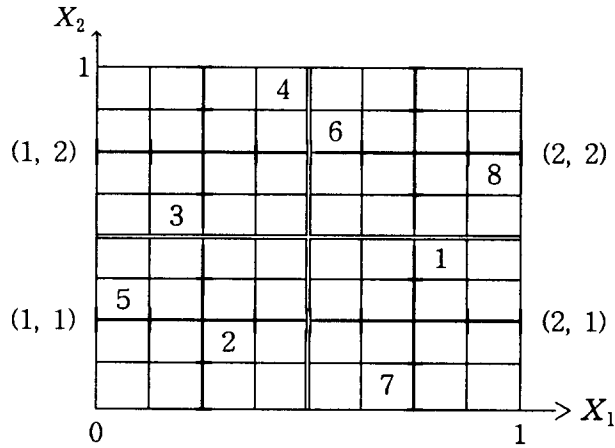
이와같은 방법으로 2 - 4의 절차를 8회 반복할 때 집합 $P, R_{it} = \{1, 2\}, i = 1, 2, 3, 4; t = 1, 2$, 그리고 $Q_{tuv} = \{1, 2\}, t = 1, 2; u = 1, 2; v = 1, 2$, 에서 원소들이 비복원추출에 의해 뽑힌 순서가 다음과 같다고 가정하자.

$$\begin{aligned}
 &P: 3, 1, 2, 2, 1, 4, 3, 4 \\
 &R_{11}: 2, 1, \quad R_{21}: 1, 2, \quad R_{31}: 2, 1, \quad R_{41}: 1, 2 \\
 &R_{12}: 1, 2, \quad R_{22}: 1, 2, \quad R_{32}: 2, 1, \quad R_{42}: 2, 1 \\
 &Q_{111}: 2, 1, \quad Q_{112}: 1, 2, \quad Q_{121}: 1, 2, \quad Q_{122}: 1, 2 \\
 &Q_{211}: 2, 1, \quad Q_{212}: 2, 1, \quad Q_{221}: 1, 2, \quad Q_{222}: 2, 1
 \end{aligned}$$

그러면 다음과 같은 층 벡터, 블록 벡터, 방 벡터의 집합을 각각 얻을 수 있다.

$$\begin{pmatrix} 2 & 1 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 1 & 2 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 1 & 2 & 2 & 2 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 \end{pmatrix}$$

이 결과를 그림으로 나타내면 <그림 4>와 같다.



<그림 4> TLHS의 한 예

한편 TLHS에 의해 추출한 표본으로부터 구한 r 의 추정량은 LHS나 SLHS의 경우와 마찬가지로 불편추정량이 됨을 보일 수 있다. (증명은 임미정, 1994, 을 참조.)

4. 프린터 구동장치에의 적용 사례

이 장에서는 앞에서 기술한 이단계 Latin Hypercube 추출법을 프린터 구동장치 모형에 적용시켜 보기로 한다.

프린터 구동장치의 설계에 있어서 중요한 파라미터는 구동장치의 이동시간(flight time or contact time)과 충격에너지이다. 이동시간은 구동장치가 종이에 닿을 때까지 경과한 시간이며 충격에너지는 종이에 닿은 시점에서의 운동에너지이다. 만약 이동시간의 편차가 크면 프린터의 기록 오차가 심하고 출력질이 좋지 못하므로 이동시간은 일정한 범위내에서 정확하게 유지되어야 한다. 또한 충격에너지가 너무 크면 종지와 리본에 손상을 주고 너무 작으면 프린터의 상(image)이 흐리므로 충격에너지 역시 지정된 범위 내에서 정확하게 유지되어야 한다. 따라서 본 논문에서는 프린터 구동장치의 성능분석을 위한 특성치로서 이동시간과 충격에너지를 사용하여 이들이 기준치로부터 벗어난 정도를 추정하였다.

