시뮬레이션 출력비 추정량의 통계적 분석*

Statistical Analysis of Simulation Output Ratios

홍 윤 기**
Yoon Gee Hong

Abstract

A statistical procedure is developed to estimate the relative difference between two parameters each obtained from either true model or approximate model. Double sample procedure is applied to find the additional number of simulation runs satisfying the preassigned absolute precision of the confidence interval. Two types of parameters, mean and standard deviation, are considered as the performance measures and tried to show the validity of the model by examining both queues and inventory systems. In each system it is assumed that there are three distinct means and their own standard deviations and they form the simultaneous confidence intervals but with control in the sense that the absolute precision for each confidence interval is bounded on the limits with preassigned confidence level. The results of this study may contribute to some situations, for instance, first, we need a statistical method to compare the effectiveness between two alternatives, second, we find the adquate number of replications with any level of absolute precision to avoid the unrealistic cost of running simulation models, third, we are interested in analyzing the standard deviation of the output measure, ... , etc.

서 론

시스템을 연구 또는 개발하는 과정에서 이용되는 도구로서 시뮬레이션이 점차 많은 역할과 비중을 차지하고 있으며 시뮬레이션의 이용도는 계속 증가 추세에 있는 실정이다. 시뮬레이션의 용도는 간단하게 말하자면 관찰 대상이 되는 시스템을 제작 또는 직접 실행하기에 앞서서 고려되는 몇몇 대안 또는 모델들을 놓고 비교·분석을 행하여 최적의 안을 찾아내는데 있다고 하겠다. 이와같은 비교·분석을 뒷받침하기 위한 또 하나의 도구로서 통계적 방법을 적용하므로서 그릇된 결론으로 야기될 수 있는 심각한 실수를 예방하고자 하는데에 그 의미가 있는 것이다.

주어진 일정 시스템에 대하여 가능하다고 인정되는 두

* 이 논문은 1992년도 교육부지원 한국학술진흥재단의 자유공모과제 학술연구조성비에 의하여연구되었음.
** 한성대학교 산업공학과

가지의 서로 다른 모델로부터 출력되는 수행측도의 차이는 흔히 사용되는 가설검정이나 신뢰구간방법을 이용할 수 있다. 여기에는 모집단이 정규분포를 따르는 경우와 비모수 접근방법의 경우로 크게 나누어 볼 수 있겠다.[2] 신뢰구간 방법으로 대표적인 것으로 대응-t(Paired-t) 신뢰구간과 수정된 두 샘플-t (Modified Two-Sample t) 신뢰구간이 있다.[4, 18, 19] 세개 이상의 모델들간의 비교·분석에 대한 연구는 Bonferroni 부등식이 가장 널리 이용되고 있으며 여기에서는 서로 다른 두 모델의 수행측도의 차이에 대한 신뢰구간을 고려 가능한 몇몇 모델들 사이에 동시에 만족하도록 동시신뢰구간을 설정하는 방법도 고려될 수 있다. [10,11,12] 한편, 여러개의 모델 가운데 최적 또는 최고의 모델을 선정하는 방법으로 순위(ranking)를 결정하는 문제나[6,15] 여러가지 중에서 상위 몇개를 선정하는 문제[15]등도 관심있게 연구되어 왔다.

본 연구는 하나의 시스템을 대상으로 서로 다른 두개의 모델을 비교 또는 구분하는데 쓰이는 적절한 방법으로 두 모델로 부터 얻어질 수 있는 관찰 대상이 되는 모수들의 상대차이를 추정하는 통계적 절차를 개발하는 데에 목적이 있다. 관심이 되는 모수들은, 예를 들어서, 평균값, 분산 또는 표준편차, 백 분위수(Percentile), 왜도(Skewness), 그리고 첨도(Kurtosis) 등을 들을 수 있다.

여러가지의 모수 가운데 본 연구에서는 평균치와 표준편차를 관심의 대상으로 선정하였으며 평균치나 표준편차도 여러가지가 함께 출력될 수 있는 시스템을 고려하여 이들을 함께 동시에 고려 할 수 있는 통계적 절차를 개발하도록 구성하였다. 연구의 주 내용은 첫째, 두 평균치간의 상대차이에 대한 신뢰구간을 설정하는 통계적 방법을 개발하고, 둘째, 두 표준편차간의 상대차이에 대한 신뢰구간을 설정하는 방법을 개발하는데 있으며, 셋째는, 동시신뢰구간의 설정을 위한 절차를 소개하고, 마지막으로 상대차이의 신뢰구간이 갖는 절대정도가 주어진 정도에 만족하도록 필요한 반복시행회수의 발견을 위한 계산절차를 개발하도록 구성되어 있다. 또한 통계적 절차의 확인을 위한 작업으로 대기행렬과 재고관리와 같이 확률과정을 따르는 시스템을 선정하여 실제로 타당성과 효율성을 연구·검토하도록 한다.

상대차이의 추정

θ를 고려하고자 하는 모수라 가정하고 서로 다른 두 모델(모델 1, 모델 2)의 상대차이를 나타내는 d를 다음과 같이 정의 하자.

$$d = \frac{\theta_1 - \theta_2}{\theta_1},$$

여기서 θ_1 = 모델 1의 모수값,
θ_2 = 모델 2의 모수값.
그러면

$$d = 1 - \frac{\theta_2}{\theta_1} = 1 - r$$

이 되고 만일 두 모수의 비 r의 $(1-\alpha)$ x 100 % 신뢰구간을 구할 수 있다면, 즉,

$$P[l \le \frac{\theta_2}{\theta_1} \le u] = 1-\alpha$$

를 만족시키는 l과 u를 산출할 수 있다면

$$P[1-u \le 1-\frac{\theta_2}{\theta_1} \le 1-l] = 1-\alpha$$

이 되어 두 모수의 상대차이에 대한 신뢰구간을 얻을 수 있다.

일반적으로 서로 다른 두 모델을 대상으로 연구하는 문제는 흔히 실제로 가설화 된 모델(Hypothesized Model 또는 Actual Model)과 기존 또는 근사모델(Approximate Model)간의 비교·분석에 관한 것들 이므로 편의상 θ_1 은 θ_T로 θ_2는 θ_A로 바꾸어 다음과 같이 표기하자.

$$d = \frac{\theta_T - \theta_A}{\theta_T} = 1 - \frac{\theta_A}{\theta_T} .$$

미리 정해진 절대정도(Absolute Precision)수준 ϵ를 만족하는 상대차이의 신뢰구간을 얻는 것이 이 연구의 첫번째 목적이라 하겠다. 여기서 절대정도는 신뢰구간의 절반에 해당하는 길이를 나타낸다. 그러므로, 포괄적인 의미의 상대차이 d는 $(1-\alpha)$의 신뢰수준으로 중간점으로 점추정치 $\hat{d}$을 갖는 신뢰간격 $l(d)$의 신뢰구간으로 추정될 수 있다. 그러면, 만약 $l(d)/2 \leqq \epsilon$이면, 우리는 대략 $(1-\alpha)$의 확률을

가지고 상대차이 d가 $(100\times\varepsilon)$ % 이내에 있다고 말할 수 있다.

