

확률적 활동 네트워크에서 사업완성시간의 적률 추정: 활동시간의 일반적 분포

(Estimating the Moments of the Project Completion Time in Stochastic Activity Networks: General Distributions for Activity Durations)

조재균^{1)*}

(Jae-Gyeun Cho)

요약 Cho[1]는 확률적 활동 네트워크 분석에서 활동시간이 상호 독립적이고 정규분포를 따른다는 가정 하에서 사업완성시간의 적률 (평균, 분산, 왜도, 첨도)을 추정하기 위한 방법을 제안하였다. 본 논문에서는 활동시간의 분포가 일반적인 분포일 때 사업완성시간의 적률을 추정하기 위한 방법을 제안한다. 제안된 방법은 활동시간 분포의 이산화를 위해 적률매칭 방법을 사용하며, 사업완성시간의 계산에 사용될 활동시간을 결정하는데 이산형 역변환 방법을 사용한다. 제안된 방법은 대규모 네트워크에 적용하기 쉽고, 몬테칼로 시뮬레이션 보다 계산적으로 효율적이며, 제안된 방법의 결과는 몬테칼로 시뮬레이션에 의한 결과와 잘 일치함을 보여준다.

핵심주제어 : 확률적 활동 네트워크, 사업완성시간, 이산화

Abstract In a previous article, for analyzing a stochastic activity network, Cho[1] proposed a method for estimating the moments (mean, variance, skewness, kurtosis) of the project completion time under the assumption that the durations of activities are independently and normally distributed. Developed in the present article is a method for estimating those moments for stochastic activity networks which allow any type of distributions for activity durations. The proposed method uses the moment matching approach to discretize the distribution function of activity duration, and then a discrete inverse-transform method to determine activity durations to be used for calculating the project completion time. The proposed method can be easily applied to large-sized activity networks, and computationally more efficient than Monte Carlo simulation, and its accuracy is comparable to that of Monte Carlo simulation.

Key Words : Stochastic Activity Network, Project Completion Time, Discretization

* Corresponding Author : jgcho@deu.ac.kr

+ 이 논문은 2017학년도 동의대학교 연구년 지원에 의하여 연구되었음.

Manuscript received April 7, 2018 / revised June 1, 2018 / accepted June 14, 2018

1) 동의대학교 정보경영학부

1. 서론

사업의 효율적 통제 및 관리를 위해 유용하게 사용되는 모형인 확률적 활동 네트워크

(Stochastic Activity Network)에서 각 활동시간은 활동의 불확정성(Uncertainty)으로 인해 양의 값을 갖는 확률변수로서 나타내어지며, 이로 인해 사업완성시간(Project Completion Time)도 확률변수가 된다. 확률적 활동 네트워크를 분석함에 있어, 사업완성시간의 분포 또는 적률(Moment)을 추정하는 것은 매우 기본적이고도 필수적이다.

사업완성시간의 분포 또는 적률을 추정하는 방법들에 관해 많은 연구들이 수행되었으며, 그 방법들은 해석적 방법(Exact Methods), 몬테칼로 시뮬레이션(Monte Carlo Simulations), 근사적 방법(Approximation Methods)으로 분류된다[2]. 해석적 방법은 정확한 결과를 얻을 수 있지만, 활동시간의 분포에 대한 매우 제한적인 가정과 계산적인 복잡성 및 많은 계산량으로 인해 다양한 활동시간 분포를 가지는 대규모 활동 네트워크에 적용하기가 어렵다. 몬테칼로 시뮬레이션은 활동시간 분포에 대해 어떠한 제한적인 가정 없이 대규모의 활동 네트워크 대해서도 쉽게 적용 가능하며, 상당히 정확한 추정치 결과를 얻을 수 있기 때문에 가장 빈번히 사용되는 방법이다. 그리고 다른 근사적 방법에 의해 얻어진 결과들의 정확도를 입증하기 위한 비교 기준으로 종종 사용된다. 반면에, 추정치의 정확도를 높이기 위해서는 많은 계산량이 요구되는 단점이 있다.

해석적 방법과 몬테칼로 시뮬레이션이 가지는 한계점들을 보완하기 위해 많은 근사적 방법들이 제안되었다. Sculli[3]는 활동시간들이 독립적이고 정규분포를 따른다는 가정하에서 사업완성시간의 평균과 분산을 근사적으로 계산했으며, Dodin[4]은 활동시간의 연속형 분포를 이산형 분포로 근사화하기 위해 이산점들(Discrete Points)과 대응되는 확률들을 결정하는 이산화 기법(Technique of Discretization)을 사용하여, 사업완성시간의 분포함수와 평균 및 분산을 계산하였다. Dodin and Sirvanci[5]는 극단값 이론(Extreme Value Theory)을 이용하여 사업완성시간의 분포함수와 평균 및 분산을 추정하였다. Mehrotra *et al.*[6]은 활동시간들이 독립적이고 동일한 분포를 갖는다는 가정하에서 주경로(Critical Path)들 간의 의존성(Interdependency) 정도를 고려하여 사업

완성시간의 평균과 분산을 계산하였으며, 활동시간이 정규분포와 지수분포인 경우에 대해 계산적 결과를 제시하였다. Abdelkader[7]는 활동시간이 와이블 분포를 따르는 경우에 적률 방법(Moments Method)을 사용하여 사업완성시간의 평균을 계산하였다. Yao and Chu[2]는 Dodin[4]의 방법을 수정함으로써 계산량을 줄이고 사업완성시간의 평균을 더 정확히 추정할 수 있는 방법을 제안하였다.

