

시뮬레이션 일정기법 - 최종공사기간의 확률 통계적 특성 추정

Probability Distribution of Project Completion Times in Simulation based Scheduling

이 동 은*
Lee, Dong-Eun

김 룰 희**
Kim, Ryul-Hee

요 약

기존의 시뮬레이션 일정기법은 최종공사기간(Project Completion Times: PCTs)이 정규분포를 따른다는 가정을 전제로 한다. 그러나 본 논문에서는 이 가정이 항상 옳은 것이 아니며, 이것이 잘못된 결과를 초래할 수 있다는 것을 검증한다. 이처럼 의문이 제기되지 않고 받아들여져 온 가정이 시뮬레이션 분석 결과에 어떠한 영향을 줄 수 있는지를 밝혀내는 리스크 정량화기법(risk quantification method)을 MATLAB 알고리즘으로 구현하였으며, 네트워크의 모델링에서부터 시뮬레이션 출력 값들로 구성된 샘플집단들에 대한 분석에 이르기까지 전 단계를 MATLAB 프로그래밍으로 구현된 알고리즘을 사용하여 제기된 의문에 대한 답을 제시하였다. 특정 네트워크를 구성하는 액티비티 기간 값들을 정의하는 확률분포함수의 종류를 다양하게 변화시켜 시뮬레이션 결과 값들 - 최종공사기간 값들 - 을 생성하고, 이처럼 생성된 시뮬레이션 출력 값들로 구성된 샘플집단들의 확률 통계적 특성을 분석하였다. 본 연구는 시뮬레이션을 기반으로 하는 일정관리기법의 신뢰성을 향상시키며, 일정관련 리스크 분석의 정확성을 향상시키는데 기여할 것이다.

키워드: 시뮬레이션, 공정관리, 확률통계이론, 공사완료기간, 퍼트, 매트랩

1. 서 론

1.1 연구배경 및 목적

건설 분야에서 사용되고 있는 공정관리기법들 중 특히 시뮬레이션 일정관리기법은 최종공사기간(최종공사비)을 확률 통계적으로 취급하여 신뢰성을 향상시키고, PERT기법이 지닌 한계점들을 보완한다는 점에서 가치를 인정받아왔다. 그러나 PERT기법은 최종공사기간의 확률을 계산할 때 정규분포를 사용하는 한계점을 지닌다. 또한, 시뮬레이션 일정관리기법과 관련된 선행 연구들은 (Ang and Tang 1975, Halpin and Riggs 1992, Lu and AbouRizk 2000) 정규분포의 평균과 표준편차를 계산하는 것이 계산상 편리하기 때문에 시뮬레이션 출력값, 즉 최종공사기간값이 정규분포를 따른다는 가정을 반복해서 사용해왔고, 이러한 가정을 토대로 시뮬레이션 출력데이터를 분석해왔다.

따라서 본 연구는 30년이 넘도록 시뮬레이션 일정관리 연구에서 지속적으로 반복 사용되어온 가정, 즉 시뮬레이션 출력 데이터가 정규분포를 따른다는 가정으로 인해 시뮬레이션 출력 데이터 분석기술이 정체된 상태에 있어왔음을 검증하고 효과적인 개선 방법론을 찾아 구현하는 것을 목적으로 하였다.

1.2 연구 방법

최종공사기간의 비정규분포 특성을 실증적으로 검증하고 리스크 정량화 절차를 효과적으로 제시하며 향후 다른 연구자들에 의해 비교연구를 할 수 있는 기회를 열어놓기 위해 이전 연구들에 사용되어온 네트워크(그림 2)를 모의실험하였다. 해당 네트워크를 Matlab을 사용하여 시뮬레이션 모델링을 하였고, 각 액티비티 기간을 (1)CPM 기법에 의한 결정론적 기간, (2)PERT 기법에 의한 확률적 기간 및 (3)Simulation기법에 의한 통계적 확률함수들로 정의하였다. 그런 다음, 각 기법들에 의해 생성된 최종 공사기간들로 구성된 샘플집단들을 3.2항에서 제시한 알고리즘을 사용하여 분석하였다.

