

방사성폐기물 처분 종합 안전성 평가에 사용되는 입력 데이터 샘플링 기법 연구

**A Study on the Data Sampling Methods in TSPA for the
Permanent Disposal of Radioactive Waste**

조 명화, 황 용수, 강 철형, 이 연명, 김 성기

한국원자력연구소
대전광역시 유성구 덕진동 150

요약

방사성폐기물 종합 안전성 평가 시에는 관련 기술기준에 의하여 확률론적 안전성 평가가 수행되어야 한다. 종합 안전성 평가 시에 확률밀도함수로 표시되는 입력 데이터 수가 매우 많아 효율적인 샘플링이 이루어지지 않을 경우 확률밀도함수를 정확히 모사하지 못하거나 많은 시간이 소요되게 된다. 현재 종합 안전성 평가 코드로 사용되는 MASCOT-K에서는 Monte Carlo 방법이나 Important Sampling 방법으로 샘플링을 하고 있다. 본 논문에서는 샘플링 기법의 다양화를 위해 Mascot-K에 Latin Hypercube Sampling(LHS) 기법을 적용하기 위한 타당성을 조사해 보았다. LHS의 효율성을 평가하기 위하여 여러 확률분포함수에 관하여 샘플링 숫자를 변화시켜 LHS와 Monte Carlo 방법을 비교하였다. 이 결과는 LHS가 모든 분포에 관하여 적은 샘플링 숫자에서도 주어진 확률분포함수를 잘 모사함을 나타내었다. 따라서 향후 연구에서는 보다 효율적인 상관 관계를 이용한 LHS 방법으로 샘플링하는 프로그램을 개발하기로 한다.

Abstract

It is essential to perform the probabilistic safety assessment to assure the safety of the permanent repository for radioactive waste. However, in practice, it is not quite so straight forward to have effective sampling to describe all input pdf's needed for the safety assessment. It is due to the complexity in the geologic medium as well as characteristics of the waste and its chemical performance. In the MASCOT-K, which is the reference safety assessment tool for the HLW repository in Korea, two useful options for sampling are provided; 1) Monte Carlo Method and 2) Important sampling. To expand its capability, the possibility to include the LHS in the MASCOT-K was tested by comparing its sampling results with those from the Monte Carlo approach. Obviously, the results from the LHS is better than those from the Monte Carlo so that it is feasible to introduce the LHS in the MASCOT-K. In the future, so called LHS with correlation shall be studied as one of the prime sampling methods for the TSPA in KAERI.

1. 서론

한국원자력연구소에서는 고준위 방사성폐기물 영구 처분장 종합 안전성 평가를 위하여 MASCOT-K 프로그램[1]을 개발하고 있다. 중·저준위 방사성 폐기물 처분 안전성을 보장하기 위한 국내 기술기준 법령[2]에서는 종합안전성 평가 시에 확률론적 평가를 수행하도록 명기되어 있다. 따라서 종합안전성 평가 코드인 MASCOT-K 개발에 있어서 확률론적 평가는 필수적일 것으로 예상된다.

고준위 방사성 폐기물의 영구 격리를 위한 심지층 처분장은 지하 500 미터 이하에 위치할 것으로 예상되어 이러한 시설 주변 암반의 물성 특성 등을 정확하게 표현하기는 매우 어려운 일이다. 따라서 처분장 특성을 기술하기 위한 입력 자료들은 대부분의 경우 법에서도 규정한 바와 같이 확률밀도함수(probabilistic density function: PDF)로 표시된다. 그러나 처분장의 특성을 기술하는 이러한 입력 자료들의 수가 매우 많으므로 효율적인 데이터 샘플링 기법을 개발하지 않을 경우 막대한 계산 시간이 요구되거나, 혹은 부정확한 데이터 샘플링 결과를 초래하게 될 것이다. 따라서 MASCOT-K의 개발에서는 보다 효율적인 샘플링 기법 연구가 수반되어야 하겠다.

