Determination of task temporal variations in construction scheduling using imprecise probability

Mehdi Modares,[1*] Rafi L. Muhanna,[2] and Robert L. Mullen[3]

[1] *Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, USA*
[2] *School of Civil and Environmental Engineering, Georgia Institute of Technology, USA*
[3] *Department of Civil and Environmental Engineering, University of South Carolina, USA*
[*] Corresponding author: mmodares@iit.edu

Abstract

In Critical Path Method (CPM) and Program Evaluation and Review Technique (PERT) for construction scheduling and monitoring, it is crucial to determine the allowable temporal variation for each task such that the project's total duration is not significantly exceeded. Using the task durations, the overall project completions as well as the allowable float in starting each task are calculated such that the project's overall completion date is not changed. However, as each event possesses uncertainties, deterministic schemes are not capable of quantification of those uncertainties. Moreover, the traditional probabilistic approaches require a significant level of available data to correctly obtain the probability distribution function for each task. As the available data is limited, these approaches may result in errors when estimating those variations.

In this work, a new formulation of CPM and PERT based on the concepts of imprecise probability is introduced. P-box formulation is employed to represent the imprecise probability; see Williamson and Downs (1990) [1], Ferson et al. (2003) [2]. This method calculates bounds on duration for each task through definition of an imprecise probability structured objective function leading to calculation of bounds on the total project duration. An example problem illustrating the application of this method is presented and compared to the use of calculations using more conventional probabilistic and interval approaches.

Keywords: Imprecise Probability, Construction Scheduling, Critical Path Methods

1 Introduction

The Critical Path Method (CPM) is an approach worldwide utilized for planning and monitoring of both industrial and construction projects. Using conventional CPM, the overall duration of a project is determined. Moreover, the project's critical tasks or activities are identified. The critical tasks are activities whose temporal variations (float times) may increase and affect the overall project duration.

However, the analysis performed in conventional CPM is deterministic and does not consider the uncertainties in the activities' durations. The early methods that consider the presence of uncertainty in project scheduling are mostly probabilistic methods which require a significant level of accuracy in the selection of the Probability Density Function (PDF). For instance, in Program Evaluation and Review Technique (PERT) method, the project durations are described in terms of three values, the minimum, expected and maximum duration (Moder et al. [3]). This method has an underlying probabilistic assumption of a Beta distribution for each activity which may exhibit complexities regarding the notion of skewness. Cottrell [4] assumed a normal distribution for each activity. This assumption leads to only two time parameters for each activity, the most likely time and the pessimistic time. However, this reduction in parameters results in an assumed symmetric distribution for each activity that may not be supported by measured data.

In recent years, iterative probabilistic (e.g. Monte Carlo simulations) and possibilistic (e.g. fuzzy) methods with general probability distributions given for the duration of each activity have been explored (Lu and AbouRizk [5]). In a Monte Carlo simulation approach for project scheduling, the probability distribution of project duration and the overall risk associated in meeting a scheduled project completion is determined. Using this approach, the duration of each activity is realized from a random number generated from the probability distribution prescribed for the activity and a deterministic CPM analysis completed. The procedure in repeated a significant number of times and the probability distribution for the project completion is computed. In most Monte Carlo simulations, the random variables associated with each activity are assumed to be independent. This may not be a realistic assumption because of the inherent dependency among project activities. For example, when production rates on an activity are adversely impacted by a weather event, the activity durations may be positively correlated. On the other hand, when there is a space constraint associated with two work teams in the same area, the activity durations may be negatively correlated.

In general, determination of the proper probability distribution for each activity requires large amounts of multi-activity data to construct the overall multidimensional probability distribution. An alternative to having sufficient data to construct the "correct" distribution is to estimate a range of distributions that represent possible activity duration. One of such methods for handling uncertainty in a system with no assumptions of a well-defined PDF is imprecise probability theory. Imprecise probability theory addresses the extraction of useful information about a process in the case that the probability distribution function (PDF) of system inputs or parameters may not be known. Contrasting to fuzzy set theory, as more information about the PDF is acquired, imprecise probability methods can sharpen its description of parameters; in the limit, of known PDF, imprecise probability reduces to conventional probability.

In this paper, the notion of imprecise probability theory is used to describe the duration of activities in a project. In particular, the activity duration is represented as a probability-box (pbox). A pbox is a set of cumulative distribution functions (CDF), the set of distributions is given by an upper and lower bounding CDFs are represented. Following that, the bounds on the total project duration are calculated.

