Project Management Development – Practice and Perspectives
8th International Scientific Conference on Project Management in the Baltic Countries
April 25-26, 2019, Riga, University of Latvia
ISSN 2256-0513, e-ISSN 2501-0263

PRACTICAL VERIFICATION OF CLT ASSUMPTION FOR PERT APPLICATION

Maksim Goman, Department of Business Informatics –
Information Engineering, Johannes Kepler University, Linz, Austria

Abstract

One of the assumptions in PERT is the possibility to apply central limit theorem (CLT) to approximate path duration times with standard normal distribution. However, CLT presumes certain conditions in order to be correctly applied. This issue is often not adressed in related literature. The aim of this paper is to examine the conditions under which PERT can accordingly be applied. We consider the conditions and aspects of their practical application in order to verify the admissibility of CLT for given activity time distributions. Lindeberg-Feller condition turned out to be the simplest technique to verify that results of a PERT analysis are free from problematic CLT application. We also summarize other issues with current probabilistic project evaluation and propose a chance constraint optimization model for probabilistic project analysis.

Key words: *PERT, central limit theorem, CLT, project management, path duration*.
JEL code: M15, O22

Introduction

Project Evaluation and Review Technique (PERT) is known by project managers and scholars. The principal idea is attractive: To produce a probabilistic analysis for the project completion time. However, there has been a lot of criticism of PERT mostly due to its model assumptions since the 1960s, e.g. (Charnes et al., 1964; Hartley and Wortham, 1966; MacCrimmon and Ryavec, 1964; Roy and Roy, 2013). Charnes et al. (Charnes et al., 1964) have already admitted erroneous usage of central limit theorem (CLT). Nevertheless, it is still taught at universities, explained in textbooks, implemented in software, and there is no other widely used and concise stochastic method today. Improvements that produce better bounds or approximate time distributions have either the same or similar assumptions as the original PERT or require large analytical or modelling efforts and therefore, are not easy to implement for approximate time estimation (Elmaghraby, 1989).

There have been a lot of extensions of PERT proposed over the years. Although, there was enough work advancing several concepts, there is no finished alternative available. We would stress two main research directions: One tries to reduce the uncertainty of resulting time estimation, the other one aims at speeding up calculations for large stochastic networks. Both do not address the basic assumptions. Every theory or algorithm has its model. If the model does not reflect the observations of reality during empirical validation or the validity of the model cannot be explained theoretically, then the results should be explained somehow. Otherwise, the outcome is problematic.

There is no obvious technique of uncertainty reduction given initial input estimations from expert estimation with potential bias, and some scholars assume mathematically nice but not empirically proven distribution types. As regards fast computations with long time horizons, computers have enough power today such that a user in practice would not probably notice the difference between tenth of a second or one second and maybe agree to wait for minutes in

return to a reliable estimation, given realistic network size assumptions. The main need is reliable reduction of uncertainty, and better and stable estimations. This bias grows with prediction time horizon and results thus become practically useless with long time horizon. This was shown already by Charnes et al. (Charnes et al., 1964). This is easy to see especially for distributions having infinite tails (e.g. exponential, normal) or large variance (e.g. uniform). Accuracy of the methods can only be verified at the end of each task and after the end of the project when time realizations will have become known. As we use mathematical models that reflect reality to certain extent, only empirical validation of the methods is possible and this is what scholars and practitioners usually do not do.
Problematic assumptions of the original PERT (Malcolm et al., 1959) include:

– Unimodal Beta distribution that allows for only three possible types and does not reflect diversity of decision maker's (DM) preferences. Building of the distribution is a mathematical abstraction (Roy and Roy, 2013). Its choice for PERT was not supported with any evidence (Roy and Roy, 2013).

– Use of expected time instead of probability distributions. After setup, PERT transforms the problem to a deterministic one. It was shown that the deterministic problem underestimates all values even in theory, e.g. with Jensen's inequality (Benati, 2006; Elmaghraby, 1989; MacCrimmon and Ryavec, 1964).

– Independent distribution functions of individual activities. This assumption simplifies the problem, but is not realistic: Tasks are at least dependent on their sequence. If earlier tasks finish late, subsequent tasks on critical path should be late as well and vice versa. There can be explicit dependencies between certain tasks. Once parallel threads collapse into one event node, the time distribution function (DF) of the next task becomes a conditional DF. In reality, the paths in stochastic networks (SN) are not independent because they share some activities. However, most research so far did not take into account the dependency between paths. Yao and Chu (Yao and Chu, 2007) showed that significant bias exists in the approximated DF of the project completion time when path dependency is ignored.