θ_A/θ_T 의 신뢰구간(l,u)를 얻어낸다는 것은 바로 비추정(Ratio Estimation)의 문제와 직결하게 된다. 먼저 두 시뮬레이션 모델로 독립시행하여 얻어낸 데이터를 이용하여 모수 θ_T와 θ_A의 추정치를 찾아낸다. 이 비추정 방법은 Fieller 방법의 응용이라 할 수 있겠으며 참고문헌 [7]에서 찾아볼 수 있다.

평균치의 상대차이

μ_X와 μ_Y를 각각 근사모델과 실제모델의 평균치라 하고 $X_i, i=1,2,\ldots,n$,는 근사모델로 부터 얻어낸 서로는 독립이고 동일한 분포를 갖는 관측치이라 하고 $Y_i, i=1,2,\ldots,n$,를 실제모델로 부터 얻어낸 역시 서로는 독립이고 동일한 분포를 갖는 관측치이라 하자. Fieller방법을 이용한 $r=\mu_X/\mu_Y$의 신뢰구간에 대하여 논하여 보자. 변수

$$\overline{X}(n)-r\overline{Y}(n)$$

을 정의하여 보자. 여기서 $\overline{X}(n)$은 n개의 관측치에 대한 확률변수 X의 표본평균을 말하고 $\overline{Y}(n)$ 역시 확률변수 Y의 표본평균이다. n을 충분히 크게 가정하면 아래의 변수는 점근적으로 표준정규분포를 나타내게 된다.

$$\frac{\overline{X}(n)-r\overline{Y}(n)}{Var\left[(\overline{X}(n)-r\overline{Y}(n)\right]^{1/2}} \rightarrow N(0,1).$$

X_i와 $Y_i, i=1,2,\ldots,n$,는 서로 독립이므로

$$Var[\overline{X}(n)-r\overline{Y}(n)]=Var[\overline{X}(n)]+r^2 Var[\overline{Y}(n)]$$

이며, 이때 우변의 추정치를 각각

$$\hat{Var}[\overline{X}(n)]=\frac{1}{n}S_X^2(n),\ S_X^2(n)=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}(n))^2,$$

$$\hat{Var}[\overline{Y}(n)]=\frac{1}{n}S_Y^2(n),\ S_Y^2(n)=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\overline{Y}(n))^2$$

로 하여 이들을 이용하면, 다음의 근사식이 성립한다.

$$P\left(-Z_{\frac{\alpha}{2}}\le\frac{\overline{X}(n)-r\overline{Y}(n)}{\left(\frac{1}{n}S_X^2(n)+r^2\frac{1}{n}S_Y^2(n)\right)^{1/2}}\le Z_{\frac{\alpha}{2}}\right)\approx 1-\alpha. \qquad (1)$$

윗 식으로 부터

$$\frac{\left(\overline{X}(n)-r\overline{Y}(n)\right)^2}{\frac{1}{n}S_X^2(n)+r^2\frac{1}{n}S_Y^2(n)}\le Z^2_{\frac{\alpha}{2}},$$

이는 다시

$$\left(\overline{Y}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_Y^2(n)\right)r^2-2\overline{X}(n)\overline{Y}(n)\,r+\left(\overline{X}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_X^2(n)\right)\le 0 \qquad (2)$$

로 바꾸어 쓸 수 있다. 식 (2) 좌변의 2차방정식의 두 근은 다음과 같이 되며,

$$r_i=\frac{\overline{X}(n)\overline{Y}(n)+(-1)^i\sqrt{V(n)}}{\overline{Y}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_Y^2(n)},\quad i=1,2, \qquad (3)$$

여기서

$$V(n)=\overline{X}(n)^2\overline{Y}(n)^2-\left(\overline{Y}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_Y^2(n)\right)\left(\overline{X}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_X^2(n)\right).$$

만약 충분히 큰값 n에 대하여 다음의 두조건

$$\overline{Y}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_Y^2(n)>0 \quad \text{과} \quad \overline{X}(n)^2-Z^2_{\frac{\alpha}{2}}\frac{1}{n}S_X^2(n)>0$$

이 만족되면 $V(n)$은 양의 실수가 되어 식 (1)은 아래와 같이 쓸 수 있게 된다.

$$P[r_1\le r\le r_2]\approx 1-\alpha.$$

이는 다시

$$P[1-r_2\le d\le 1-r_1]\approx 1-\alpha \qquad (4)$$

로서 상대차이 d는 하한으로 $(1-r_2)$을 상한으로는 $(1-r_1)$을 갖는 신뢰구간으로 추정될 수 있다. 그리고 구간의

중간값은

$$\hat{d} = 1 - \frac{\overline{X}(n)\overline{Y}(n)}{\overline{Y}(n)^2 - Z^2_{\frac{\alpha}{2}}\frac{1}{n}S^2_Y(n)} \approx 1 - \frac{\overline{X}(n)}{\overline{Y}(n)}$$

과 같이 된다. 이때 신뢰간격의 절반(Half of the Width)은

$$\frac{\sqrt{V(n)}}{\overline{Y}(n)^2 - Z^2_{\frac{\alpha}{2}}\frac{1}{n}S^2_Y(n)} \tag{5}$$

이 된다.

이제 미리 지정 해 놓은 절대정도를 만족하는 신뢰구간을 찾기 위해 이종추출절차를 적용하여 보자. 먼저 초기 표본크기로서 n_0 만큼의 시뮬레이션 시행을 하고 식(5)를 이용하여 신뢰구간의 절반을 계산하면 바로 이것이 현 시점에서의 절대정도가 된다. 만약 간격의 절반이 ε보다 작거나 같으면 이 시점에서 더 이상의 시뮬레이션을 중단하고 요구되는 절대정도에 만족하다는 결론을 내린다. 그러나 구간의 절반이 ε보다 크면 n_0보다 많은 시행회수(n)가 요구되어 이를 찾는 작업에 착수하게 된다. 여기서 몇가지 가정해 두어야 할 것은 n_0보다 추가로 요구되는 시행회수(n)로 부터 얻어내는 μ_X, μ_Y, σ_X^2, 그리고 σ_Y^2의 추정치들은 n_0때 보다 그다지 크게 변화하지 않고 거의 같은 수준일 것 이라는 점이다. 그래서 다음의 관계식을 만족하는 가장 작은 n을 찾아 보기로 한다.

$$\frac{\sqrt{V(n,n_0)}}{\overline{Y}(n_0)^2 - Z^2_{\frac{\alpha}{2}}\frac{1}{n}S^2_Y(n_0)} \le \varepsilon \tag{6}$$

여기서

$$V(n,n_0) = \overline{X}(n_0)^2\overline{Y}(n_0)^2 - \left(\overline{Y}(n_0)^2 - Z^2_{\frac{\alpha}{2}}\frac{1}{n}S^2_Y(n_0)\right)\left(\overline{X}(n_0)^2 - Z^2_{\frac{\alpha}{2}}\frac{1}{n}S^2_X(n_0)\right).$$

식(6)의 좌변은 단지 식 (5)의 절반간격을 n과 n_0를 이용하여 표현한 식에 불과하다. 식(6)의 n에 관한 방정식을 풀어서 구해낸 해 가운데 큰 양의 수를 택하여 이를 다시 그 수 보다는 크고 가장 가까운 정수로 대치하여 n으로 사용한다. 두근 가운데 큰 양의 해는

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

이며, 이때

$a = \varepsilon^2 Y^4(n_0)$,

$b = -Z^2_{\frac{\alpha}{2}}\left(\overline{X}(n_0)^2 S^2_Y(n_0) + \overline{Y}(n_0)^2 S^2_X(n_0) + 2\varepsilon^2 + \overline{Y}(n_0)^2 S^2_Y(n_0)\right)$,

$c = Z^4_{\frac{\alpha}{2}} S^2_Y(n_0)\left(S^2_X(n_0) + \varepsilon^2 S^2_Y(n_0)\right)$.