Cho[1]는 활동시간이 상호 독립적이고 정규분포를 따른다는 가정 하에서 사업완성시간의 적률(평균, 분산, 왜도, 첨도)을 추정하기 위한 방법을 제안하였다. 본 논문에서는 활동시간의 분포가 일반적인 분포를 가질 때 사업완성시간의 적률을 추정하기 위한 방법을 제안한다. 제안된 방법은 근사적 방법과 몬테칼로 시뮬레이션을 결합한 방법으로, 활동시간의 분포함수를 이산화하기 위하여 적률매칭 방법(Moment Matching Approach)을 사용하며, 사업완성시간의 계산에 사용될 활동시간값을 결정하기 위해 이산형 역변환 방법(Discrete Inverse-Transform Method)을 사용한다. 활동시간의 분포함수를 이산화하는 목적은 계산적 효율성을 높이고, 활동시간이 심하게 비대칭이거나 꼬리부분이 두텁고 긴 분포를 가질 경우에 사업완성시간의 안정적인 적률 추정치를 얻기 위함이다.

2절에서는 연속형의 활동시간 분포를 이산형 분포로 근사화하기 위해 이산점들과 대응되는 가중치를 결정하기 위한 이산화 기법을 설명한다. 3절에서는 사업완성시간의 적률을 추정하기 위하여 Cho[1]의 절차를 수정한 절차를 제안하고 예증한다. 4절에서는 예제 네트워크들에 대해 계산적 실험을 수행함으로써 제안된 방법의 성능을 평가하고, 몬테칼로 시뮬레이션과 비교한다. 마지막으로 5절에서는 본 연구에 대한 결론을 맺는다.

2. 이산화 기법

본 절에서는 통계적 허용차 설계(Statistical Tolerance Design)에서 시스템 구성요소의 수준

값과 대응되는 가중치를 결정하는 방법을 설명한다. 이것은 본 연구에서 활동시간 분포의 이산화를 위해 이용된다.

통계적 허용차 설계는 허용차 분석(Tolerance Analysis)과 허용차 배분(Tolerance Allocation)의 두 단계로 이루어져 있다. 허용차 분석 단계에서는 시스템 구성요소의 허용차가 시스템 특성치의 변동에 미치는 영향을 평가하게 되며, 허용차 배분 단계에서는 시스템 특성치의 허용차를 조립성, 비용 등을 고려하여 시스템 구성요소에 합리적으로 배분하게 된다[8].

허용차 분석을 위해서는 시스템 구성요소들의 분포 또는 적률이 주어졌을 때, 시스템 특성치의 분포 또는 적률을 추정할 필요가 있다. 허용차 분석을 위해 테일러급수 전개 방법(Taylor Series Expansion Method), 1차/2차 신뢰도 방법(First/Second-Order Reliability Method), 몬테카를로 시뮬레이션(Monte Carlo Simulation), 수치적 분법(Numerical Quadrature Method), 실험계획법(Experimental Design Technique) 등의 다양한 방법들이 제안되었다[9].

실험계획법에 의한 허용차 분석에서, Taguchi[10]는 직교표(Orthogonal Array, OA)를 이용한 실험과 ANOVA(Analysis of Variance)에 의한 실험 데이터의 분석을 제안한다. Taguchi[10]의 3수준 OA 실험에서, 시스템 구성요소(실험계획법에서는 요인이라 부름)의 수준값은 $(\mu_i - \sigma_i \sqrt{3/2}, \mu_i, \mu_i + \sigma_i \sqrt{3/2})$ 그리고 대응되는 가중치는 $(1/3, 1/3, 1/3)$ 로 결정할 것을 추천한다. 여기서, μ_i 와 σ_i 는 각각 요인 i 의 평균과 표준편차를 나타낸다. D'Errico and Zaino[11]는 시스템 특성치의 고차 적률을 더 정확하게 추정할 수 있도록 Taguchi[10] 방법을 수정한 방법을 제안한다. 그들의 방법은 요인 i 의 수준값은 $(\mu_i - \sigma_i \sqrt{3}, \mu_i, \mu_i + \sigma_i \sqrt{3})$ 로 하고, 대응되는 가중치는 $(1/6, 4/6, 1/6)$ 을 부여한다. 그리고 시스템 특성치의 값들을 계산하는데 있어 3수준 완전요인계획법(Full Factorial Design)을 사용한다. 그런데 Taguchi[10] 방법과 D'Errico and Zaino[11] 방법에서 요인의 수준값 선택은 요인이 정규분포를 따른다는 가정을 하고 있다. Seo

and Kwak[12]은 D'Errico and Zaino[11]의 방법을 확장하여 시스템 구성요소가 정규분포가 아닌 일반적인 분포인 경우에도 적용 가능한 수준값과 가중치의 결정 방법을 제안하고 있다. Seo and Kwak[12]에 의해 제안된 수준값과 대응되는 가중치의 결정 방법은 다음과 같다.