현재 MASCOT-K에서는 Important sampling[3]과 Monte Carlo 샘플링과 같은 샘플링 기법을 제공하는데, 본 논문에서는 원자력 발전소의 확률론적 안전성 분석 시 많이 사용되는 Latin HyperCube Sampling(LHS)[4]를 MASCOT-K 프로그램에 도입하는 타당성을 예비 분석해 보고자 한다.

2. 본론

처분장에서 사용하는 여러 가지 데이터들은 많이 불확실성을 가지기 때문에 입력모수에 대한 선택이 중요한 의미를 가진다. 따라서 입력모수(input parameter)를 선택할 경우 일반적인 결정 모수보다는 그 데이터를 잘 설명할 수 있는 함수로 확률변수가 취할 수 있는 모든 값에 대응하는 확률을 나타내는 분포인 확률밀도함수를 사용한다. 이러한 PDF를 처분장에서 발생하는 데이터들에 대한 여러 가지 불확실성을 보완하여 고려할 수 있으며 더 적절한 분포에 대한 모수를 구할 수 있다.

처분 연구에서 input data 기술에 흔히 사용되고 있는 분포로는 Normal distribution, Log-Normal distribution, Uniform distribution, Log-Uniform distribution, Triangular distribution 등을 사용하는데 여러 가지 sampling 기법들을 이용하여 데이터를 추출함으로써 분포를 어떻게 잘 재현하는가에 대해서 고찰해 보고자 한다.

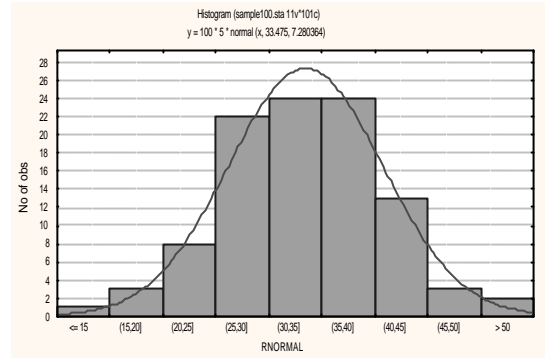
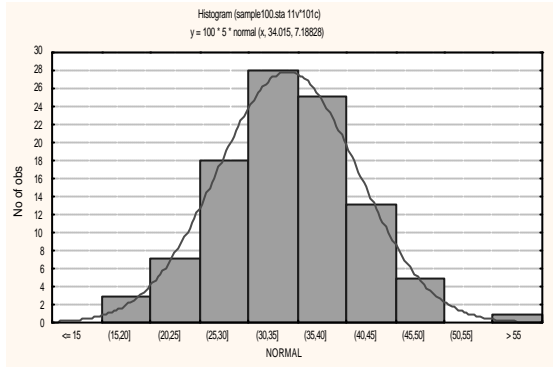
일반적으로 다양한 분포에 대한 sampling을 할 경우는 그 분포의 개수만큼의 multi-dimensional distribution을 얻게 되며 따라서 표본 추출도 차원의 수만큼의 다차원적 표본을 추출하게 된다. 하지만 LHS를 사용함으로써 다차원의 분포에서 추출하는 표본에 대한 차원축소를 유도해 낼 수 있는 것이다. 모집단에서 표본을 추출할 확률이 동일한 Monte Carlo에서는 차원축소를 행할 경우는 자료에 대한 정보가 줄어들 뿐만 아니라 sampling 개수가 증가할 수록 정확한 분포를 유도해 낼 수 있다. 그러나 LHS는 그러한 오차를 최대한 줄이면서 분포를 더 잘 설명할 수 있는 장점을 가지고 있다.