이제 식(4)를 이용하여 상대차이 d에 대한 신뢰구간을 구하면 모든 n개의 표본들은 포함확률(Coverage Probability) $1-\alpha$ 와 절대정도 ε을 공히 만족하게 된다.

한편, 기존 또는 근사모델의 모수값 μ_X가 정확한 값을 가지는 경우에는 지금까지 사용하였던 모든 식에서 $\overline{X}(n)$의 표본분산을 $Var\ [\overline{X}(n)] = S^2_X(n)/n = 0$ 으로 해서 사용하면 된다. 한편, 실제모델의 모수값 μ_Y에 대한 신뢰구간

$$\left(\overline{Y}(n) - Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}},\ \overline{Y}(n) + Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}}\right)$$

이용하여

$$P\left(\frac{\mu_X}{\overline{Y}(n) + Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}}} \le \frac{\mu_X}{\mu_Y} \le \frac{\mu_X}{\overline{Y}(n) - Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}}}\right) \approx 1-\alpha \tag{7}$$

와 같이 표현된다. 여기서 요구되는 조건으로는

$$\overline{Y}(n) - Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}} > 0$$

인데, 양의 실수를 모수로 갖고 충분한 크기의 n을 고려할때 이 조건은 만족된다고 볼 수 있다. 식 (7)의 부등호 양변에 각각

$$\overline{Y}(n) - Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}} > 0 \text{ 과 } \overline{Y}(n) + Z_{\frac{\alpha}{2}}\frac{S_Y(n)}{\sqrt{n}} > 0$$

을 곱하여 주면 $r = \mu_X/\mu_Y$에 대한 신뢰구간의 하한과 상한으로 각각

$$\frac{\mu_X\overline{Y}(n) - Z_{\frac{\alpha}{2}}\mu_X\frac{S_Y(n)}{\sqrt{n}}}{\overline{Y}(n)^2 - Z^2_{\frac{\alpha}{2}}\frac{S^2_Y(n)}{n}},\quad \frac{\mu_X\overline{Y}(n) + Z_{\frac{\alpha}{2}}\mu_X\frac{S_Y(n)}{\sqrt{n}}}{\overline{Y}(n)^2 - Z^2_{\frac{\alpha}{2}}\frac{S^2_Y(n)}{n}}$$

으로 표현되며 이는 앞에서 설명한 r의 신뢰구간에 $\bar{X}(n)$ 대신 μ_X를 그리고 $S_X^2(n)/n=0$ 으로 놓은 결과와 일치 한다.

표준편차의 상대차이

평균값이 아닌 표준편차의 상대차이에 대해서 고려하여 보기로 하자. 두개의 분산 또는 표준편차의 비에 대한 구간추정치는 F분포를 이용하여 구할 수 있으나 이때에는 모집단의 정규성을 인정하여야 하는 제약이 뒤따르게 된다. 그러나 정규성을 지니지 않은 경우에 표준편차의 상대차이에 대한 구간추정치에 관한 연구는 저자로서는 아직 찾아볼 수 없으며, 이와같은 연구는 매우 중요하고 의미있는 정보를 제공하여 준다고 하는 점에서 소홀히 다루어져서는 아니되겠다. 여기에서도 앞서 연구에서와 유사하게 다음과 같은 변수를 정의하자.

$$S_X^2(n) \quad - \quad r^2 S_Y^2(n),$$

여기서 $S_X^2(n)$ 와 $S_Y^2(n)$은 각각 σ_X^2와 σ_Y^2의 불편추정치이며 $r^2=\sigma_X^2/\sigma_Y^2$이라고 하고 여기에서도 전과 마찬가지로 충분히 큰값 n에 대하여 다음의 관계가 성립하게 된다.

$$\frac{S_X^2(n)-r^2S_Y^2(n)}{Var[(S_X^2(n)-r^2S_Y^2(n)]^{\frac{1}{2}}} \rightarrow N(0,1).$$

또한 X_i와 Y_i는 서로 독립이므로 윗식의 분모는 추정치를 이용하여 다음과 같이 쓸 수 있다.

$$\widehat{Var}[S_X^2(n)-r^2S_Y^2(n)]=\widehat{Var}[S_X^2(n)]+r^4\widehat{Var}[S_Y^2(n)]$$

여기서

$$\widehat{Var}\,[S_X^2(n)]=S_X^4(n)\left(\frac{2}{n-1}+\frac{\hat{\gamma}_X(n)}{n}\right)=S_X^4(n)\,d_X^2(n),$$

$$\widehat{Var}\,[S_Y^2(n)]=S_Y^4(n)\left(\frac{2}{n-1}+\frac{\hat{\gamma}_Y(n)}{n}\right)=S_Y^4(n)\,d_Y^2(n),$$

이며$\hat{\gamma}_X(n)$ 과 $\hat{\gamma}_Y(n)$는 각각 변수 X와 Y의 첨도(Kurtosis)의 추정치이며 윗식의 함수 $d^2(n)$은 큰괄호안의 표현식을 대표하고 있다.

다음의 포함확률(Coverage Probability)을 고려하여 보자.

$$P\left(-Z_{\frac{\alpha}{2}} \le \frac{S_X^2(n)-r^2S_Y^2(n)}{\left(\widehat{Var}[S_X^2(n)]+r^4\widehat{Var}[S_Y^2(n)]\right)^{\frac{1}{2}}} \le Z_{\frac{\alpha}{2}}\right)\approx 1-\alpha, \tag{8}$$

이를 다시 정리하여 쓰면 다음과 같으며

$$P\left[\frac{\left(S_X^2(n)-r^2S_Y^2(n)\right)^2}{\widehat{Var}[S_X^2(n)]+r^4\widehat{Var}\,[S_Y^2(n)]} \le Z_{\frac{\alpha}{2}}^2\right]\approx 1-\alpha,$$

이때r^2의 신뢰구간의 상한과 하한은 아래에 주어진 r^2에 대한 2차 방정식의 두근을 구하여 얻을 수 있게 된다.

$$\left(S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\ \widehat{Var}[S_Y^2(n)]\right)r^4-2\,S_X^2(n)S_Y^2(n)r^2$$
$$+\left(S_X^4(n)-Z_{\frac{\alpha}{2}}^2\widehat{Var}[S_X^2(n)]\right)=0.$$

두근을 표현하여 보면 다음과 같이 표현된다.