시스템이 한 개의 요인으로 구성되어 있는 경우에, 요인을 x , 시스템 특성치를 $g(x)$ 라 할 때, $g(x)$ 의 원점에 관한 k 차 적률은 절점(Node)의 수가 m 인 구적법식(Quadrature Formula)을 이용하여 다음과 같이 근사화할 수 있다.

$$\begin{aligned} E\{g^k\} &= \int_{-\infty}^{\infty} [g(x)]^k f(x) dx \\ &\cong w_1 [g(\mu + \alpha_1 \sigma)]^k + w_2 [g(\mu + \alpha_2 \sigma)]^k \\ &\quad + \dots + w_m [g(\mu + \alpha_m \sigma)]^k \end{aligned} \quad (1)$$

단, μ , σ , $f(x)$ 는 x 의 평균, 표준편차, 확률밀도 함수를 나타낸다. $g(x)$ 의 1~4차까지의 적률을 정확히 추정하기 위해서는, 식(1)에서 m 의 값은 최소한 3이어야 한다. $(\alpha_1, \alpha_2, \alpha_3)$ 와 (w_1, w_2, w_3) 은 다음의 식(2)와 같은 적률매칭 방정식을 풀어 구할 수 있다.

$$\begin{aligned} M_k &= \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx \\ &= w_1 (\alpha_1 \sigma)^k + w_2 (\alpha_2 \sigma)^k + w_3 (\alpha_3 \sigma)^k \quad (k = 0 \sim 5) \end{aligned} \quad (2)$$

단, M_k 는 x 의 평균에 관한 k 차 적률이다. 식(2)에서 $\mu + \alpha_i \sigma$ 를 l_i 로 대체하면 식(2)는 식(3)과 같이 나타낼 수 있다.

$$\begin{aligned} w_1 + w_2 + w_3 &= 1 \\ w_1 l_1 + w_2 l_2 + w_3 l_3 &= \mu \\ w_1 (l_1 - \mu)^2 + w_2 (l_2 - \mu)^2 + w_3 (l_3 - \mu)^2 &= \sigma^2 \\ \frac{w_1 (l_1 - \mu)^3 + w_2 (l_2 - \mu)^3 + w_3 (l_3 - \mu)^3}{\sigma^3} &= \sqrt{\beta_1} \\ \frac{w_1 (l_1 - \mu)^4 + w_2 (l_2 - \mu)^4 + w_3 (l_3 - \mu)^4}{\sigma^4} &= \beta_2 \\ w_1 (l_1 - \mu)^5 + w_2 (l_2 - \mu)^5 + w_3 (l_3 - \mu)^5 &= M_5 \end{aligned} \quad (3)$$

단, $\sqrt{\beta_1}$ 와 β_2 는 $g(x)$ 의 왜도와 첨도를 나타낸다. 식(3)의 연립방정식을 풀어 (l_1, l_2, l_3) 와 (w_1, w_2, w_3) 를 구하는 것은 매우 복잡하기 때문에, Seo and Kwak[12]은 식(3)에서 5차 적률 조건인 마지막 등식을 삭제하고, l_2 를 $\mu + \Delta$ 라 둔 후, 연립방정식을 풀어 수준값 (l_1, l_2, l_3) 와 가중치 (w_1, w_2, w_3) 를 식(4) 및 식(5)와 같이 구한다.

시스템이 n 개의 독립적인 요인으로 구성되어 있는 경우에는, 적구적법 규칙(Product Quadrature Rule)에 의하여 일차원 구적법을 n 차원의 각각에 적용 가능하며, 시스템 특성치 $g(x_1, \dots, x_n)$ 의 1~4차 적률들은 다음과 같이 구할 수 있다.

$$\begin{aligned} \mu_g &= \sum_{i_1=1}^3 w_{1i_1} \cdots \sum_{i_n=1}^3 w_{ni_n} g(l_{1i_1}, \dots, l_{ni_n}) \\ \sigma_g^2 &= \sum_{i_1=1}^3 w_{1i_1} \cdots \sum_{i_n=1}^3 w_{ni_n} \{g(l_{1i_1}, \dots, l_{ni_n}) - \mu_g\}^2 \\ \sqrt{\beta_{1g}} &= \left[\sum_{i_1=1}^3 w_{1i_1} \cdots \sum_{i_n=1}^3 w_{ni_n} \{g(l_{1i_1}, \dots, l_{ni_n}) - \mu_g\}^3 \right] / \sigma_g^3 \\ \beta_{2g} &= \left[\sum_{i_1=1}^3 w_{1i_1} \cdots \sum_{i_n=1}^3 w_{ni_n} \{g(l_{1i_1}, \dots, l_{ni_n}) - \mu_g\}^4 \right] / \sigma_g^4 \quad (6) \end{aligned}$$

단, l_{ij} 와 w_{ij} 는 각각 요인 i 의 j 번째 수준과 가중치를 나타낸다. 그리고 μ_g , σ_g^2 , $\sqrt{\beta_{1g}}$, β_{2g} 는 각각 $g(x_1, \dots, x_n)$ 의 평균, 분산, 왜도, 첨도를 나타낸다. 식(6)은 3^n 완전요인계획법을 사용하여 계산할 수 있다.

3. 사업완성시간의 적률 추정 절차

2절에서의 논의를 바탕으로, 사업완성시간의 적률 (평균, 분산, 왜도, 첨도)을 추정하기 위한 절차는 다음과 같다. 활동과 사업완성시간은 2절에서의 요인과 시스템 특성치에 각각 대응되며, 수준값은 1절에서의 이산점에 대응된다.