앞에서 언급했듯이 LHS와 Monte Carlo에 의해 추출한 sample들이 각 분포들을 얼마나 잘 설명하는가를 그림 1-3을 통하여 비교하였다. 일반적으로 그 분포를 정확히 묘사하기 위해선 표본의 수를 증가시키는데, 이러한 경우 추출이 어렵거나 시간과 비용이 많이 드는 등의 여러 가지 문제점이 발생하기 때문에 표본의 수를 줄이는 것이 효율적이다. LHS 방

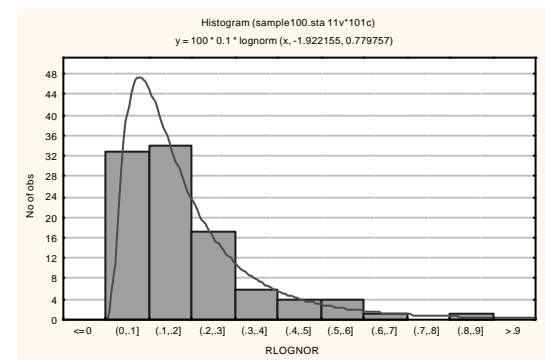
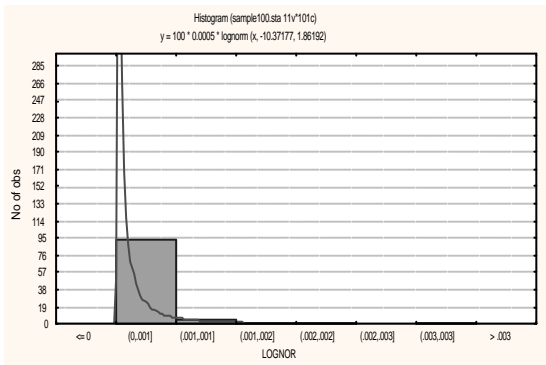
법론을 사용하는데 있어서 가장 큰 장점은, 적은 sample들을 추출하여도 좋은 효과를 얻을 수 있다는 것이다. 아래의 그림들은 sampling 개수를 100, 50, 10으로 변화시킬 경우 LHS와 Monte Carlo 샘플링기법을 사용할 경우, Normal distribution, Log-normal distribution, Uniform distribution, Log-Uniform distribution, Triangular distribution와 같은 다양한 분포들에 관한 샘플링 결과물들의 특성들을 도시한 것이다.

LHS

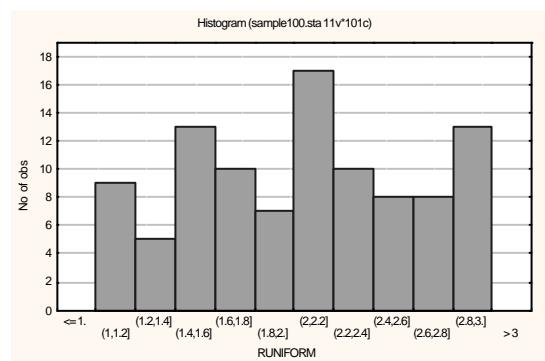
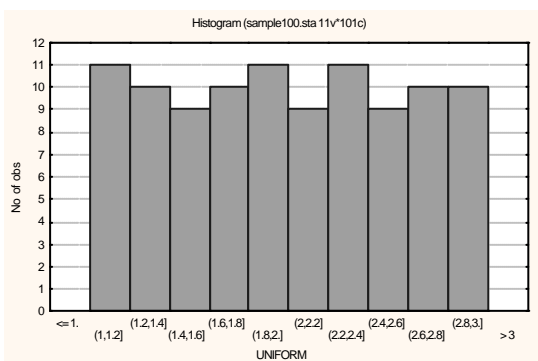
Monte Carlo