$$r_i^2=\frac{S_X^2(n)\,S_Y^2(n)+(-1)^i\sqrt{W(n)}}{S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\,\widehat{Var}[S_Y^2(n)]},i=1,2,$$

$$W(n)=S_X^4(n)S_Y^4(n)-\left(S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\,\widehat{Var}[S_Y^2(n)]\right)\left(S_X^4(n)-Z_{\frac{\alpha}{2}}^2\,\widehat{Var}[S_X^2(n)]\right).$$

평균값의 상대차이에서와 같은 조건을 적용하자면, 즉, 여기서 사용되는 반복시행회수 n이 다음의 두조건을 만족한다고 하면

$$S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\,\widehat{Var}[S_Y^2(n)]>0,\ S_X^4(n)-Z_{\frac{\alpha}{2}}^2\,\widehat{Var}[S_X^2(n)]>0,$$

$W(n)>0$ 이 성립하고 식 (8)은 간단하게 아래와 같이 쓸 수 있게 된다.

$$P[\ r_1^2\le r^2\le r_2^2\ ]\approx 1-\alpha$$

이는 다시

$$P[1-r_2\le d\le 1-r_1]\approx 1-\alpha$$

이 되며 이는 곧 상대차이d는 하한으로$(1-r_2)$와 상한으로 $(1-r_1)$을 갖는 신뢰구간을 형성함을 알 수 있다. 또한 일반적으로 사용하는 충분히 큰 n에 대해서$Var[S_X^2(n)]$ 과

$Var[S_Y^2(n)]$ 이 대략 0의 값을 갖는다는 사실을 이용하면 신뢰구간의 중간점(midpoint)은

$$\frac{1}{2}[(1-r_2)+(1-r_1)]=1-\frac{S_X(n)}{S_Y(n)}$$

이 되고 신뢰구간의 절반에 해당하는 길이를 HW(Half-Width)라 했을때 $\mathrm{HW}=\frac{1}{(r_2-r_1)}$ 이 된다.

이종추출절차(Double Sampling Procedure)를 적용하기 위하여 우선 시뮬레이션의 초기 반복시행회수를 n_0라고 하였을때 얻어지는 신뢰구간의 절반에 해당하는 길이는

$$HW(n_0)=\frac{1}{2}(r_2(n_0)-r_1(n_0))$$

이 되며 만약 $HW(n_0)\leq\varepsilon$ 이면 요구에 충분히 만족되므로 더 이상의 추출 또는 시뮬레이션을 시행할 필요가 없게된다. 반면에 $HW(n_0)\rangle\varepsilon$인 경우에는 추가로 소요되는 시뮬레이션 반복시행회수를 구하기 위하여 다음의 표현식으로부터 n을 구하여 n_0보다 큰 가장 가까운 정수를 구한다.

$$HW(n,n_0)=\frac{1}{2}(r_2(n,n_0)-r_1(n,n_0))=\varepsilon \tag{9}$$

위의 식 (9)에 필요한 두개의 근은 아래의 관계식으로부터 구하여질 수 있다.

$$r_i(n,n_0)=\left(\frac{S_X^2(n_0)\,S_Y^2(n_0)+(-1)^i\sqrt{W(n,n_0)}}{S_Y^4(n_0)-Z_{\frac{\alpha}{2}}^2\,S_Y^4(n_0)\,\phi_Y^2(n,n_0)}\right)^{\frac{1}{2}},\quad i=1,2,$$

$$W(n,n_0)=S_X^4(n_0)S_Y^4(n_0)-\left(S_Y^4(n_0)-Z_{\frac{\alpha}{2}}^2\,S_Y^4(n_0)\,\phi_Y^2(n,n_0)\right)$$

$$\left(S_X^4(n_0)-Z_{\frac{\alpha}{2}}^2\,S_X^4(n_0)\,\phi_X^2(n,n_0)\right),$$

$$\phi_X^2(n,n_0)=\frac{2}{n-1}+\frac{\hat{\gamma}_X(n_0)}{n},\ \phi_Y^2(n,n_0)=\frac{2}{n-1}+\frac{\hat{\gamma}_Y(n_0)}{n}. \tag{10}$$

계산을 돕기 위하여 식 (10)에서 n-1 대신 n을 사용하여 n에 대한 2차 방정식의 두근 가운데 근의 공식을 이용하면

$$\frac{-b+\sqrt{b^2-4ac}}{2a}$$

으로 구하여지며, 이때 a, b, 그리고 c 는 각각 아래와 같이 된다.

$$a=4\,\varepsilon^2\left(\varepsilon^2\,S_Y^8(n_0)-S_X^2(n_0)S_Y^6(n_0)\right),$$

$$b=-Z_{\frac{\alpha}{2}}^2\,[S_X^4(n_0)S_Y^4(n_0)\left(4+\hat{\gamma}_X(n_0)+\hat{\gamma}_Y(n_0)\right)$$

$$+4S_X^2(n_0)S_Y^6(n_0)\,\varepsilon^2(2+\hat{\gamma}_Y(n_0))$$

$$-8S_Y^8(n_0)\varepsilon^4\left(2+\hat{\gamma}_Y(n_0)\right)],$$

$$c=Z_{\frac{\alpha}{2}}^4\,[4\,S_Y^8(n_0)\varepsilon^4\left(2+\hat{\gamma}_Y(n_0)\right)^2-S_X^4(n_0)S_Y^4(n_0)$$

$$\left(2+\hat{\gamma}_X(n_0)\right)\left(2+\hat{\gamma}_Y(n_0)\right)].$$

평균값의 상대차이에서와 마찬가지로 만약 σ_X^2이 알려진 경우에는 $S_X^{2(n)}$ 대신에 σ_X^2를 그대로 대입하고 $Var[S_X^2(n)]=0$으로 고정시켜서 위의 식을 취급하면 된다. 또한 여기서 취급하는 반복시행회수와 양의 실수인 모수값을 감안하면

$$S_Y^2(n)-Z_{\frac{\alpha}{2}}\left(Var[S_Y^2(n)]\right)^{\frac{1}{2}}>0$$

이 되며 이 경우의 σ_X/σ_Y에 관한 $100(1-\alpha)$%신뢰구간은 하한과 상한이 각각 다음과 같이 표현된다.

$$\frac{\sigma_X^2\left(S_Y^2(n)-Z_{\frac{\alpha}{2}}\left(\hat{Var}[S_Y^2(n)]\right)^{\frac{1}{2}}\right)}{S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\hat{Var}[S_Y^2(n)]},$$

$$\frac{\sigma_X^2\left(S_Y^2(n)+Z_{\frac{\alpha}{2}}\left(\hat{Var}[S_Y^2(n)]\right)^{\frac{1}{2}}\right)}{S_Y^4(n)-Z_{\frac{\alpha}{2}}^2\hat{Var}[S_Y^2(n)]}$$

상대차이의 동시신뢰구간

시스템을 연구 또는 개발하는데 시뮬레이션 모델을 이용하였을때 얻어지는 출력에는 단순히 하나의 모수(Parameter)에 관하여만 분석하는 경우도 있지만 경우에 따라서는 여러가지의 다른 모수들을 함께 관심을 가지고 분석을 시도하게 된다. 일반적으로 종료 시뮬레이션