제안된 절차는 활동시간의 수준값과 대응되는 가중치를 결정하기 위하여 Seo and Kwak[12]에 의해 제안된 방법을 사용한다. 그리고 사업완성시간의 계산에 사용될 활동시간값을 결정하기 위하여, 대규모 활동 네트워크에 대해서는 계산량이 지나치게 많아지는 3수준 완전요인계획법 대신에 이산형 역변환 방법[13]을 사용한다. 제안된 절차를 흐름도로 나타내면 Fig. 1과 같다.

$$(l_1, l_2, l_3) = \left\{ \frac{B - \sqrt{C}\sigma}{A}, \mu + \Delta, \frac{B + \sqrt{C}\sigma}{A} \right\} \quad (4)$$

$$(w_1, w_2, w_3) = \begin{bmatrix} \frac{F - E\sqrt{C}}{C\sigma - D\sqrt{C}} \\ \frac{(\beta_2 - \beta_1 - 1)\sigma^4}{\Delta^4 - 2\sqrt{\beta_1}\sigma\Delta^3 + (\beta_2 - 3)\sigma^2\Delta^2 + 2\sqrt{\beta_1}\sigma^3\Delta + (\beta_2 - \beta_1)\sigma^4} \\ \frac{F + E\sqrt{C}}{C\sigma + D\sqrt{C}} \end{bmatrix}^T \quad (5)$$

단,

$$\begin{aligned} A &= 2(-\Delta^2 + \sqrt{\beta_1}\sigma\Delta + \sigma^2), \\ B &= (-\sqrt{\beta_1}\sigma - 2\mu)\Delta^2 + ((\beta_2 - 1)\sigma^2 + 2\mu\sqrt{\beta_1}\sigma)\Delta + \sqrt{\beta_1}\sigma^3 + 2\mu\sigma^2, \\ C &= (\beta_1 + 4)\Delta^4 - 2\sqrt{\beta_1}(\beta_2 + 3)\sigma\Delta^3 + (\beta_2^2 - 6\beta_2 + 6\beta_1 - 3)\sigma^2\Delta^2 + 2\sqrt{\beta_1}(3\beta_2 - 2\beta_1 + 1)\sigma^3\Delta \\ &\quad + (4\beta_2 - 3\beta_1)\sigma^4, \\ D &= 2\Delta^3 - 3\sqrt{\beta_1}\sigma\Delta^2 + (\beta_2 - 3)\sigma^2\Delta + \sqrt{\beta_1}\sigma^3, \\ E &= \Delta^3 - \sqrt{\beta_1}\sigma\Delta^2 - \sigma^2\Delta, \\ F &= \sqrt{\beta_1}\Delta^5 - (\beta_2 + \beta_1 - 3)\sigma\Delta^4 + (\beta_2 - 7)\sqrt{\beta_1}\sigma^2\Delta^3 + (\beta_2 + 3\beta_1 - 5)\sigma^3\Delta^2 + 5\sqrt{\beta_1}\sigma^4\Delta + 2\sigma^5 \end{aligned}$$

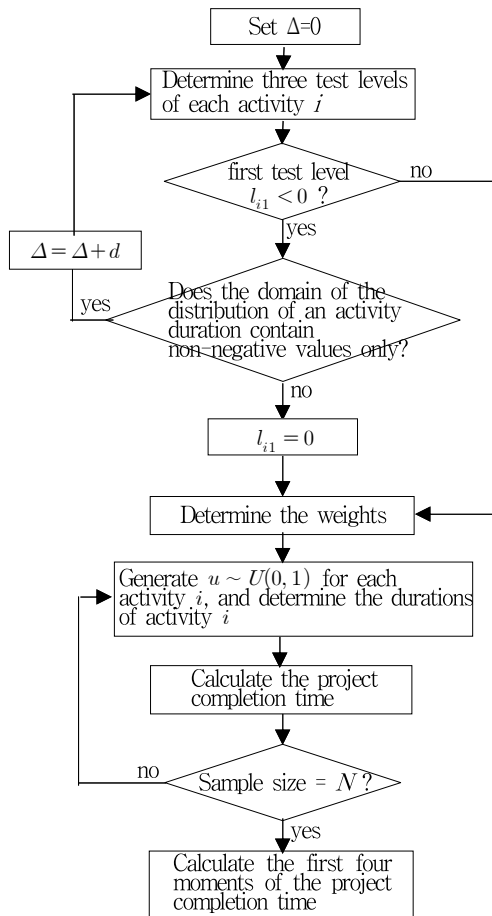


Fig. 1 The Proposed Procedure for Estimating the Moments of the Project Completion Time.

제안된 절차에서, 각 활동의 처음 3개의 수준값은 식(4)에서 $\Delta = 0$ 라 두고 계산되어진다. 이때 1수준값 l_1 은 음수값이 될 수도 있다. 이러한 경우에 제안된 절차에서의 처리방법은 다음과 같다. 활동시간 분포의 도메인이 비음(Non-Negative)의 값들만을 포함하는 경우(예를 들면, 4절의 Table 1에서 Normal을 제외한 모든 분포), l_1 이 비음의 값을 가질 때까지 Δ 를 점차적으로 증가시키고, 최종적인 Δ 값을 가지고 식(5)를 사용하여 가중치를 계산한다. 반면에, 활동시간 분포의 도메인이 음수의 값들을 포함할 수도 있는 경우(예를 들면, Normal), $l_1 = 0$ 라 두고, $\Delta = 0$ 값을 가지고 식(5)를 사용하여 가중치를 계산한다. 그리고 이산형의 활동시간 분포에 대해 확률변수값을 생성하기 위하여 이산형 역변환

방법을 사용한다. 이산형 역변환 방법에 의한 확률변수의 생성 방법은 다음과 같다. 이산형 확률변수 X 가 $x_1, x_2, \dots (x_1 < x_2 < \dots)$ 의 값을 가지며, 확률질량함수가 $p(x_i) = P\{X = x_i\}$ 일 때, 누적분포함수 $F(x) = P\{X \leq x\} = \sum_{x_i \leq x \text{인 모든 } i} p(x_i)$.