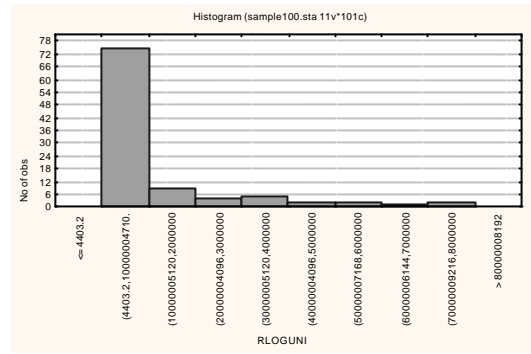
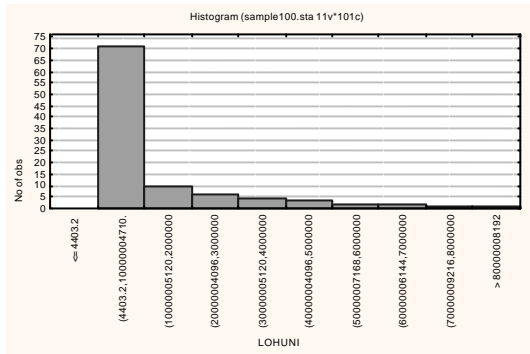
Normal Distribution



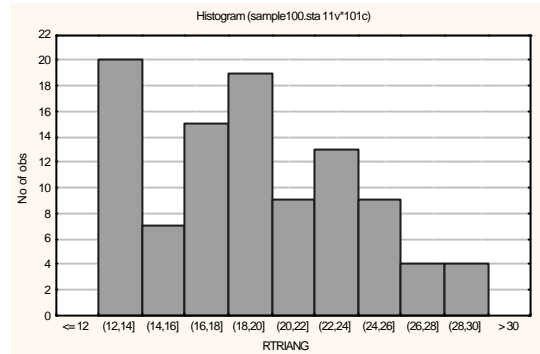
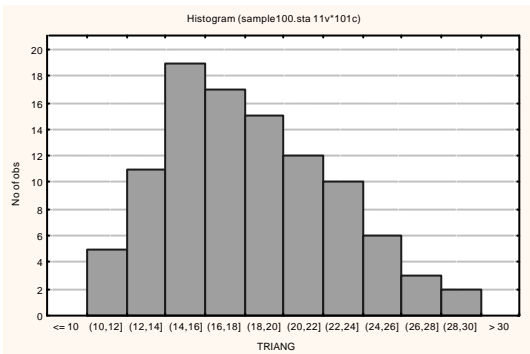
Log-Normal Distribution



Uniform Distribution



Log-Uniform Distribution

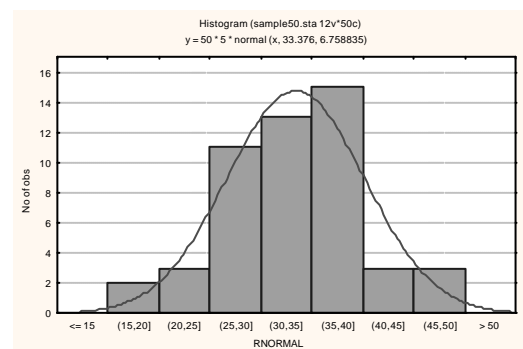
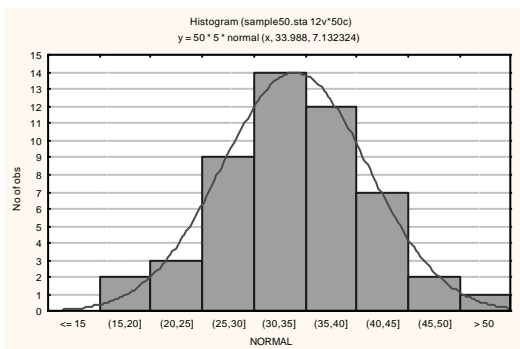