(Terminating Simulation) 또는 안정상태 시뮬레이션(Non-terminating Simulation) 공히 여러개의 모수들을 함께 고려하게 되는 것이 일반적이다. 예를 들어서, 대기행렬의 경우에 관심이 될 수 있는 모수들은 고객의 평균 대기시간, 평균 고객의 수, 또는 시스템의 평균 이용률 등을 들을 수 있으며 재고관리 시스템의 경우에는 평균 주문비용, 평균 유지비용, 그리고 평균 품절비용 등을 들을 수 있다. 여기서 이들 서로 다른 모수들을 동시에 취급할 수 있는 통계적 기법의 필요성이 요구됨은 물론이거니와 관심이 되는 모수의 수는 연구를 진행하는 의사결정자에게 전적으로 의존하게 된다. 그런데 한가지 주의하여야 할 점은 두개의 다른 모델로 부터 얻어진 추정치로 실지 모수들의 상대차이로 구간추정치를 구하는 것이 이 연구의 목적이지만 만약 관심의 대상이 되는 모수의 수(K)가 두개 이상인 경우 과연 실지로 동시에 고려되는 모수의 수를 K_e라 할때 이 값이 얼마가 되는지 먼저 이론적 또는 실험을 통하여 찾아야 한다. K와 K_e의 구분은 매우 중요한 의미를 갖는다. 왜냐하면 K개 각각의 모수에 대한 n_i는 결국 구하고자 하는 $n^* = \max(n_1, n_2, n_3, \ldots, n_{Ke})$를 구하는데 이용되고 신뢰구간에 쓰여지는 Z값은 $Z_{\alpha/(2K)}$가 아닌 $Z_{\alpha/(2Ke)}$이 되어야 한다는 사실이다. 만약 $K_e \leq K$인 경우에는 당연히 $Z_{\alpha/(2Ke)} \geq Z_{\alpha/(2K)}$의 관계가 성립하며 옳은 Z값의 사용은 결국 시뮬레이션 반복회수를 작게하는 결과를 낳게 해 준다는 사실을 주시하여야 한다.

동일한 조건하에서 실행된 시뮬레이션으로 얻어진 서로 다른 모수들의 추정치들 간에는 일반적으로 서로는 독립이 아니고 적지 않은 상관관계를 가지게 된다. 만약 K개의 모수가운데 i번째 모수의 상대차이를 $d_i, i = 1,2,\ldots,K$,라 할때 K개의 상대차이에 대한 신뢰구간을 $l(d_1), l(d_2), \ldots, l(d_K)$라 하고 $P[d_i \in l(d_i)] = 1-\alpha_i$라 하자. 실질적으로 동시에 고려되는 모수의 수는 K_e이므로 이들간에 서로 독립이 아닌 경우 일반적으로

$$P\left[\bigcap_{i=1}^{K_e}\{d_i \in l(d_i)\}\right] \neq \Pi_{i=1}^{K_e}(1-\alpha_i)$$

의 관계가 성립하지만 *Bonferroni*의 부등식을 이용하여 보면

$$P\left[\bigcap_{i=1}^{K_e}\{d_i \in l(d_i)\}\right] \geq 1 - \sum_{i=1}^{K_e}\alpha_i$$

의 관계를 알 수 있다. 여기서 K_e개의 모수들이 모두 동일하게 중요하다고 가정을 할때 만약 $\alpha_1 = \alpha_2 = \cdots\cdots = \alpha_K = \alpha/K_e$라고 하면

$$P\left[\bigcap_{i=1}^{K_e}\{d_i \in l(d_i)\}\right] \geq 1-\alpha$$

임을 알 수 있으며 여기서 $(1-\alpha)$는 전반적인 동시 신뢰구간의 포함확률(Coverage Probability)를 나타낸다.

Bonferroni 부등식을 고려한 이종 표본추출 절차(Double Sampling Procedure)를 간략하게 소개하면 아래와 같다.

1) 먼저 신뢰계수 $(1-\alpha)$를 정하고 관심이 되는 모수로서 $d_1, d_2, \ldots, d_K$를 채택하고 실지로 사용되는 K_e를 발견한다.
2) 미리 정해 놓은 시뮬레이션 반복시행회수(n_0) 만큼의 실험을 실시한다. 여기서 n_0는 얻어지는 관측치의 평균값이 정규성(Normality)을 만족하도록 선택된 반복회수이어야 한다.
3) $\alpha_i = \alpha/K_e$로 지정을 해 놓고 식 (6, 또는 10)을 이용하여 각 모수 d_i에 대한 n_i를 계산하여 $(n_i - n_0)$만큼의 추가로 요구되는 반복시행회수가 계산한다. $i = 1,2,\ldots,K$
4) 마지막으로 $n^* = \max(n_1, n_2, n_3, \ldots, n_K)$를 구하고 $n - n_0$만큼의 추가 반복시행을 거친 다음 각각은 $(1-\alpha/K_e)$를 신뢰계수로 갖는 K개의 상대차이에 대한 신뢰구간, $l(d_1), l(d_2), \ldots, l(d_K)$을 계산한다.

이와같은 절차를 통하여 추가로 시뮬레이션을 실행한 후 얻어지는 두 모수의 상대차이의 신뢰구간의 절반이 정해놓은 절대정도 ε보다 작게 됨을 기대할 수 있게 된다. 결과적으로 요약하면

$$P\left[\bigcap_{i=1}^{K_e}\{d_i \in l(d_i)\}\right] \geq 1-\alpha \text{ 이고 } \frac{1}{2}l(d_i) \leq \varepsilon,\ i = 1,2,\ldots,K,$$

이며, 여기서 $l(d_i)$는 i번째 신뢰구간의 간격을 의미한다.

이종추출절차의 시뮬레이션 적용 예

이종추출절차를 위한 첫번째 반복시행에서는 우선 표본평균과 표본분산의 정규성을 만족하기 위한 충분한 시행회수가 필요하다. 또한 모평균, 모분산, 그리고 첨도에 비

하여 추가로 시행되는 관측치들은 그다지 큰 변화가 없도록 역시 충분한 시행회수가 요구된다. 이와같은 요구를 만족시키기 위한 초기 반복시행회수(n_0)를 타당성있게 발견하기 위하여 대기행렬 및 재고관리문제에 대하여 각각 시뮬레이션을 실시하고 관측치의 평균값에 대한 정규성을 검토하여 본 결과 대기행렬의 경우 약 1,000회 그리고 재고관리 문제의 경우 약 500회를 적정 초기 시행회수로 발견할 수 있었다.