2. 문헌고찰

시뮬레이션 일정관리기법을 설명하는 문헌들 중 다수는 최종공사기간에 대해 다루고 있고 (Sculli 1983, Dodin and Sirvanci 1990, Barraza et al 2004, Lee 2005, Lee and Arditi 2006), 어떤 연구들은 최종공사기간의 확률분포함수를 결정하는 문제를 다룬다(Cottrell 1999, Barraza et al 2004). 이러한 연구들은 3가지 영역, 즉 수학적 기법(e.g., Dodin 1985, Fisher et al 1985, Kulkarni and Adlakha 1986, and Hagstrom 1990), 근사적 기법(e.g., Sculli 1983, Dodin 1985b, Sculli and Wong 1985, Golenko-Ginzburg 1989, and Dodin and Sirvanci 1990), 시뮬레이션 기법(e.g., Sculli 1983, Barraza et al 2004, Lee 2005, Lee and Arditi 2006)들로 분류될 수 있다.

* 일반회원, 경북대학교 건설공학부 전임강사, 공학박사 (교신저자), dolee@knu.ac.kr

** 일반회원, 경북대학교 대학원, 석사과정, archi_ryul@knu.ac.kr

3. 방법론

3.1 액티비티 기간의 확률분포함수 정의

다양한 확률분포함수를 사용하여 액티비티 기간을 정의함에 따라 최종공사기간이 어떠한 확률 통계적 특성을 띄는지 분석하였다. 액티비티 기간은 Matlab의 난수생성기를 사용하여 특정 확률분포함수로 정의되었고, 각각의 확률분포함수에 대해 30,000개씩의 액티비티 기간을 생성한 후 샘플집단의 확률분포함수를 추정하였다.

3.2 최종공사기간 확률분포함수 추정 알고리즘

그림1에 제시된 시뮬레이션 모델링 및 분석 알고리즘은 다양한 확률분포함수로 액티비티 기간을 정의하고 모의실험을 실행하며, 서로 다른 확률분포함수를 사용했을 때 발생하는 리스크를 정량적으로 추정하는 것을 용이하게 한다. 본 알고리즘은 최종공사기간의 확률분포의 추정과 관련된 리스크를 과학적 실험에 근거하여 정량적으로 취급한다. 이를 사용하여 시뮬레이션 출력 데이터의 확률분포 특성을 파악하는 것은 시뮬레이션 출력 데이터 분석기술을 진보시킨다. 알고리즘에 대한 자세한 설명은 다음과 같다.

- **단계 ①:** 일정 네트워크가 AOA(Activity on Arrow)기법을 활용하여 모형화된다. 이벤트 시간값 및 액티비티 기간값들이 정의된다. 이벤트 시간 값들은 0으로 초기화되며, 액티비티 기간값들은 최빈값에 해당하는 결정론적 값들로 초기화된다.
- **단계 ②:** 액티비티들 사이의 선후행 관계가 단계 ①에서의 변수들이 계산되는 순서를 정의함으로 확립된다.
- **단계 ③:** 시스템 사용자에게 의해 할당된 시간은 결정론적 CPM값, 확률적 PERT-beta값 및 7가지 확률분포함수의 인수 값들로 변환되어 정의된다(1=Beta, 2=Exponential, 3=Normal, 4=Lognormal, 5=Uniform, 6=Triangular, 7=Weibull, 8=CPM, 9=PERT-Beta).
- **단계 ④:** 기간 추정 값들은 초기에 CPM계산을 위해 결정론적인 기간이 설정된다. 그러나 COV가 20%라는 가정을 사용하여, 7가지 확률분포 함수들에 상응하는 모수들을 생성한다.
- **단계 ⑤:** 최대 축차 반복횟수가 시스템 사용자에게 의해 설정되며, 현재 축차반복 횟수는 0으로 설정된다.
- **단계 ⑥:** 액티비티 기간들이 단계 ③에서 선택된 확률분포 함수로부터 생성된다.
- **단계 ⑦:** 단계 ⑥에서 생성된 액티비티 기간들을 사용하여 전진계산알고리즘이 실행된다. 최댓값을 계산하는 MAX 함수를 사용하여 후행 액티비티의 이른 시작시간 (Early Start time: ES) 을 찾는다.
- **단계 ⑧:** 최종 공사기간은 매 축차반복에 있어서 마지막 노드의 이벤트 시간과 동일하다. 지정된 축차반복 횟수만큼 시뮬레이션이 진행되는 동안, 최종공사기간 값들이 계산되고 하나의 벡터에 저장된다.
- **단계 ⑨:** 사용자에게 의해 단계 ⑤에서 설정된 최대 축차반복횟수가 현재 축차반복 횟수보다 큰지를 비교하고, 단계 ⑥에서 ⑧까지 지정된 횟수만큼 반복한다.