Triangular Distribution

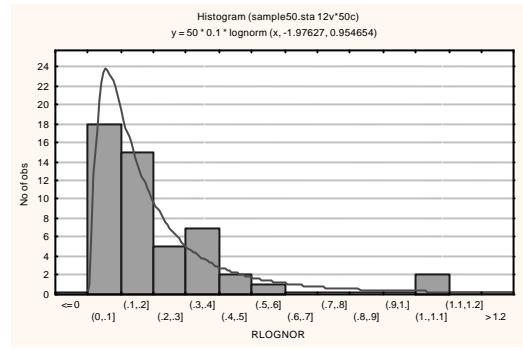
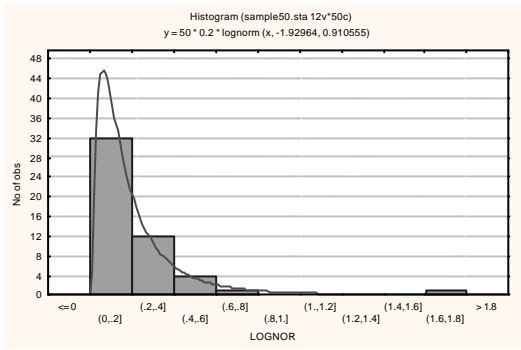
그림 1. Sampling 개수가 100 일 경우 주요 확률밀도분포함수에 대한 LHS와 Monte Carlo 샘플링 결과 분포도

LHS

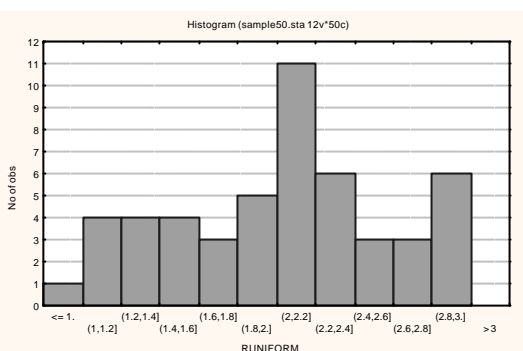
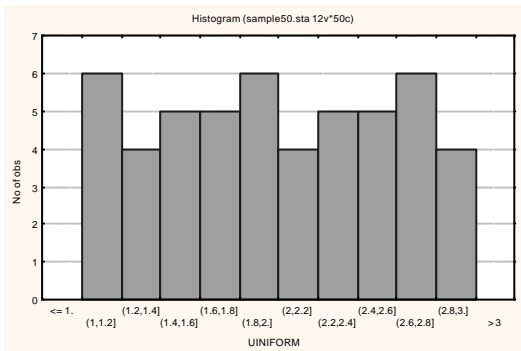
Monte Carlo



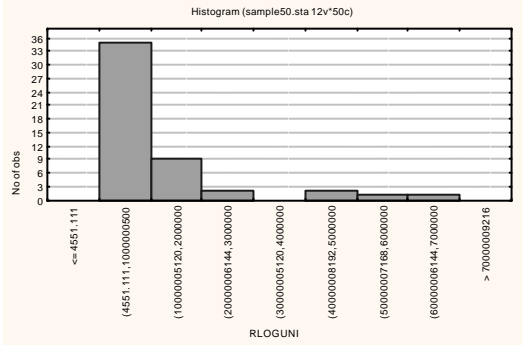
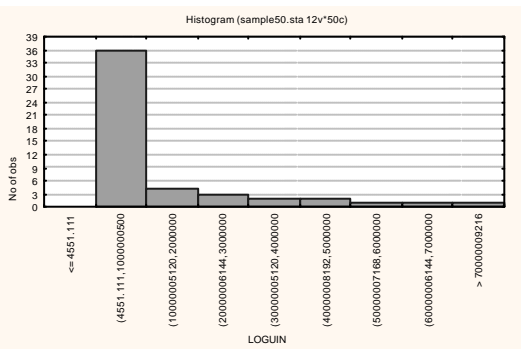
Normal Distribution