1.1 Overview

This work is a symbiosis of two historically independent fields, project scheduling (using stochastic analysis), and the imprecise probability theory (based on P-box). Prior to introducing the developed method, descriptions of both fields are presented. Following that, the developed methodology for project scheduling using imprecise probability is introduced. The methodology is then illustrated using numerical examples.

2 Previous Applications of Stochastic Methods to CPM and PERT

To accurately perform project scheduling and planning, especially for a large-scale project, the time duration of each activity must be estimated with sufficient reliability. In order to achieve this reliability, numerous probabilistic methods have been utilized to represent and quantify the temporal uncertainty of each activity. One of those methods is to use Subjective Beta Distribution Fitting based on the available data for each activity (AbouRizk et al. [6]). Using this method, a beta distribution is fitted to an activity thru statistical approaches. In case the activities' end-points are known, the two distribution shape parameters are obtained using moment matching or maximum likelihood approaches; whereas, if the activities' end-points are unknown, the method of least squares can be used as the aforementioned methods may fail (AbouRizk et al. [7]).

In case there are not sufficient data available, subjective approaches (e.g. expert opinion) can be used. Using these approaches, the two shape parameters of the beta distribution can be calculated thru five subjective estimation methods. These five methods are based on using: 1) mean and variance, 2) mean and mode, 3) mode and variance, 4) mode and an arbitrary percentile, and 5) two arbitrary percentiles. The probability parameter (mean and variance) obtained from this method leads to accurate estimation of the PDF shape parameters that can be used in simplified CPM/PERT simulations (AbouRizk et al. [6]).

Another probabilistic approach for representing and quantifying uncertainty in fitting the beta distribution has been thru using "Three Time Estimates" that determine the activity duration used (McLaughlin and Pickhardt [8]). Those estimates include: 1) the "most optimistic" time, i.e., minimum time needed to perform an activity, 2) the "most pessimistic" time, i.e., maximum time needed to perform an activity, and 3) the "most likely" time i.e., normal time needed to perform an activity (occurrence estimation is based on the results from sufficient repetition).

3 Imprecise Probability Structures Based on P-boxes

3.1 Continuously Bounded P-boxes

Consider $F(x)$ as the Cumulative Distribution Function (CDF) for the random variable X. When the distribution parameters are uncertain, for every x, an interval $[\underline{F}(x), \overline{F}(x)]$ generally can be found to bound the possible values of $F(x)$, i.e., $\underline{F}(x) \leq F(x) \leq \overline{F}(x)$. Such a pair of two CDFs $\underline{F}(x)$ and $\overline{F}(x)$ construct a so-called *probability box* or *probability bounds* Ferson et al. [2].

Figure 1 shows the probability box for a normal distribution with an interval mean of $[2.0, 3.0]$ and a standard deviation of 0.5. In this simple example, it is easy to verify that $\underline{F}(x)$ is the CDF of the normal variable with a mean of 3 and $\overline{F}(x)$ is the one with a mean of 2. Probability box represents a general framework to represent imprecisely specified distributions. It can represent not only distributions with unknown parameters, but also distributions with unknown type or even unknown dependencies.

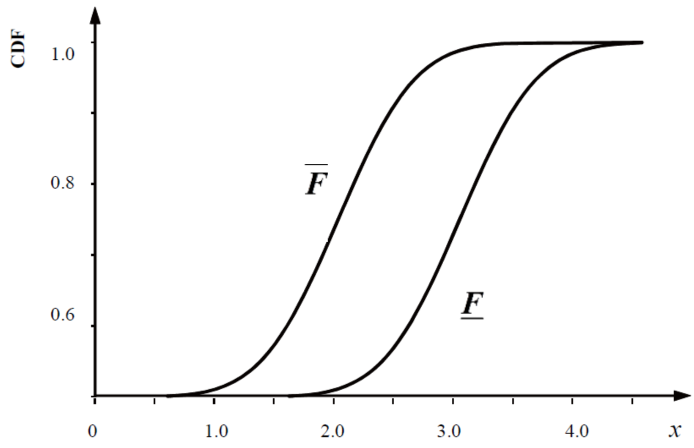

Figure 1: A probability box defined by a normal distribution with a mean of [2.0; 3.0] and a standard deviation of 0.5

3.2 Discrete Bounded P-boxes

An alternative to continuously bounded P-boxes, is the use of discrete, interval based P-box structures as a data type and a library of extensions of standard arithmetic operators on such structures. A discrete P-box structure consists of a collection of interval values, each of which has an associated probability.