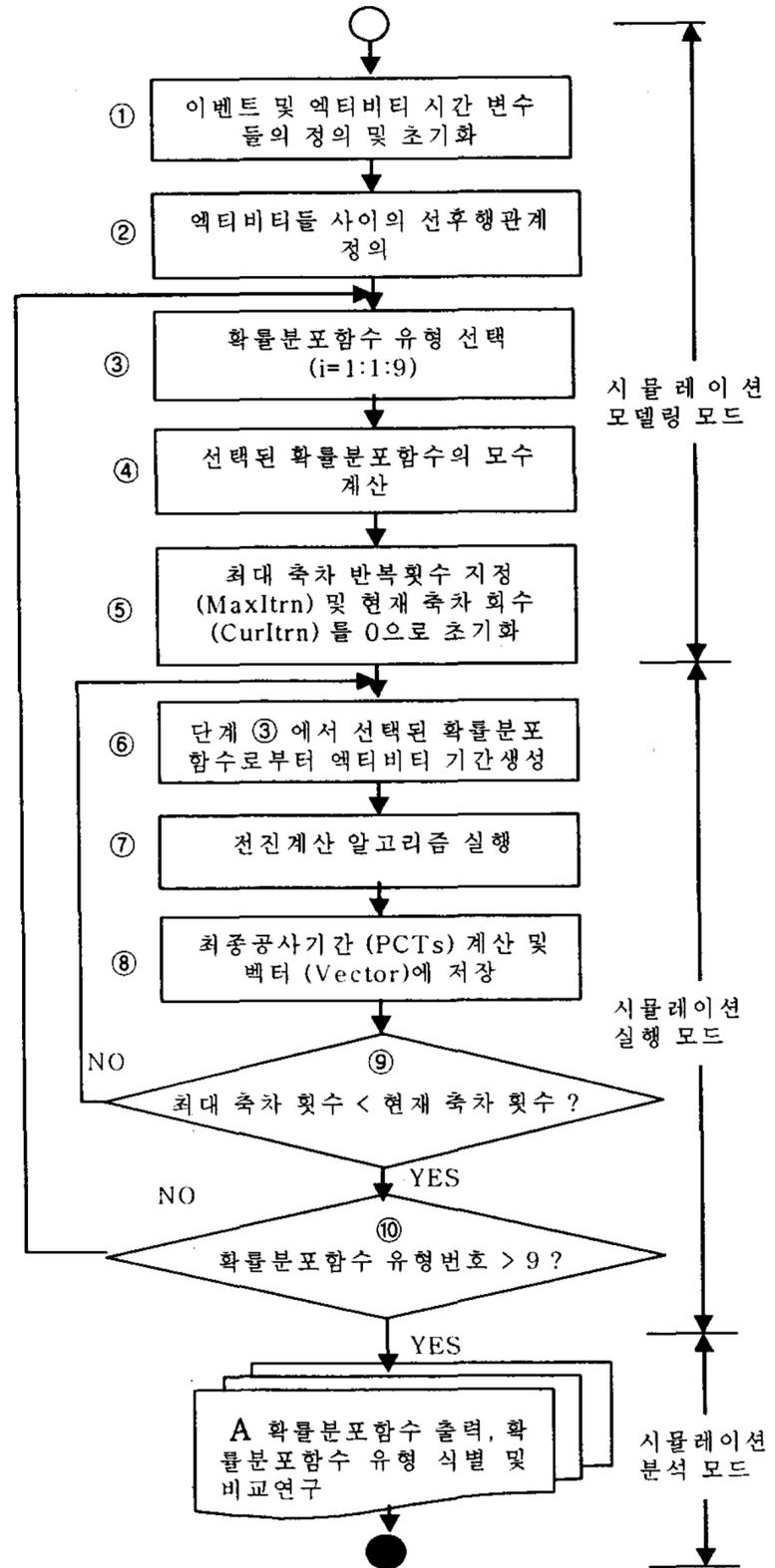


그림1. 시뮬레이션을 활용한 최종공사기간의 확률분포함수를 추정하는 알고리즘

- **단계 ⑩:** 확률분포함수 유형과 관련하여 사용자의 선호 혹은 필요에 따라, 어떤 분석에 적용될 확률분포함수의 유형이 확장될 수 있다.
- **단계 A:** 최종공사기간을 가장 적합하게 묘사하는 확률분포함수의 유형 및 통계치(인수값 포함)를 확률분포 적합성실험을 사용하여 계산하고, 결정론적 CPM에 의해 계산된 기간이내에 프로젝트를 완성할 확률을 계산한다.

4. 모의실험

그림2의 네트워크를 사용하여 시뮬레이션 절차와 방법론을 모의 실험 하였다. 