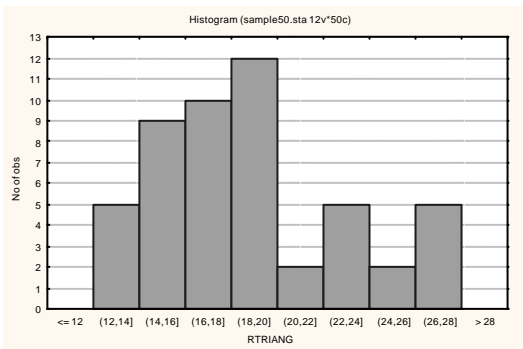
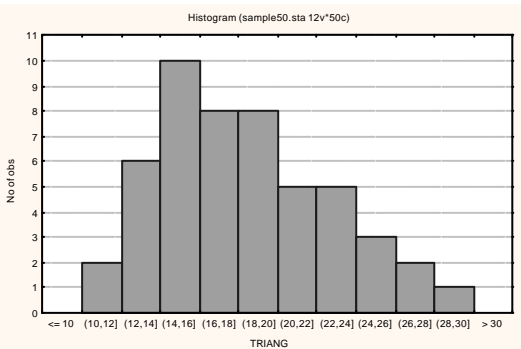
Log-Normal Distribution



Uniform Distribution



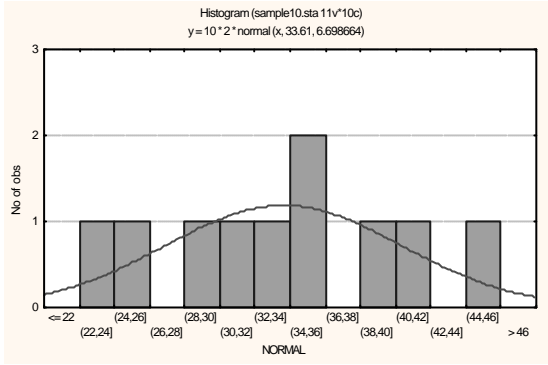
Log-Uniform Distribution



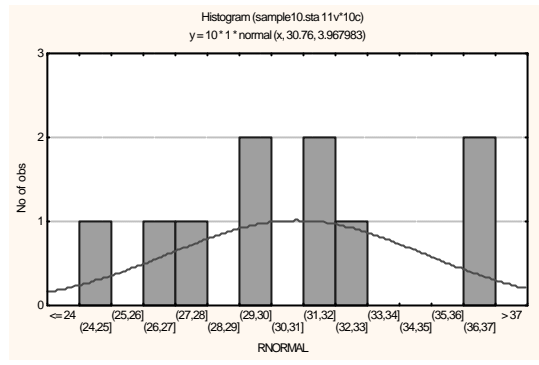
Triangular Distribution

그림 2. Sampling 개수가 50 일 경우 주요 확률밀도분포함수에 대한 LHS와 Monte Carlo 샘플링 결과 분포도

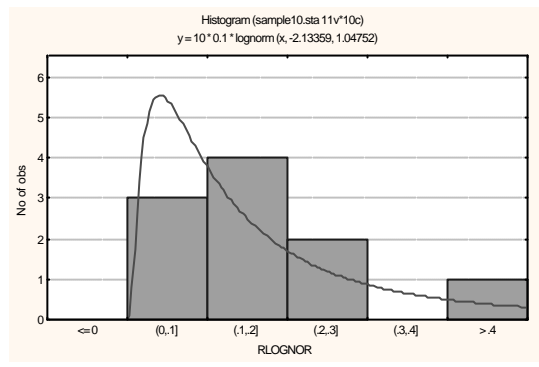
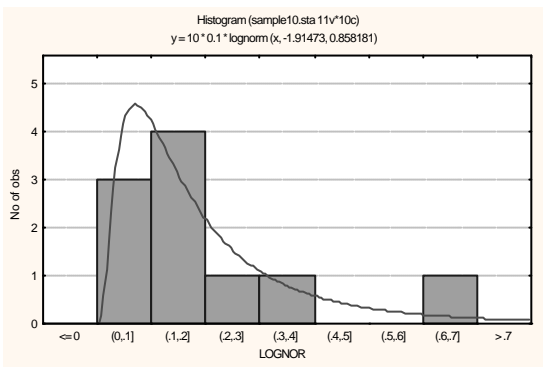
LHS