통계적 절차와 관련하여 이해를 돕기 위하여 다음의 표들을 만들어 이 연구에서 사용된 절차를 살펴 보기로 한다. 두가지의 표를 작성하여 예를 들어 보기로 한다. 먼저 〈표 A-1〉은 대기행렬문제의 예로서 시스템이 비어있는 상태에서 시작하여 500명 고객까지의 데이타를 가지고 관심이 되는 모수들로 고객의 평균 대기시간(Wq), 기다리는 평균 고객의 수(Lq), 그리고 시스템의 평균 가동율(ρ)의 세가지로 하고 이들 각각의 평균값과 표준편차에 대하여 상대차이와 절대정도를 알아보기 위하여 초기 시행회수로 얻어진 데이타를 이용하여 두 모델(M/M/1 과 Er/Er/1)의 모수들간 상대차이에 대한 신뢰구간과 구간의 절반에 해당하는 길이와 이를 절대정도에 만족시키기 위하여 필요한 반복시행회수를 계산하여 나타내고 있다. 모두 여섯개의 신뢰구간 가운데 L_q의 표준편차는 두 모델간에 -121.62%의 상대차이가 있는 것으로 추정이 되며 구간추정치의 절반(HW)은 40.6185%에 해당되며 이를 절도정도 7%가 되게 하려면 총 31,715만큼의 반복시행이 요구됨을 알 수 있다. 〈표 A-2〉는 추가로 요구되는 반복회수(31,715-1,000) 만큼 모의실험을 추가로 시행한 후 다시 〈표 A-1〉과 유사한 방법으로 상대차이의 신뢰구간을 계산하여 표로 나타낸 것으로서 여기에서 주목할 것은 여섯가지의 모수가운데 상대차이의 절반이 가장 큰것이 절대정도 6.9325%로서 7%에 거의 인접하고 있고 나머지 다섯가지의 절대정도들은 모두 7%내에 들어있는 점이다. 그리고 위의 여섯가지의 모수들에 대하여 두 모델간 상대차이의 90% 동시신뢰구간으로 이들 각각은 절대정도 7%내에서 보여주고 있음을 표에서 확인할 수 있다. 예로서, L_q의 평균값에 대한 상대차이는 하한이 -91.46%, 상한이 -87.08%인 신뢰구간을 보여주고 있다. 재고관리 시스템의 예에서는 (s,S) 재고정책으로 (20,80) 과 (20,40)의 두 모델을 놓고 평균 주문비용(O), 재고 평균 유지비용(H), 재고 평균 품절비용(S)의 세가지변수에 대하여 이들 각각의 평균과 표준편차에 대한 상대차이의 신뢰구간과 절대정도를 알아 보기로 한다. 〈표 B-1〉은 초기 500회의 시행결과 재고 품절비용(S)에서 두 모델간의 상대차이로 -34.07%의 추정치를 얻을 수 있었고 이때 절대정도는 14.55%로서 요구되는 2%를 만족하기 위하여 25,712회의 총 모의실험회수가 요구됨을 보여주고 있다. 〈표 B-2〉는 추가로 (25,712-500)의 모의실험을 행한 후 상대차이의 신뢰구간과 절대정도를 나타낸 표로서 여섯개의 모수 모두 2%내에 들어 있음을 알 수 있을 뿐만 아니라 이들 가운데 최고는 거의 2%에 가까운 1.9897%를 나타내고 있음을 알 수 있다. 여기에서도 마찬가지로 여섯개의 상대차이에 대한 신뢰구간은 90%이상의 신뢰수준으로 동시에 만족한다는 사실이다.

〈표 A-1〉 대기행렬 시스템의 예, 초기 1,000회 반복시행 결과 상대차이의 절대정도와 요구되는 반복시행회수(n^*), n_0=1,000, M/M/1 vs Er/Er/1.

	실제모델	근사모델
도착 시간간격 분포	Erlang-2	지수분포
서비스 시간간격 분포	Erlang-2	지수분포
평균 도착시간	1.00	1.00
평균 서비스시간	0.90	0.90
모수의 수(K_e) : 6	반복시행회수	: 1,000
유의수준(α) = 0.10		절대정도(ε)=7.00%

	모 델		상대차이 신뢰구간(%)		중간값	절반 간격	요구반복 시행회수 (n^*)
	실 제	근 사	하 한	상 한			
E[Wq]	3.5543	6.7855	(-103.43	-79.07)	-91.25	12.1838	3021
S(Wq)	1.9810	4.2780	(-155.90	-83.23)	-119.57	36.3370	25408
E[Lq]	3.5847	6.8818	(-105.15	-79.55)	-92.35	12.7987	3333
S(Lq)	2.0798	4.5439	(-162.24	-81.01)	-121.62	40.6180	31715
E[ρ]	0.8926	0.8878	(0.02	1.05)	0.54	0.5170	5
S(ρ)	0.0372	0.0484	(-39.99	-20.89)	-30.44	9.5486	1850

통계적 절차의 타당성 검토

지금까지 언급한 이종추출절차를 이용한 두 모수의 상대차이에 대한 시뮬레이션방법의 타당성을 검토하기 위하여 실제로 이론값이 알려진 대기행렬의 모델을 선정하여 검토를 실시하여 보기로 하자. 대기행렬의 경우 고객의 도착은 포아송과정을 따르고 서비스의 완료가 지수분포를 나타내며 제공하는 서어버의 수가 하나인 M/M/1의 경우를 근사모델로하고 이에 대하여 실제모델이 Er/Er/1이라고 가정했을때의 예를 들어 설명하기로 하자. 또한 관심

〈표 A-2〉 대기행렬 시스템의 예, 요구되는 반복시행 결과 상대차이의 절대정도, $n^*=31,715$, M/M/1 vs Er/Er/1.

	실제모델	근사모델
도착 시간간격 분포	Erlang-2	지수분포
서비스 시간간격 분포	Erlang-2	지수분포
평균 도착시간	1.00	1.00
평균 서비스시간	0.90	0.90
모수의 수(Ke) : 6	반복시행회수:	31,715
유의수준(α) = 0.10		절대정도(ε)=7.00%

	모델		상대차이 신뢰구간(%)		중간값	절반 간격 % (HW)
	실제	근사	하한	상한		
E[Wq]	3.5247	6.6362	(-90.38	-86.20)	-88.29	2.0901
S(Wq)	1.9580	4.0547	(-113.99	-100.46)	-107.22	6.7670
E[Lq]	3.5561	6.7303	(-91.46	-87.08)	-89.27	2.1904
S(Lq)	2.0528	4.2983	(-116.47	-102.60)	-109.53	6.9325
E[ρ]	0.8920	0.8857	(0.61	0.80)	0.71	0.0924
S(ρ)	0.0369	0.0492	(-35.03	-31.66)	-33.34	1.6835

〈표 B-1〉 (s,S) 재고관리 시스템의 예, 초기 500회 반복시행 결과 상대차이의 절대정도와 요구되는 반복시행회수(n^*), no=500, (20,80) vs (20,40).