그때, $[0, 1)$ 사이의 난수값 u 를 생성한 후, $u \leq F(x_j)$ 가 되는 가장 작은 양의 정수값 J 를 결정하고, $X = x_J$ 를 확률변수값으로 하게 된다. 따라서, $i = 1$ 일때는 $u \leq F(x_1) = p(x_1)$ 이면 확률변수값은 $X = x_1$ 이 되고, $i \geq 2$ 일때는 $F(x_{i-1}) < u \leq F(x_i)$ 이면 확률변수값은 $X = x_i$ 가 된다.

네트워크에 있는 활동들의 수를 n 이라 할 때, 제안된 절차는 다음과 같다.

단계 0. $\Delta = 0$ 라 둔다.

단계 1. 각 활동 $i (i = 1, \dots, n)$ 의 3개의 수준값 (l_{i1}, l_{i2}, l_{i3})을 식(4)와 같이 결정한다.

단계 2. $l_{i1} < 0$ 이면, 단계 3으로 간다. 그렇지 않으면 단계 4로 간다.

단계 3. 각 활동시간 분포의 도메인이 비음의 값들만을 포함하는 경우에는 $\Delta = \Delta + d$ 라 두고 단계 1로 간다. 그렇지 않으면 $l_{i1} = 0$ 라 두고 단계 4로 간다.

단계 4. 단계 2 또는 단계 3에서의 최종적인 Δ 값을 가지고 식(5)를 사용하여 가중치를 결정한다.

단계 5. 각 활동 $i (i = 1, \dots, n)$ 에 대해 $[0, 1)$ 사이의 난수값 u 를 생성한 후, 각 활동의 활동시간 t_i 를 다음과 같이 결정한다.

$$u \leq w_{i1} \text{ 이면 } t_i = l_{i1},$$

$$w_{i1} < u \leq w_{i1} + w_{i2} \text{ 이면 } t_i = l_{i2}$$

$$\text{그렇지 않으면 } t_i = l_{i3}$$

단계 6. 단계 5에서 결정된 t_i 값들을 가지고, CPM(Critical Path Method)을 이용하여 사업완성시간 Y_i 를 계산한다.

단계 7. 단계 5와 단계 6을 N 번 반복한다.

단계 8. 사업완성시간의 평균, 분산, 왜도, 첨도를 다음과 같이 계산한다.

평균: $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$

분산: $\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2$

왜도: $\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^3 / \left[\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 \right]^{3/2}$

첨도: $\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^4 / \left[\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 \right]^2$

Fig. 2에 있는 예제 네트워크에 대해 제안된 절차를 예증하면 다음과 같다. 각 활동시간은 화살표 위에 표시된 평균을 갖는 지수분포를 따른다고 가정한다. 그리고 $d = 0.01$, $N = 100,000$ 라 둔다.

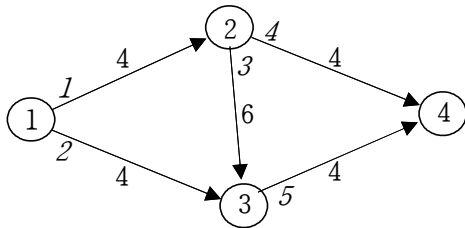


Fig. 2 An Example Network 1. (The numbers on the arcs represent mean activity durations, and italic numbers represent activity numbers.)

단계 1. 활동 1, 2, 4, 5의 최초의 3개의 수준값은 (-1.7980, 4.0, 17.7980), 그리고 활동 3의 최초의 3개의 수준값은 (-2.6969, 6.0, 26.6969).

단계 2. 활동 1~5에 대해 단계 3으로 감.

단계 3. 활동 1, 2, 4, 5에 대해 $\Delta = 1.08$ 에서 3개의 수준값은 (0.0082, 5.08, 18.9364), 그리고 활동 3에 대해 $\Delta = 1.61$ 에서 3개의 수준값은 (0.0023, 7.61, 28.3946).

단계 4. 활동 1, 2, 4, 5에 대해 가중치는 (0.3347, 0.6207, 0.0446), 그리고 활동 4에 대해 가중치는 (0.3336, 0.6218, 0.0446).

단계 5. 활동 1에 대해, $0 \leq u \leq 0.3347$ 이면 $t_1 = 0.0082$, $0.3347 < u \leq 0.9554$ 이면 $t_1 = 5.08$, $0.9554 < u < 1$ 이면 $t_1 = 18.93643$ 를 선택한다. 나머지 활동들에 대해서도 유사한 방법으로 t_i ($i = 2, 3, 4, 5$)를 결정한다.

단계 6. $t_1 = 0.0082$, $t_2 = 5.08$, $t_3 = 7.61$, $t_4 = 5.08$, $t_5 = 0.0082$ 일 때, $Y_1 = 7.6264$.

단계 7. 단계 5와 단계 6을 100,000번 반복한다.

단계 8. 사업완성시간의 평균, 분산, 왜도, 첨도는 Table 2의 “Exponential”행과 “Network 1”열에 보여진다.