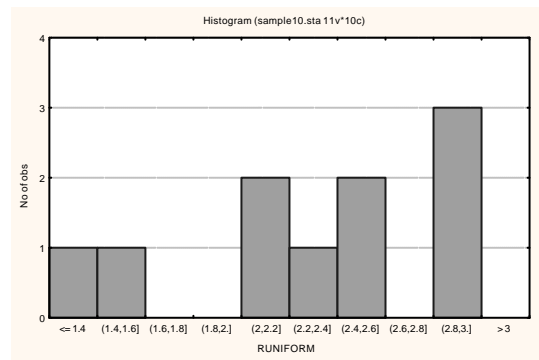
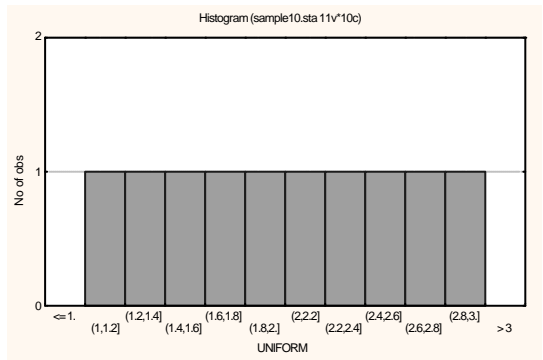
Monte Carlo



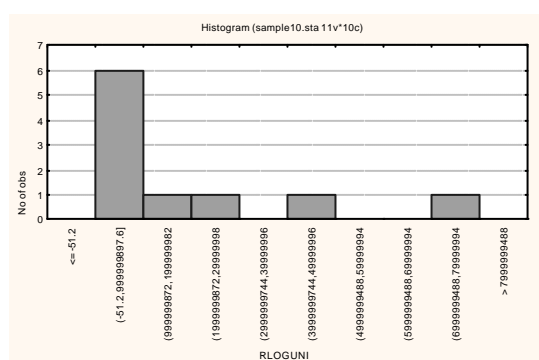
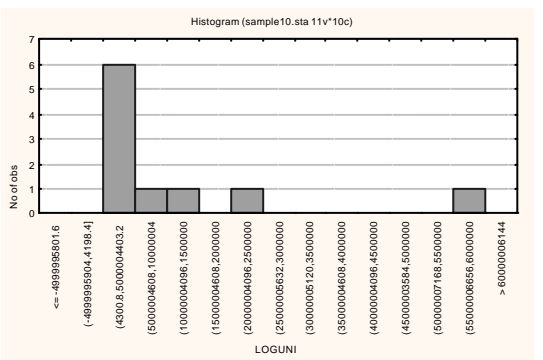
Normal Distribution



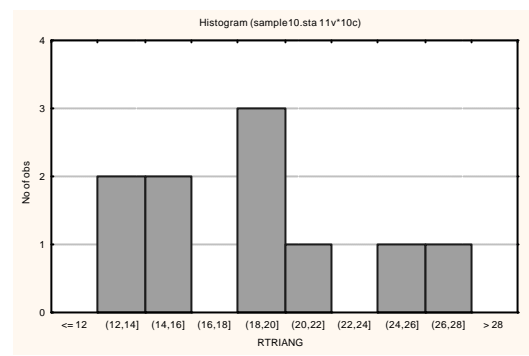
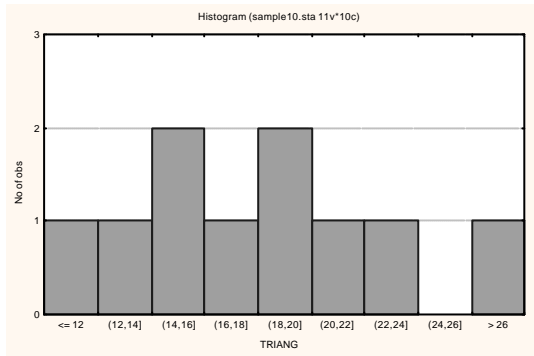
Log-Normal Distribution