프린터 구동장치는 영구자석, 전동자(armature), 전자석, 그리고 스프링으로 구성되어 있는데, 전동자의 동작에 관한 수학적 모형으로는 Chen 등(1984)이 제안한 다음과 같은 연립미분방정식을 사용하였다.

$$m(d^2x/dt^2) = Bli - Kx - F_{paper}(x), \quad V = iR + L(di/dt) + B(dx/dt),$$

x : 전동자의 위치, t : 시간, i : 전류, B, K, R, L, m, l : 입력변수

F_{paper} : 전동자가 종이에 닿는 순간에 전동자에 작용한 힘으로 Hendriks(1983)의 다음 식을 인용하였음;

$$F_{paper} = C \left\{ I(dx/dt \geq 0) \left(\frac{x-x_p}{d} \right)^{2.7} + I(dx/dt < 0) \left(\frac{x_{max}}{d} \right)^{-8.3} \left(\frac{x-x_p}{d} \right)^{11} \right\}$$

C : 리본, 종이의 질(stiffness)과 indenter의 면적의 곱에 비례하는 상수

x_p : 전동자가 종이에 닿았을 때의 위치, d : 리본과 종이의 두께

x_{max} : 인쇄 기판(printing substrate)의 최대 깊이(maximum indentation)

<표 1> 입력변수들의 분포

변수	단위	평균	표준편차
R(저항)	ohms	61.3	1.5
L(자기감응계수)	millihenrys	2.92	0.15
m (질량)	grams	0.403	0.03
B(자기유도)	kilogauss	9.67	0.4
K(스프링상수)	newton/meter	117.5	10
l (코일 conductor의 길이)	meters	6.67	0.2

여기서 여섯개의 입력변수들이 서로 독립이고 <표 1>에 있는 평균과 표준편차를 가지는 정규분포를 따른다고 가정하면, 이동시간과 충격에너지 또한 특정한 분포를 따르는 확률변수들이 된다. 이동시간과 충격에너지가 기준치로부터 벗어난 정도를 추정하기 위해 Stein(1987)에서와 마찬가지로 다음 네가지의 함수식의 기대값을 추정하였다.

$$h_{IE}^1(x) = \begin{cases} 0 & \text{if } 12 \leq \text{impact energy} \leq 14 \\ 1 & \text{otherwise} \end{cases}$$

$$h_{IE}^2(x) = (\text{impact energy} - 13)^2$$

$$h_{CT}^1(x) = \begin{cases} 0 & \text{if } .14 \leq \text{contact time} \leq .16 \\ 1 & \text{otherwise} \end{cases}$$

$$h_{CT}^2(x) = (\text{contact time} - .15)^2$$

함수 $h_{IE}^1(x)$ 와 $h_{CT}^1(x)$ 는 베르누이 확률변수이고 $h_{IE}^2(x)$ 와 $h_{CT}^2(x)$ 는 제곱오차가 된다. 앞에서 기술한 표본추출방법들을 사용하여 입력변수들의 분포로부터 128개의 표본을 추출하고 위의 네가지 함수식에 대한 기대값의 추정치들을 표준오차면에서 비교하기 위해 500번의 Monte Carlo 모의실험을 실시한 결과가 <표 2>에 요약되어 있다. 각 표본으로부터 추정치들을 계산하기 위해서는 위에서 주어진 x 와 t 에 관한 연립미분방정식을 풀어야 하는데 이를 위해 IMSL의 서브루틴 IVPAG를 사용하였다.

<표 2>에서 보는 바와 같이 네가지 추출법으로 구한 함수 $h_{IE}^1(x)$, $h_{IE}^2(x)$, $h_{CT}^1(x)$, $h_{CT}^2(x)$ 의 기대값의 추정량은 모두 불편추정량이고 TLHS로 구한 추정량의 표준오차가 다른 추출법에 의한 추정량의 표준오차보다 상당히 작음을 알 수 있다. 특히 함수 $h_{IE}^2(x)$ 와 $h_{CT}^2(x)$ 의 경우에 TLHS로 구한 추정량의 표준오차는 랜덤추출법의 경우보다 두 배 이상 작고 SLHS보다도 많이 감소하였음을 볼 수 있다. 따라서 프린터 구동장치의 성능분석시에 TLHS방법으로 입력변수들을 추출하면 입력변수 각각의 확률분포를 잘 대표하고 추출변동이 작아지므로 충격에너지와 이동시간을 다른 추출법에 비해 보다 정확하게 예측할 수 있을 것이다.