액티비티 기간은 지수분포함수로 정의하였으며, 할당된 숫자는 최빈값이다. CPM의 경우 최빈값들이 그대로 사용되었고, PERT의 경우는 COV가 20%라는 가정 하에 낙관 및 비관적 기

간 값들을 계산하였다. 또한 시뮬레이션 기법의 경우는 확률분포함수의 유형에 따라 COV가 20%라는 가정을 사용하여 모수추정을 수행하였다. 상기의 3가지 기법들을 사용하여 최종공사기간을 묘사하

는 확률분포함수 유형 및 인수값을 추정하였다. 액티비티 기간들을 결정론적 CPM값, 확률적 PERT값, 그리고 확률분포함수로 정의한 후, 최종 공사기간이 어떠한 거동을 나타내는지 분석하였다.

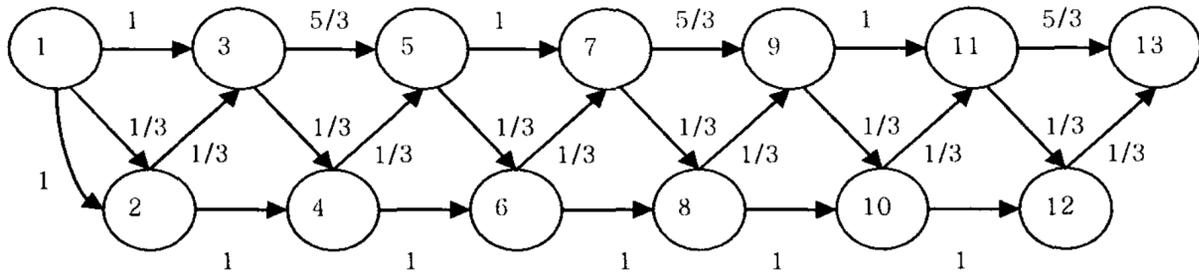


그림 2. 실험 1의 네트워크 데이터 - 지수분포로 정의된 액티비티들 (Abdelkader 2004 재사용)