4. 계산적 결과 및 분석

제안된 방법의 성능을 정확도와 계산적 효율 관점에서 평가하기 위하여, Fig. 2~Fig. 4에 있는 예제 네트워크들에 대해 계산적 실험을 수행한다. Fig. 2~Fig. 4에서 화살표 위에 있는 숫자들은 평균 활동시간을 나타낸다. 제안된 방법을 적용함에 있어, Table 1에 나타난 활동시간의 분포 및 모수들이 사용되며, Bowman[14]의 연구에서 사용된 분포 및 모수들과 동일하다.

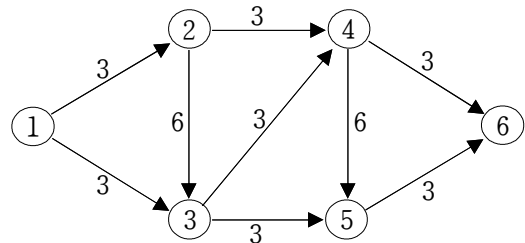


Fig. 3 An Example Network 2. (The numbers on the arcs represent mean activity durations.)

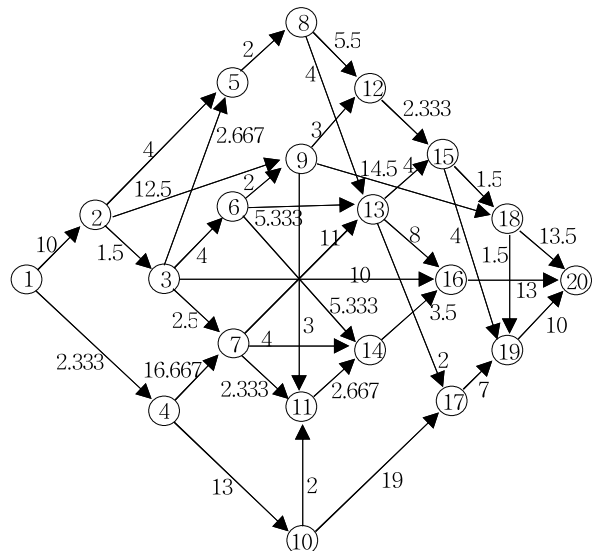


Fig. 4 An Example Network 3. (The numbers on the arcs represent mean activity durations.)

Table 1 Distributions and Parameters

Distributions	Parameters
Exponential(λ)	Scale $\lambda = 1/\text{평균}$
Gamma(α, β)	Scale $\alpha \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ Shape $\beta = \text{평균}/\alpha$
Normal(μ, σ^2)	$\mu = \text{평균}, \sigma = u\mu, u \sim U(1/10, 1/3)$
Uniform(a, b)	$a = (1-u) \times \text{평균}, b = (1+u) \times \text{평균},$ $u \sim U(0.2, 1.0)$
Beta(α, β)	Shape $\alpha, \beta \in \{1, 2, 3\}$
Triangular(a, c, b)	$c = 3 \times \text{평균} / \{(1-u_1) + 1 + (1+u_2)\},$ $a = (1-u_1)c, b = (1+u_2)c,$ $u_1 \sim U(0.5, 1.0), u_2 \sim U(0.5, 2.0)$

총 18개의 경우(3개의 네트워크와 6개의 분포)에 대해, 샘플크기 N=100,000의 제안된 방법에 의해 추정된 적률과 CPU 시간은 Table 2와 같다. 제안된 방법의 결과를 비교하기 위하여, 샘플크기 100,000의 몬테칼로 시뮬레이션에 의해 얻어진 결과도 Table 2에 보여진다.

Table 2 Moments of the Project Completion Time

Distribution	Network 1		Network 2		Network 3	
	Proposed method	MCS	Proposed method	MCS	Proposed method	MCS
Exponential	15.42 ⁽¹⁾	15.46	22.91	22.82	79.81	81.28
	61.96 ⁽²⁾	61.62	86.75	89.58	681.87	618.75
	1.32 ⁽³⁾	1.20	1.18	1.09	0.98	0.95
	5.24 ⁽⁴⁾	5.30	5.02	4.98	3.56	4.59
	409 ⁽⁵⁾	396	766	791	3180	3215
Gamma	16.13	16.05	23.91	23.64	71.72	72.08
	93.91	96.13	131.93	137.08	258.48	236.08
	1.66	1.42	1.38	1.25	0.67	0.79
	6.34	6.18	5.60	5.55	3.48	4.28
	400	1208	794	1929	3217	8017
Normal	13.99	14.01	21.00	20.99	55.72	55.72
	5.33	5.25	1.77	1.76	16.68	16.07
	0.01	0.03	0.01	0.01	0.47	0.45
	2.98	2.91	3.01	3.01	3.09	3.53
	383	506	711	929	3124	4089
Uniform	14.03	14.03	21.03	21.04	60.29	60.27
	9.36	9.38	14.39	14.11	41.18	41.33
	0.06	0.07	0.01	0.03	0.09	0.10
	2.39	2.36	2.20	2.19	3.00	3.03
	404	427	722	722	3210	3103
Beta	1.66	1.65	2.96	2.97	5.26	5.26
	0.12	0.12	0.27	0.27	0.32	0.33
	0.22	0.24	0.01	0.00	0.08	0.08
	2.71	2.73	2.80	2.79	2.90	2.89
	384	4336	717	7717	3254	33041
Triangular	14.00	14.00	21.01	21.00	60.07	60.09
	6.82	6.89	11.03	10.96	47.13	46.49
	0.19	0.19	0.19	0.19	0.15	0.17
	2.77	2.75	2.84	2.84	2.75	2.89
	394	415	718	735	3235	3256

(1) mean (2) variance (3) skewness (4) kurtosis (5) CPU time (in milliseconds)

제안된 방법과 몬테칼로 시뮬레이션 프로그램들은 Microsoft Visual C++로 코딩되었으며, Intel Core i3-4150 3.5 GHz CPU, 4.0 GB RAM 사양의 PC에서 실행되었다.