<표 2> 4개 함수의 추정치의 평균과 표준오차

함 수	추출방법	평 균	표준오차
$h_{IE}^1(x)$	RS	1.1172×10^{-1}	2.7174×10^{-2}
	LHS	1.1395×10^{-1}	2.4930×10^{-2}
	SLHS	1.1350×10^{-1}	2.0054×10^{-2}
	TLHS	1.1211×10^{-1}	1.8892×10^{-2}
$h_{IE}^2(x)$	RS	3.9678×10^{-1}	4.9940×10^{-2}
	LHS	3.9750×10^{-1}	3.7338×10^{-2}
	SLHS	3.9287×10^{-1}	2.7288×10^{-2}
	TLHS	3.9582×10^{-1}	2.2927×10^{-2}
$h_{CT}^1(x)$	RS	9.6672×10^{-2}	2.4792×10^{-2}
	LHS	9.9969×10^{-2}	1.9699×10^{-2}
	SLHS	9.7781×10^{-2}	1.9165×10^{-2}
	TLHS	9.8563×10^{-2}	1.7281×10^{-2}
$h_{CT}^2(x)$	RS	3.6979×10^{-5}	4.4413×10^{-6}
	LHS	3.7668×10^{-5}	2.8313×10^{-6}
	SLHS	3.7266×10^{-5}	2.2460×10^{-6}
	TLHS	3.7246×10^{-5}	1.6942×10^{-6}

5. 맺는 말

TLHS는 Monte Carlo 모의실험 결과가 보여 주듯이 추정량의 표준오차를 감소시킨다는 점에서 바람직한 표본추출법이나 표본 크기에 다소 제약을 받는다는 단점이 있다. 즉, M 을 층을 나누기 위한 구간의 수, K 를 입력변수의 수, n 을 각 층에서의 블록의 수라 하면 표본크기 N 이 $N = n \cdot M^K$ 을 만족시켜야 TLHS를 사용할 수 있다. 그러나 N 이 이 관계를 만족시키지 않더라도 $2 \leq Q \leq K$ 인 정수 Q 에 대해 $N = n \cdot M^Q$ 을 만족시킨다면 Q 개의 입력변수는 TLHS를 사용하여 추출하고 나머지 $K - Q$ 개의 입력변수는 LHS를 사용하여 추출하는 방법을 적용할 수 있으므로 실제적으로 큰 제약조건이라고 볼 수는 없을 것이다.

참고문헌

- [1] 임미정 (1994). 이단계 Latin Hypercube 추출법과 그 응용, 석사학위 논문, 충남대학교.
- [2] 최선영 (1993). 층화 Latin Hypercube 추출법과 이를 이용한 연료봉 내압의 위험도 분석, 석사학위 논문, 충남대학교.
- [3] Chen, J.S-C. Wang, S., and Zug, P. (1984). Optimal Tolerance Design of a Printer Actuator, *Transactions of the ASME-Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 106, 510-517.
- [4] Hendriks, R. (1983). Bounce and Chaotic Motion in Impact Printer Hammers, *IBM Journal of Research and Development*, Vol. 27, 273-280.
- [5] Iman, R. L., Helton, J. C. and Campbell, J. E. (1981a). An Approach to Sensitivity Analysis of Computer Models : Part I -Introduction, Input Variable Selection and Preliminary Variable Assessment, *Journal of Quality Technology*, Vol. 13, No. 3, 174-181.
- [6] Iman, R. L., Helton, J. C. and Campbell, J. E. (1981b). An Approach to Sensitivity Analysis of Computer Models : Part II - Ranking of Input Variables, Response Surface Validation, Distribution Effect and Technique Synopsis, *Journal of Quality Technology*, Vol. 13, No. 4, 232-240.
- [7] IMSL (1992). *IMSL Library User's Manual*, IMSL, Inc., Houston, Texas.
- [8] McKay, M. D., Beckman, R. J. and Conover, W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics*, Vol. 21, No. 2, 239-245.
- [9] Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling, *Technometrics*, Vol. 29, No. 2, 143-151.

Two-Stage Latin Hypercube Sampling and Its Application

Mi-Jeong Im⁴⁾, Woo-Joo Kwon⁵⁾, Joo-Ho Lee⁶⁾

Abstract

When modeling a complicated system with a computer model, it is of vital importance to choose input values efficiently. The Latin Hypercube sampling(LHS) proposed by McKay et al.(1979) has been most widely used for choosing input values for a computer model. We propose the two-stage Latin Hypercube sampling(TLHS) which is an improved version of the LHS for producing input values in estimating the expectation of a function of the output variable. The proposed method is applied to simulation study of the performance of a printer actuator and it is shown to outperform the other sampling methods including the LHS in accuracy.

-
- 4) Mission Analysis Department, Yongin Satellite Control Station, Yongin, Kyungki-do, 449-800, KOREA.
 - 5) Fuel Quality Control Department, Korea Atomic Energy Research Institute, 150 Dukjin-dong, Yusong-gu, Daejeon, 305-353, KOREA.
 - 6) Department of Statistics, Chungnam National University, Daejeon, 305-764, KOREA.