초기 재고수준	60 개		
수요량의 구분 (Di=0,1,2,3)	4 가지		
수요량의 확률분포; P(D≦Di)	0.167 0.5 0.833 1		
평균 수요 간격	0.1 개월		
배송 지연 범위	0.5 내지 1 개월		
모의실험 길이	120 개월		
o = 32 (Ordering)	i = 3 (Increment)	h = 1 (Holding)	pi = 5 (Backlog)
실제 모델 (s,S) : (20,80)			
근사 모델 (s,S) : (20,40)			
모수의 수(K_e) : 6	반복시행회수 : 500		
유의수준(α) = 0.10	절대정도(ε)=2.00%		

	모델		상대차이 신뢰구간(%)		중간값	절반 간격	요구 반복시행 회수(n^*)
	실제	근사	하한	상한			
E[O]	84.8978	97.7325	(-15.63	-14.58)	-15.11	0.53	34
S[O]	2.5836	2.9160	(-25.45	-1.57)	-13.51	11.94	17387
E[H]	26.8870	9.1355	(65.78	66.26)	66.02	0.24	7
S[H]	0.9408	0.5163	(39.02	50.68)	1.00	5.83	4166
E[S]	9.5911	18.3257	(-95.93	-86.37)	-91.15	4.78	2849
S[S]	1.8368	2.4434	(-48.63	-19.52)	-34.07	14.55	25712

〈표 B-2〉 (s,S) 재고관리 시스템의 예, 요구되는 반복시행 결과 상대차이의 절대정도, $n^*=25,712$, (20,80) vs (20,40).

초기 재고수준	60 개		
수요량의 구분 (Di=0,1,2,3)	4 가지		
수요량의 확률분포; P(D≦Di)	0.167 0.5 0.833 1		
평균 수요 간격	0.1 개월		
배송 지연 범위	0.5 내지 1 개월		
모의실험 길이	120 개월		
o = 32 (Ordering)	i = 3 (Increment)	h = 1 (Holding)	pi = 5 (Backlog)
실제 모델 (s,S) : (20,80)			
근사 모델 (s,S) : (20,40)			
모수의 수(K_e) : 6	반복시행회수 : 25,712		
유의수준(α) = 0.10	절대정도(ε) = 2.00%		

	모델		상대차이 신뢰구간(%)		중간값	절반 간격 % (HW)
	실제	근사	하한	상한		
E[O]	84.9148	97.6740	(-15.10	-14.95)	-15.03	0.0748
E[O]	2.6439	2.9757	(-14.25	-10.87)	-12.56	1.6891
E[H]	26.8700	9.1238	(66.01	66.08)	66.04	0.0352
S[H]	0.9779	0.5386	(44.10	45.74)	44.92	0.8199
E[S]	9.5726	18.3081	(-91.92	-90.59)	-91.26	0.6618
S[S]	1.8243	2.4151	(-34.39	-30.41)	-32.40	1.9897

의 대상이 되는 모수들을 고객의 대기시간(Wq), 기다리는 고객의 수(Lq), 그리고 시스템의 가동율(ρ)의 세가지로 제한하고 여기에 부연하여 이들 각각의 표준편차도 함께 고려하기로 하자. 초기 반복시행회수(n_0)를 1,000회로 하여 고객의 도착율과 서비스의 완료율을 각각 $\lambda=1.0$ 그리고 $\mu=1/.90$ 로 하여 20회의 독립적인 반복시행을 실시하여 보기로 한다. 한편 절대정도 ε은 .07로 하기로 하자. 먼저 1,000회의 모의실험을 거쳐서 얻어진 데이타를 가지고 표를 완성하여 상대차이들의 신뢰구간을 구한 다음 상대정도를 만족하는 추가 모의실험반복회수를 계산하여 모의실험을 실시한다. 다시 〈표A-2〉와 같이 작성하여 여섯가지 모수들의 상대차이의 신뢰구간 절반에 해당하는 길이가 절대정도 ε를 크게 벗어나면 성공, 이들중 하나라도 크게 벗어나는 경우에는 실패라고 하면, 20회 독립시행 결과 8.0%를 넘는 한가지 경우를 제외하면 나머지 19경우에는 모두 절도정도 7%를 크게 벗어나지 않고 있음을 알 수 있었다. 7%주변에서 약간의 변화는 표본평균 또는 표본분산의 정규성에 기인한다고 볼 수 있겠다. 실험결과를

요약한 내용이 아래의 〈표 C〉에 나타나 있다.

〈표 C〉 대기행렬 시스템의 예, 20회 독립시행 결과 나타난 최대 절대정도, n_o=1,000, 모수의 수 K_e=6, M/M/1 vs Er/Er/1, ε=7%, 유의수준 α=.10.

시행회수	1	2	3	4	5	6	7	8	9	10
최대(%) 절대정도	7.88	7.53	7.25	7.11	7.84	6.96	6.98	7.72	7.56	7.39
시행회수	11	12	13	14	15	16	17	18	19	20
최대(%) 절대정도	8.02	6.95	7.79	7.52	7.80	7.11	7.85	7.62	6.93	6.68

(s,S) 재고관리 문제에서는 주문비용(O), 유지비용(H), 그리고 품절비용(S)에 대하여 이들 각각의 평균 및 표준편차를 모수로하는 전체 여섯개의 모수에 대한 상대차이를 이용하여 검토하여 보기로 하되 실제모델로 (20,80)을 근사모델로 (20,40)을 가정하여 실험을 한다. 대기행렬에서와 같이 우선 500회의 모의실험을 거쳐서 얻어진 데이타를 가지고 표를 완성하여 상대차이들의 신뢰구간을 구한 다음 절대정도를 만족하는 추가 모의실험반복회수를 계산하여 모의실험을 실시한다. 대기행렬의 경우와 마찬가지로 여섯가지 모수들의 상대차이의 신뢰구간 절반에 해당하는 길이가 절대정도 ε을 크게 벗어나는지의 여부를 확인하면 거의 모든 경우에 절도정도 2%주변을 보여주고 있음을 알 수 있어서 대기행렬의 경우보다는 표본평균 또는 표본분산의 출력면에서 정규성을 지닌 안정성에 기인한 통계적절차의 이상적인 면을 보여주고 있다. 20회 독립된 실험 결과가 아래의 〈표 D〉에 나타나 있다.

〈표 D〉 재고관리 시스템의 예, 20회 독립시행 결과 나타난 최대 절대정도, no=500, 모수의 수 Ke=6, (s, S)=(20,80) vs (20,40), ε=2%, 유의수준 α=.10.

시행회수	1	2	3	4	5	6	7	8	9	10
최대(%) 절대정도	1.99	1.95	2.08	2.00	2.24	2.05	2.05	1.95	2.16	2.16
시행회수	11	12	13	14	15	16	17	18	19	20
최대(%) 절대정도	1.95	1.91	2.30	2.23	1.99	1.80	2.07	1.93	2.01	1.97

결 론

시뮬레이션모델은 복잡한 시스템을 연구하는데 가장 편리하고 효율적인 도구로서 점차 큰 역할을 담당하고 있다. 모델을 설정해 놓은 전문가들이 유효 적절한 통계적 방법을 필요로 하는 것은 당연한 일이 아닐 수 없다. 모델에 나타나는 몇가지 중요한 모수(Parameter)들의 出力(Output)을 효율적으로 다루므로서 적절한 회수의 시뮬레이션 반복시행을 통하여 비용을 감소함과 동시에 신뢰성 높은 결과를 기대할 수 있게 된다.