우선, 7가지 확률분포함수를 사용하여 액티비티 기간을 정의한 후 30,000회의 시뮬레이션을 실행하여 7개의 샘플집단들을 생성하였다. 표1은 CPM, PERT, 그리고 다양한 확률분포함수들을 사용하여 시뮬레이션 결과를 제시하고 있다. 시뮬레이션기법을 적용한 경우, 다음의 두 가지 방법을 사용하여 통계치를 추정하였다. (1) 최종 공사기간을 정규분포로 가정하는 방법과 (2)확률분포 적합성 실험을 통해 최적 확률분포함수를 찾아 확률을 구하는 방법. 또한 결정론적 CPM기법에 의해 계산된 8.3333내에 끝낼 확률을 계산

하였다. 표 1에 제시된 것처럼, (1) CPM은 8.3333, (2)PERT는 평균 = 8.3333; 표준편차 = 1.1785를 계산한다. 개략적으로 최종공사기간의 확률분포함수들은 3가지 그룹 (예, 그룹A - Normal, Lognormal, and Beta, 그룹B - Uniform and Triangular, and 그룹C - Weibull and Exponential)으로 분류될 수 있었다. 그룹 A, B 그리고 C에 속하는 확률분포함수들이 액티비티 기간을 정의하기 위해 사용했을 때, 각 그룹에 속하는 확률분포함수들은 유사한 통계적 거동을 나타내었다.

방법	액티비티 확률분포함수	정규분포 가정의 경우 최종공사기간의 통계치			적합도가 최적인 확률분포함수로 추정된 통계치						
		확률분포함수 종류	평균	표준편차	8.3333내에 끝낼 확률	최적 확률분포함수 종류	최후 추정값	평균	표준편차	8.3333내에 끝낼 확률	
시뮬레이션	그룹 A	Exponential	Normal	10.8058	3.0017	21.29%	Lognormal	-7580.53	10.7868	3.2305	23.83%
		Weibull	Normal	10.8442	3.4079	24.80%	Lognormal	-734.361	10.4494	2.9714	25.67%
	그룹 B	Uniform	Normal	8.62264	0.8691	36.96%	Normal	-3835.28	8.62264	0.8691	36.96%
		Triangular	Normal	8.47445	0.6308	41.15%	Normal	-2873.86	8.47445	0.6308	41.15%
	그룹 C	Lognormal	Normal	8.3754	0.3164	45.44%	Lognormal	-91.7136	8.40673	0.3294	41.91%
		Normal	Normal	8.3589	0.3214	46.82%	Normal	-84.6774	8.3589	0.3214	46.82%
	Beta	Normal	8.3380	0.3043	49.39%	Normal	-68.2551	8.3380	0.3043	49.39%	
PERT	PERT-Beta		8.3333	1.3889	50%			8.3333	1.1785	50%	
CPM	Deterministic		8.3333	-	-			8.3333	-	100%	

표.1 정규분포 가정의 경우 vs 적합도 최적 확률분포함수로 추정된 시뮬레이션 결과비교

그룹A (표1)에 제시된 것처럼, 액티비티 기간을 (1)지수 확률분포로 모델링했을 때 평균 = 10.8058; 표준편차= 3.0017, 그리고 (2)와이블 분포로 모델링했을 때 평균 = 10.8442; 표준편차= 3.4079라는 결과를 생성한다. 반면, 확률분포함수 적합도 시험에 의해 추정된 통계치는 두 분포 모두 최적 확률분포함수로 로그확률분포함수가 추정되었다. 반면, 확률분포함수 적합도 시험에 의해 추정된 통계치는 지수 확률분포의 경우 평균 = 10.7868; 표준편차= 3.2305, 와이블 분포의 경우 평균 = 10.4494; 표준편차= 2.9714라는 결과를 생성한다. 결정론적 CPM기법에 의해 얻은 8.3333이내에 프로젝트를 완료할 가능성을 기존 방법론(정규분포를 가정)에 의해 분석한 결과, 21.29%와 24.80%를 얻는 반면, 확률분포함수 적합도 시험에 의해 23.83%, 25.67%를 얻어 각각 2.54% 및 0.87%의 편차를 보인다. 무엇보다도 유의할 특징은 다른 그룹들에 비해 상대적으로 현저하게 보수적 특성을 나타낸다는 점이다.