Uniform Distribution



Log-Uniform Distribution



Triangular Distribution

그림 3. Sampling 개수가 10 일 경우 주요 확률밀도분포함수에 대한 LHS와 Monte Carlo 샘플링 결과 분포도

3. 결과 및 토의

MASCOT-K가 현재 사용하고 있는 sampling기법은 Monte-Carlo method와 Important Sampling등이 있는데, 여기서는 Stratified Sampling(층화추출)이라는 제약조건을 준 LHS(Latin Hypercube Sampling)라는 기법을 사용하여 좀 더 적은 비용으로 표본을 추출 하여도 분포를 잘 설명함을 그림 1~3에서 도시하였다. 특히 Normal distribution이나 Uniform distribution에서는 LHS 추출법이 Monte Carlo보다 분포를 더 잘 묘사하는 표본을 추출함을 알 수 있으며, 그림 3에서 잘 나타난 바와 같이 LHS은 sampling 개수가 적더라도 확률분포를 Monte Carlo 방법론보다 잘 묘사함을 알 수 있다. 여기서 sampling 개수를 100, 50, 10으로 줄이면 Monte Carlo는 이상점(Outliers)이 발생하여 분포를 적절히 표현하기 못하지만, LHS는 Uniform distribution에서와 같이 sampling 개수가 매우 적은 경우에도 확률밀도함수를 정확하게 표현하는 등 유용성이 매우 높았었다.

따라서 종합 안전성 평가 코드로 사용되고 있는 MASCOT-K에서 Monte Carlo 방법이나 Important Sampling방법의 샘플링 뿐만 아니라, 본 논문의 결과에서 잘 나타난 바와 같이 샘플링 효율성이 높은 Latin Hypercube Sampling(LHS) 기법을 적용하는 것이 타당성이 높음을 알 수 있다.

LHS를 사용함에 있어서 가장 중요한 점은 어떠한 방법으로 구간을 정하는가에 따라서 표본들의 구성이 변화하기 때문에, 각 변수별로 구하고 싶은 표본의 수에 따라 프로그램에서 어떻게 구간을 적절히 나눌 수 있는가에 있다.

본 논문에서는 단순히 LHS를 이용한 sampling의 값만으로 각 분포에 대하여 비교하였지만 다양한 분포들간의 상관관계(Correlate Relation)를 유도하여 보다 효율적인 샘플링을 수행할 수 있는 바 이러한 상관 관계를 이용한 효율적인 샘플링 시에는 분포마다 얻어진 data 값의 차이가 크므로 순위(Rank)를 이용하는 것이 요구된다.

향후 연구에서는 이러한 상관 관계를 이용한 보다 효율적인 LHS 샘플링 기법을 개발하여 MASCOT-K의 부 프로그램으로 개발하여 종합 안전성 평가 시 보다 효율적인 샘플링 기법을 제공하고자 한다.

감사의 글

본 연구는 과학기술부의 원자력연구개발사업의 일환으로 수행되었음

참고 문헌

1. J E Sinclair and P J Agg, "MASCOT and MOP Programs for Probabilistic Safety Assessment, Part A: Overview", Nirex NSS/R336 part A, 1994.
2. 과학기술부, "중저준위 방사성폐기물 처분시설의 성능 등에 관한 규정, 과기부 고시 제 94-2", 과학기술부, 1994.
3. J E Sinclair and P J Agg, "MASCOT and MOP Programs for Probabilistic Safety Assessment, Part C: technical Details", Nirex NSS/R336 part C, 1994.
4. R L Iman and M J Shortencarier, "A Fortran 77 Program and user's Guide for the Generation of LH and RS for Use with Computer Models", NUREG/CR-3624, SAND83-2365, 1984.