In a uniform discretization, the associated probability for each interval is the same. For example, consider a P-box defined by bounding lognormal distribution with interval mean of [2.47, 11.08] and interval standard deviation of [2.76, 12.38]. In Figure 2, the continuous P-box and the interval discretization of the P-box into 10 intervals is shown for these bounding lognormal distributions. While the continuous distributions are not bounded to the right, the discrete P-box is truncated when the cumulative distribution is greater than a specified tolerance. In this example, a tolerance of 0.01 is used (Figures 2 and 3).

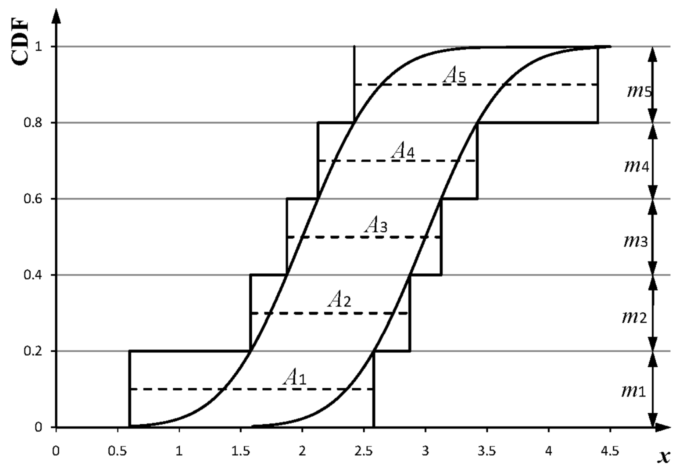

Figure 2: A P-Box defined by bounding lognormal distribution with mean of [2.47, 11.08] and standard deviation of [2.76, 12.38] (**Enclosure Discretization**)

While other researchers have used non-conservative mid-point discretization as shown in Figure 3, this paper uses a discretization that encloses the original P-box (Figure 2). The arithmetic of discrete P-Box structures is discussed in the work of Ferson et al. [2], Williamson and Downs [1], who gives detailed description of algorithms for arithmetic operators with either the assumption of independence between variables or the consideration of any-dependency between variables. Copulas can be employed to provide for other dependency (bounds) on random variables. (Ferson and Hajagos. [9]).

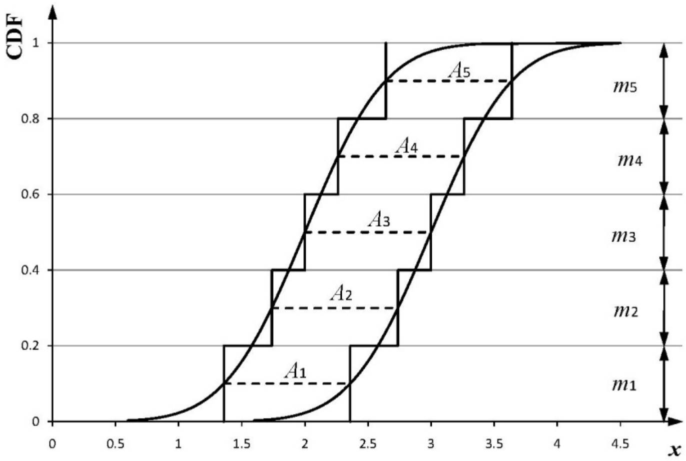

Figure 3: A P-Box defined by bounding lognormal distribution with mean of [2.47, 11.08] and standard deviation of [2.76, 12.38] (**Mid-point Discretization**)

4 Numerical Example - Simple Network for Building Construction

4.1 Problem Definition - Simple Network for Building Construction

This example calculation is based on a model from Ahuja et al. [10] (Chapter 15) which was used initially to illustrate a PERT analysis (Table 1). Figure 4 depicts the activity information schematically.