제안된 방법의 정확도를 몬테칼로 시뮬레이션과 비교하기 위하여, Table 2에 나타난 적률 추정치로부터 다음과 같은 차이를 계산한다.

- 평균, 분산, 첨도에 대해서는 백분율 차이
즉, $\frac{|\text{제안된 방법} - \text{몬테칼로 시뮬레이션}|}{\text{몬테칼로 시뮬레이션}} \times 100\%$

- 왜도에 대해서는 차이
즉, $|\text{제안된 방법} - \text{몬테칼로 시뮬레이션}|$

그 결과는 Table 3과 같다. 총 18개의 경우에 대해 (Table 3에서 'Distribution'열의 'All'행), 평균, 분산, 왜도, 첨도의 차이의 평균값은 각각 0.31%, 2.45%, 0.05, 3.95%로 상당히 작으며, 최대값도 각각 1.80%, 10.20%, 0.24, 22.44%로 작아서, 제안된 방법에 의한 적률 추정치는 몬테칼로 시뮬레이션에 의한 적률 추정치와 서로 잘 일치함을 볼 수 있다.

Table 3 The Accuracy of the Proposed Method

Distribution	Moments	Average	Max
Exponential	mean (%)	0.83	1.80
	variance (%)	4.67	10.20
	skewness	0.08	0.12
	kurtosis (%)	8.12	22.44
Gamma	mean (%)	0.70	1.10
	variance (%)	5.20	9.50
	skewness	0.16	0.24
	kurtosis (%)	7.39	18.69
Normal	mean (%)	0.03	0.10
	variance (%)	1.97	3.80
	skewness	0.01	0.02
	kurtosis (%)	4.96	12.46
Uniform	mean (%)	0.00	0.00
	variance (%)	0.87	2.00
	skewness	0.01	0.02
	kurtosis (%)	0.91	1.27
Beta	mean (%)	0.30	0.60
	variance (%)	1.00	3.00
	skewness	0.01	0.02
	kurtosis (%)	0.48	0.73
Triangular	mean (%)	0.00	0.00
	variance (%)	1.00	1.40
	skewness	0.01	0.02
	kurtosis (%)	1.86	4.84
All	mean (%)	0.31	1.80
	variance (%)	2.45	10.20
	skewness	0.05	0.24
	kurtosis (%)	3.95	22.44

다음으로, Table 2의 CPU 시간들을 살펴보면, 각 네트워크 대해 제안된 방법의 CPU 시간들은

분포별로 유사하나 몬테칼로 시뮬레이션의 CPU 시간들은 그렇지 않음을 볼 수 있다. 그 이유는 다음과 같다. 제안된 절차와 몬테칼로 시뮬레이션 절차는 (a) 각 활동의 활동시간을 결정하고 (3절의 제안된 절차에서 단계 0 ~ 단계 5), (b) CPM을 사용하여 사업완성시간을 계산하고, (c) 사업완성시간의 적률을 계산하는 단계들로 이루어져 있다. 제안된 절차에서는 단계 (a)를 수행하는데 필요한 CPU 시간은 연속형의 분포함수를 이산화하는데 소요되는 시간과 [0, 1) 사이의 난수를 사용하여 활동들의 활동시간을 결정하는데 소요되는 시간의 합과 같다. 그런데 연속형의 분포함수를 이산화하는데 소요되는 시간이 전체 CPU 시간에서 차지하는 비율은 매우 작다 (예를 들어, 총 18개의 경우에 대해 그 비율은 0.65% 보다 작음). 그래서, 단계 (a)를 수행하는데 필요한 CPU 시간의 대부분은 [0, 1) 사이의 난수를 생성하는데 소요된다. 반면에, 몬테칼로 시뮬레이션 절차에서는 단계 (a)를 수행하는데 필요한 CPU 시간은 주어진 확률분포로부터 확률변수값을 생성하는데 소요되는 시간과 같다. 그리고 제안된 절차와 몬테칼로 시뮬레이션 절차에서 단계 (b)와 단계 (c)는 동일하며, 활동시간의 분포와는 무관하다. 따라서, 제안된 절차의 CPU 시간들은 활동시간의 분포와는 상관없이 유사하나, 몬테칼로 시뮬레이션의 CPU 시간들은 활동시간의 분포에 따라 차이를 보인다. 결과적으로, 제안된 방법과 몬테칼로 시뮬레이션의 계산적 효율성은 [0, 1) 사이의 난수를 생성하는데 소요되는 시간과 확률분포로부터 확률변수값을 생성하는데 소요되는 시간의 차이에 의해 결정된다.

뿐만 아니라, Table 2를 보면 제안된 방법의 CPU 시간들은 몬테칼로 시뮬레이션의 CPU 시간들 보다 작거나 비슷함을 알 수 있다. 제안된 방법의 계산적 효율성을 몬테칼로 시뮬레이션과 비교하기 위하여, 다음과 같은 비(Ratio)를 계산한다.

$$\frac{\text{제안된 방법의 의한 CPU 시간}}{\text{몬테칼로 시뮬레이션에 의한 CPU 시간}} \times 100\%$$

그 결과는 Table 4와 같다. 지수, 균등, 삼각분포의 경우에는 제안된 방법의 CPU 시간들과 몬

테칼로 시뮬레이션의 CPU 시간들이 비슷하나, 감마와 정규분포의 경우에는 제안된 방법이 몬테칼로 시뮬레이션 보다 계산적인 효율성이 높음을 알 수 있다. 특히, 베타분포의 경우에는 제안된 방법의 계산적 효율성이 몬테칼로 시뮬레이션에 비해 매우 높음을 알 수 있다.