그룹B (표1)에 제시된 것처럼, 기존 방법론 및 확률분포함수 적합도 시험에 의해 추정된 통계치 모두 액티비티 기간들을 (1)균등분포함수로 모델링했을 때 평균 = 8.62264; 표준편차= 0.8691, (2)삼각분포함수로 모델링했을 때 평균 = 8.47445; 표준편차= 0.6308라는 결과를 생성한다. 정규분포에 대한 적합도 시험(Lilliefors

test)을 시행한 결과 2개의 표본집단 모두 정규분포성을 나타낸다는 사실이 입증되었다. 그러나 T-test 결과는 두개의 서로 다른 확률분포함수로부터 얻은 결과들이 하나의 동일한 모집단내에 있는 것으로 받아들일 수 없음을 보였다. 실험결과 균등분포와 삼각분포함수를 사용하여 모델링된 네트워크는 최종공사기간이 균등분포 혹은 삼각분포가 아니라 정규분포에 근사하였다.

그룹C (표1)에 제시된 것처럼, 액티비티 기간들을 (1)정규분포함수로 모델링하였을 때, 기존 방법론에 의하면 평균 = 8.3589; 표준편차= 0.3214, (2)로그정규분포함수로 모델링하였을 때, 평균 = 8.3754; 표준편차= 0.3164, 그리고 (3)베타분포함수로 모델링하였을 때, 평균= 8.33809; 표준편차= 0.3043의 결과를 생성한다. 반면, 확률분포함수 적합도 시험에 의해 결과를 분석하면, 정규분포함수와 베타분포함수는 차이가 없지만 로그분포함수로 모델링된 경우 최종공사기간을 묘사하는 최적 확률분포함수가 로그확률분포함수가 최적으로 추정되며, 평균 = 8.40673; 표준편차= 0.3294라는 결과를 생성한다. 따라서 CPM기법에 의해 계산되어진 8.3333이내에 프로젝트를 완료할 가능성이 두 가지 방법의 경우 45.44%와 41.91%로 3.53%의 차이를 보인다.

그룹B의 확률분포함수 멤버들은 그룹C의 확률분포함수 멤버들

보다 더 폭넓은 분포도를 나타내며, 그룹B와 C는 대칭적 특성을 나타낸다. 반면 그룹A는 최종공사기간의 분포가 양왜도형(positively skewed)으로 비대칭 특성에 가깝다.

실험 I의 네트워크를 개발된 알고리즘(그림1)을 사용하여 수치 분석한 결과(표1), 프로젝트를 8.3333 이내에 끝낼 확률이 다음 순서에 따라 점차 감소하는 결과를 나타낸다. [CPM>PERT>Beta>Normal>Lognormal>Triangular>Uniform>Weibull>Exponential] = [100>50>49.39>46.82>41.91>41.15>36.96>25.67> 23.83%]. 특히 그룹A에 속하는 확률분포함수가 사용된 경우 최종공사기간의 확률분포함수는 항상 정규분포를 따르는 것으로 받아들이기 어렵다. 그러므로 시뮬레이션 연구에서 출력 값들이 정규분포를 따른다는 가정이 잘못된 결과를 유도할 수 있음이 입증되었고 최적 확률분포함수를 추정해야 할 필요성이 정당화되었다. 본 논문에 제시된 방법론을 사용하여 관련된 리스크를 정량화할 수 있다.

5. 발견 및 결론

본 연구를 통해 최종공사기간의 확률분포함수는 기존연구들에서 받아들여져 온 것처럼 항상 정규분포를 따르는 것은 아니라는 사실이 검증되었다. 또한, 기존의 시뮬레이션 일정관리기법 및 연구들이 편의상 시뮬레이션 출력 데이터가 정규분포를 따른다고 가정함으로써 정확성이 절충되었지만, 본 연구에서 개발된 방법론은 최적 확률분포함수를 추정하여 기존방법론을 보완함으로써 더 신뢰할만한 결과를 얻는데 기여하고 있다. 본 방법론은 시뮬레이션 출력 데이터를 묘사하는 확률분포함수를 추정함으로써 예측의 정확성을 높일 수 있기 때문에 건설 공정관리 실무 중 특히 입찰단계에서 그 힘을 드러낼 수 있다.