주어진 시스템의 연구를 위하여 가정된 실제의 모델과 근사모델로서 모두 시뮬레이션 모델로 완성하고 이들의 반복시행으로 부터 얻어지는 관심대상이 되는 모수들의 추정치를 이용하여 이들 두 모델간의 상대차이를 추정하고자 하는 것이 이 연구의 주된 목적이다. 이종추출절차를 적용하되 두 모수값의 상대차이에 대한 점추정치보다 한층 더 효율적인 정보를 획득하기 위하여 신뢰구간을 형성하고 미리 설정해 놓은 절대정도를 충족시키기 위하여 예상되는 추가 반복회수를 계산한다. 추가로 모의실험한 결과 데이타들을 가지고 다시 신뢰구간을 구하면 이때에 얻어지는 신뢰구간의 절반에 해당하는 길이는 절대정도를 만족하는 값을 가지게 된다. 수행측도의 평균값이 정규성을 보장하고 분산의 변화를 최소로 하기 위하여 충분한 반복시행회수를 찾아내어 실시하여야 한다. 이 연구에서는 수행측도의 모수로서 평균값 뿐만 아니라 표준편차도 고려하여 포함시키고 있다. 대기행렬과 재고관리 문제에서 세가지의 평균 수행측도치와 이들 각각의 표준편차를 더하여 모두 여섯개의 모수들을 고려하여 타당성 검토와 결과 분석에 이용하였다. 서로 다른 두 종류의 시스템을 이용한 실험결과를 토대로 이 연구에서 개발된 이종추출절차를 이용한 통계적 방법의 타당성과 효율성이 입증되었다.

본 연구결과가 가져다 주는 기여 또는 기대효과는 다음에 열거하는 몇가지로 요약할 수 있겠다. 첫째, 한 시스템을 놓고 두개의 서로 다른 모델을 비교.분석하는 문제에서 만약 이 시스템의 출력이 확률법칙에 따라 움직이는 확률변수라 한다면 이 연구에서 얻어지는 통계적기법을 이용하여 신뢰있는 충분한 정보를 얻을 수 있게 된다는 것이다. 둘째, 시뮬레이션에서 흔히 범하기 쉬운 과대한

비용의 지출을 적절한 시행반복회수를 통해서 방지할 수 있다는 것이다. 세째, 의사결정자의 결정에 따른 절대정도(Absolute Precision)의 선택으로 이 연구의 활용성도 다각적인 측면에서 가능하다는 것이다. 네째, 평균치의 분석뿐만 아니라 나아가 2차 적률을 고려하여 표준편차의 상대차이에 대한 신뢰구간 문제를 다룬다는 점이며, 다섯째로는, 시뮬레이션의 기본원리와 조금의 컴퓨터 지식을 갖추고 기초적인 통계학의 원리를 습득하고 있는 사용자들이라면 어느 분야에 종사하든지 쉽게 이해하고 유효적절한 시기에 간단히 사용할 수 있다는 것이다. 마지막으로 첨단 과학문명의 발달과 더불어 계획되고 투자되는 대규모 비용의 연구를 시행하는 여러 전문가들에게 시뮬레이션의 중요성은 강조가 됨은 말할 나위가 없겠으나 이 시뮬레이션을 수행하고 분석하는 과정에서 올바른 또는 타당성있는 모델의 사용법은 더욱 더 중요한 것이다. 실지로 재화를 투자하기에 앞서 올바른 판단의 기준으로서 이용이 기대되는 본 연구의 결과는 앞으로 여러 연구개발분야에서 사전에 익히고 실제로 적용될 수 있는 효과적인 방법론으로 기대된다.

참고문헌

[1] Bratley, P., Fox,B.L., and Schrage, L.E., A Guide to Simulation, Spring-Verlag, New York, 1983.

[2] Cnover, W.J., *Practical Nonparametric Statistics,* 2nd Ed., John Wiley, Now York, 1980.

[3] Cramer, H., *Mathematical Methods of Statistics,* Princeton University Press, Princeton, 1946.

[4] Devore, J.L., *Probability and Statistics for Engineering and Sciences,* Brooks/Cole, Monterey, California, 1982.

[5] Dudewicz, E.J. and Ralley, T.G., *The Handbook of Random Number Generation and Testing with TEST RAND Computer Code,* American Sciences Press, Inc., Columbus, 1981.

[6] Dudewicz, E.J. and Dalar, S.R., "Allocation of Obersations in Ranking and Selection with Unequal Variances," *Sankhya,* B37, pp. 28-78, 1975.

[7] Fieller, E.C., "Some Problems in Interval Estimation," *Journal of the Royal Statistical Society,* Series B, 2, pp. 175-185, 1954.

[8] Fishman, G.S., *Principles of Discrete Event Simulation,* John Wiley and Sons, New York, 1978.

[9] Fishman, G.S. and Moore, L.R., "A Statistical Evaluation of Multiplicative Congruential Random Mumber Generators with Modulus 231 - 1," *Journal of the American Statistical Association,* 77, pp. 129-136, 1982.

[10] Gupta, S.S. and Hsu, J.C., "Computer Package for Ranking, Selection, and Multiple Comparisons with the Best," *Proc. 1984 Winter Simulation Conference,* Dallas, pp. 251-257, 1984.

[11] Hsu, J.C., "Constrained Two-Sided Simultaneous Confidence Intervals for Multiple Comparisons with the Best," *Ann. Statistics,* 12, pp. 1136-1144, 1984.

[12] Hsu, J.C. and Edwards, D.G., "Sequential Comparisons with the Best," *Journal of the American Statistical Association,* 78, pp. 958-964, 1983.

[13] Kleijnen, J.P.C., *Statistical Techniques in Simulation,* pt. II, Dekker, New York, 1974.

[14] Kleijnen, J.P.C., *Statistical Tools for Simulation Practitioners,* Dekker, New York, 1987.

[15] Koenig, L.W. and Law, A.M., "A Procedure for Selecting a Subset of Size m Containing the l Best of k Independent Normal Populations, with Applications to Simulation," *Commun. Statist. Simulation and Computation,* 14, pp. 719-734, 1985.

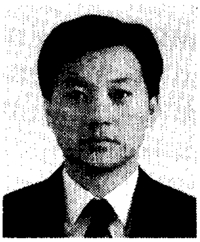
[16] Law, A.M. and Kelton, W.D., *Simulation Modeling and Analysis,* 2nd Ed., McGraw Hill Book Co., New York, 1991.

[17] Lewis, P.A.W., Goodman, A.S., and Miller, J.M., "A Pseudo-random Number Generator for the System 1360," *IBM Systems Journal,* 8, pp. 136-146,1969.

[18] Scheffe, H., *The Analysis of Variance,* John Wiley and Sons, New York, 1959.

[19] Welch, B. L., "The Significance of the Difference between Two Means when the Population Variances are Unequal," Biometrika, 25, pp. 350-360, 1938.

● 저자소개 ●



홍 윤 기

1980년 고려대학교 공과대학 산업공학과 졸업(공학사)
1985년 University of Southern California 산업공학과(OR 석사)
1989년 University of Southern California 산업공학과(산업공학 박사)
1989년~1991년 California State University at Northridge, 경영과학과 조교수
1991년~현재 한성대학교 이공대학 산업공학과 조교수
관심분야: Simulation Output Analysis, Wargame Simulation, Stochastic Combat Analysis