Table 4 Ratio of CPU Time of Proposed Method to that of MCS (%)

Distribution	Network 1	Network 2	Network 3
Exponential	103.3	96.8	98.9
Gamma	33.1	41.2	40.1
Normal	75.7	76.5	76.4
Uniform	94.6	100.0	103.4
Beta	8.9	9.3	9.8
Triangular	94.9	97.7	99.4

5. 결론

본 논문에서는 확률적 활동 네트워크의 분석을 위해 사업완성시간의 적률을 추정하기 위한 방법을 제안하였다. 제안된 방법은 근사적 방법과 몬테칼로 시뮬레이션이 결합된 방법이며, 활동시간의 분포함수를 이산화하기 위하여 적률매칭 방법을 사용하며, 사업완성시간의 계산에 필요한 활동시간값을 결정하기 위해 이산형 역변환 방법을 사용한다.

몬테칼로 시뮬레이션과 마찬가지로 제안된 방법은 활동시간의 분포로서 특정한 분포가 아닌 일반적인 분포에 대해서도 적용 가능하며, 대규모의 활동 네트워크에 적용이 쉽다. 그리고 계산적 실험 결과, 제안된 방법의 결과는 몬테칼로 시뮬레이션에 의해 얻어진 결과와 잘 일치함을 보여준다. 뿐만 아니라, 지수, 균등, 삼각분포의 경우에는 제안된 방법의 CPU 시간들과 몬테칼로 시뮬레이션의 CPU 시간들이 비슷하나, 감마와 정규분포의 경우에는 제안된 방법이 몬테칼로 시뮬레이션 보다 계산적 효율성이 높음을 알 수 있다. 특히, 베타분포의 경우에는 제안된 방법의 계산적 효율성이 몬테칼로 시뮬레이션에 비해 매우 높음을 알 수 있다.

References

- [1] Cho, J. G., "Estimating the Moments of the Project Completion Time in Project Networks," *Journal of the Korea Industrial Information Systems Research*, Vol. 22, No. 1, pp. 61-67, 2017.
- [2] Yao, M. J. and Chu, W. M., "A New Approximation Algorithm for Obtaining the Probability Distribution for Project Completion Time," *Computers and Mathematics with Applications*, Vol. 54, pp. 282-295, 2007.
- [3] Sculli, D., "The Completion Time of PERT Networks," *Journal of the Operational Research Society*, Vol. 34, No. 2, pp. 155-158, 1983.
- [4] Dodin, B., "Approximating the Distribution Functions in Stochastic Networks," *Computers and Operations Research*, Vol. 12, No. 3, pp. 251-264, 1985.
- [5] Dodin, B. and Sirvanci, M., "Stochastic Networks and the Extreme Value Distribution," *Computers and Operations Research*, Vol. 17, No. 4, pp. 397-409, 1990.
- [6] Mehrotra, K., Chai, J. and Pillutla, S., "A Study of Approximating the Moments of the Job Completion Time in PERT Networks," *Journal of Operations Management*, Vol. 14, pp. 277-289, 1996.
- [7] Abdelkader, Y. H., "Evaluating Project Completion Times when Activity Times are Weibull distributed," *European Journal of Operational Research*, Vol. 157, pp. 704-715, 2004.
- [8] Yum, B. J., Kim, S. J., Seo, S. K., Byun, J. H. and Lee, S. H., "The Taguchi Robust Design Method: Current Status and Future Directions," *Journal of the Korean Institute of Industrial Engineers*, Vol. 39, No. 5, pp. 325-341, 2013.
- [9] Lee, S. H. and Chen, W., "A Comparative Study of Uncertainty Propagation Methods for Black-Box-Type Problems," *Structural and Multidisciplinary Optimization*, Vol. 37, pp. 239-253, 2009.
- [10] Taguchi, G., "Performance Analysis Design," *International Journal of Production Research*, Vol. 16, pp. 521-530, 1978.
- [11] D'Errico, J. R. and Zaino Jr., N. A., "Statistical Tolerancing Using a Modification of Taguchi's Method," *Technometrics*, Vol. 30, No. 4, pp. 397-405, 1988.
- [12] Seo, H. S. and Kwak, B. M., "Efficient Statistical Tolerance Analysis for General Distributions Using Three-Point Information," *International Journal of Production Research*, Vol. 40, pp. 931-944, 2002.
- [13] Law, A. M. and Kelton, W. D., *Simulation Modeling and Analysis*, New York: McGraw-Hill, 1982.
- [14] Bowman, R. A. "Efficient Estimation of Arc Criticalities in Stochastic Activity Networks," *Management Science*, Vol. 41, No. 1, pp. 58-67, 1995.



조재균 (Jae-Gyeun Cho)

- 정회원
- 연세대학교 응용통계학과 학사
- 한국과학기술원 산업공학과 석사
- 한국과학기술원 산업공학과 박사
- 한국전자통신연구원 선임연구원
- 현재: 동의대학교 정보경영학부 교수
- 관심분야 : 프로젝트 관리, e비즈니스 모델 및 전략