참고문헌

1. Abdelkader, H. Y., (2004). "Evaluating project completion times when activity times are Weibull distributed." *European Journal of Operational Research*. 157, 704-715.
2. Ang, A.H-S., and Tang, W. H.(1975). "Probability Concepts in Engineering Planning and Design - Analytical models of random phenomena." Vol. 1, Wiley, NY.
3. Barraza, A. Gabriel, Back, W. Edward, and Mata, Fernando, (2004). "Probabilistic Forecasting of Project Performance using Stochastic S Curves." *Journal of Construction Engineering and Management*, 130(1). 25-32.
4. Cottrell, D. W, (1999). "Simplified Program Evaluation and Review Technique (PERT)." *Journal of Construction Engineering and Management*, 125(1), 16-22.
5. Dodin, B.M, (1985b). "Approximating the Distribution Functions in Stochastic Networks", *Computers and Operations Research*. 12(3).251-264.
6. Dodin, B.M., (1985a). "Bounding the Project Completion Time Distribution in PERT Networks", *Operations Research*. 24(4). 862-881.
7. Dodin, Bajis and Mete Sirvanci, (1990)."Stochastic Networks and The Extreme Value Distribution", *Computers and Operations Research*. 17(4) 397-409.
8. Fisher, D.L., D. Saisi and W.M. Goldstein, (1985)."Stochastic PERT Networks: OP Diagrams, Critical Paths and the Project Completion Time.", *Computers and Operations Research*. 12(5), 471-0482.
9. Golenko-Ginzburg, Dimitri, (1989)."A New Approach to the Activity-Time Distribution in PERT.", *Journal of Operational Research Society*. 40(4), 389-393.
10. Hagstrom, J.N., (1990)."Computing the Probability Distribution of Project Duration in a PERT Network", *Networks*. Vol. 20, 231-244.
11. Halpin, D. W., and Riggs, L. S. (1992). *Planning and analysis of construction operations*, John Wiley & Son, Inc, New York. 270-274.
12. Lee, D.-E, and Arditi, D., (2006). "Automated Statistical Analysis in Stochastic Project Scheduling Simulation." *J. of Constr. Eng and Mangt. ASCE*. 132(3). 268-277.
13. Lee, D.-E. (2005)."Probability of Project Completion Using Stochastic Project Scheduling Simulation (SPSS)". *J. of Constr. Eng and Mangt. ASCE*. 131(3). 310-318.
14. Lu, M. and AbouRizk, S. M. (2000). "Simplified CPM/PERT Simulation Model." *J. Constr. Eng. Mgnt*, ASCE, 126(3), 219-226.
15. Sculli, D. (1983). "The Completion Time of PERT Networks." *The Journal of the Operational Research Society*, 34(2), 155-158.
16. Sculli, D. and K.L. Wong, (1985)."The Maximum and Sum of Two Beta Variables and the Analysis of PERT Networks.", *Omega*. 13(3), 233-240.

Abstract

This paper verifies that the normality assumption that the simulation output data, Project Completion Times (PCTs), follow normal distribution is not always acceptable and the existing belief may lead to misleading results. A risk quantification method, which measures the effect caused by the assumption, relative to the probability distribution of PCTs is implemented as an algorithm in MATLAB. To validate the reliability of the quantification, several series of simulation experiments have been carried out to analyze a set of simulation output data which are obtained from different type of Probability Distribution Function (PDF) assigned to activities' duration in a network. The method facilitates to find the effect of PDF type and its parameters. The procedure necessary for performing the risk quantification method is described in detail along with the findings. This paper contributes to improving the reliability of simulation based scheduling method, as well as increasing the accuracy of analysis results.

Keywords : Simulation, Scheduling, Stochastic Probability Theory, Project Completion Time, PERT, MATLAB