Table 1: Activity information for example problem

Activity	**Description**	**Start Node**	**End Node**	**Min Duration**	**Expected Duration**	**Max Duration**
A	Order/prefab metal building	0	2	20	22	25
B	Clear site Weather	0	1	5	10	15
C	Underground and foundation	1	2	5	10	16
D	Erect prefab building	2	3	8	10	20
E	Finish interior	3	4	9	10	11

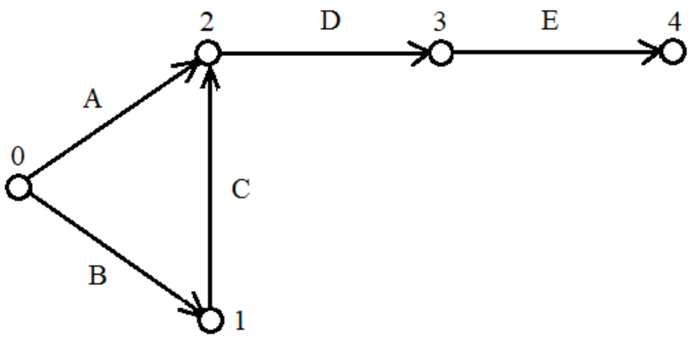

Figure 4: Schematic depiction for the activity information

4.2 Analytical Results

Table 2 enumerates the results obtained by conventional CPM and PERT methods and probability approaches through confident limited (as reported by Lu et al. [5]), as well as those using imprecise probability structures including the one with no assumption on probability distribution. For probabilistic and Imprecise Probability approaches, the expected duration for each activity is considered to have an independent normal distribution with the mean set to the PERT time and a standard deviation is: (Max time-Min time)/6.

Table 2: Expected overall duration of the project using different methods

Analytical Method	Overall Expected Project Duration (Days)
CPM	42
PERT Activity Time=(Min Time+4*Expected Time+Max Time)/ 6	43.5
Probability Approach (The 95% confident limited)	[43.20, 44.10]
Imprecise Probability (Independent-30 point enclosure discretization)	[43.02, 44.21]
Imprecise Probability (Independent-100 point enclosure discretization)	[43.37, 43.83]
Imprecise Probability (With any Dependencies)	[41.57, 47.49]
Interval Analysis (No assumption on distribution of activity duration, "worst-case interval")	[37,62]

4.3 Observations

The results in Table 2 show that using an imprecise probability structure for event duration yields a series of more objective bounds for the overall duration mean than conventional CPM, PERT, and probabilistic approaches. Moreover, higher levels of discretization narrow the bounds more precisely. Furthermore, consideration of any dependencies yields wider results. Direct implementation of interval time durations with no assumption on the distribution yields the widest results compared to imprecise probability approach.

5 Summary and Conclusions

In this work, a new formulation for construction scheduling based on the concept of imprecise probability is introduced. This method can obtain the bounds for the mean of overall project duration with consideration of uncertainties in the values and distribution for duration of each activity. The results obtained from this method have a higher level of confidence and robustness due to objective evaluation of variations in the parameter distributions. This objectivity makes it attractive to introduce imprecise probability concepts in the field of construction scheduling and management.

References

[1] R. Williamson and T. Downs, Probabilistic arithmetic. i. numerical methods for calculating convolutions and dependency bounds, *International Journal of Approximate Reasoning* **4**, 89 (1990).

[2] S. Ferson, V. Kreinovich, L. Ginzburg, K. Sentz, and D. S. Myers, Constructing probability boxes and Dempster- Shafer structures (2003).

[3] J. J. Moder, C. R. Phillips, and E. W. Davis, *Project management with CPM, PERT and precedence diagramming* (Van Nostrand Reinhold, 1983).

[4] W. D. Cottrell, Simplified program evaluation and review technique (pert), *Journal of Construction Engineering and Management* **125**, 16–22 (1999).

[5] M. Lu and S. M. Abourizk, Simplified CPM/PERT simulation model, *Journal of Construction Engineering and Management* **126**, 219–226 (2000).

[6] S. M. Abourizk, D. W. Halpin, and J. R. Wilson, Visual interactive fitting of beta distributions, *Journal of Construction Engineering and Management* **117**, 589–605 (1991).

[7] S. M. Abourizk, D. W. Halpin, and J. R. Wilson, Fitting beta distributions based on sample data, *Journal of Construction Engineering and Management* **120**, 288–305 (1994).

[8] F. S. McLaughlin and R. C. Pickhardt, *Quantitative techniques for management decisions* (Houghton

Mifflin, 1979).
[9] S. Ferson and J. G. Hajagos, Arithmetic with uncertain numbers: rigorous and (often) best possible answers, *Reliability Engineering and System Safety* **85**, 135–152 (2004).
[10] H. N. Ahuja, S. P. Dozzi, and S. M. Abourizk, *Project management: techniques in planning and controlling construction projects* (John Wiley, 1994).