– Large number of tasks in a path in order to approximate the sum of their time distributions with normal distribution applying CLT. This assumption is also wrong for projects with parallel activities. Obviously, almost all projects have parallel tasks. Leemis et al. (Leemis et al., 2006) and Elmaghraby (Elmaghraby, 1989) argue that in case of parallel networks with independent and identically distributed (iid.) stochastic activity durations, the resulting time distribution is skewed so therefore CLT is inappropriate. Another issue is possible dependence of tasks. In both cases, the DF becomes conditional DF and simple CLT is not valid. There exists multidimensional CLT, but its application was only considered for normally distributed times, e.g. by (Monhor, 2011).

- CLT introduces ambiguity about initial time distributions of tasks (Roy and Roy, 2013), i.e. they had initially beta DF functions (DFs), but going back from the resulting normal distribution of the whole project, we can assume that they are marginally normal. In this way, there is little use of constructing initial beta distribution of tasks.

As a result, PERT considers only one critical path and obtains only one of many possible lower bounds of total project's time. There are many competing critical paths and possible lower bounds in stochastic project time network.

There have been many attempts to use normally distributed times in PERT analysis, a review can be found in (Udoumoh and Ebong, 2017) Choice of a normal distribution was due to simplified modelling and calculations, especially for multivariate case. Assumptions of

ISSN 2256-0513, e-ISSN 2501-0263

independence of times, sequences, and paths usually accompany normality of activity time distributions. Unfortunately, the biased assumptions and lack of empirical evidence prevent us from accepting these models in practice. Computation time for a project network is not an issue today for realistic sizes of PERT networks. DF, cumulative distribution function (CDF) and inverse CDF functions were implemented in statistical tools and programming libraries. Aggregation of DFs can be done in a fraction of second today provided that DFs are known. On the other hand, the choice of normal distribution needs more explanations of its negative range of values and infinite tails. According to (Udoumoh and Ebong, 2017), some scholars use truncated normal distribution which is a solution to the problems with infinite tails. However, the result is not a normal distribution, and lacks the desired property of the normal family and the choice thus does not offer any advantages over any other distribution. Use of normal distribution was also found in (Monhor, 2011) and (Prékopa et al., 2004). Based on convenient qualities of multivariate normal distribution with correlations, a new approach to probabilistic critical path was presented. Although the papers provide an important step forward in identification of probabilistic critical path, because of the assumption of convergence of activity times on a path to normal distribution and model oversimplification, these models still do not seem fully practical.

We consider that any method or algorithm should be applied only under the conditions that were defined for them. Without meeting these requirements, the results are problematic because the method was not designed for arbitrary conditions and can not be expected to return consistent and reasonable results. Therefore, how can we rely on a priori unproved result in management? One of the conditions for PERT is applicability of CLT for time distributions. We believe that in order to obtain consistent estimations we should at least determine that the underlying conditions are satisfied.

Having described related problems with the PERT method, we formulate the research question: Can it be assured that PERT returns mathematically correct results through the use of CLT?

We are going to consider ways of mathematical verification of initial random time estimations of tasks in order to determine applicability of CLT to the data and assure validity of PERT results. This analysis assumes continuous time DFs, but it generalizes to discrete case. This can at least remove the bias of inappropriate CLT application to the given data, and constitutes a contribution to current PERT analysis.

We will also define an improved stochastic model for probabilistic project analysis including time, cost, quality, resources and other constraints. This new *probabilistic* PERT model aims at relying on more realistic assumptions and is based on chance constraint model. We believe that the term *probabilistic PERT* suits the new model better and we refer the usual PERT as *original* PERT.

The paper contains the following parts. We will consider four main approaches for checking the applicability of CLT to given data. Next, examples for data verification for CLT applicability will be given. The new model for future PERT improvement is proposed after a short discussion of potential directions for methods used in project analysis. Summary and future work directions are presented in the concluding section.

Verification of conditions for CLT application

According to original PERT, CLT is applied to approximate path duration times with standard normal distribution, independently of the DFs of individual activity times. There is an informal rule of thumb that there should be at least 30 activities on a path for proper CLT application and (Ludwig et al., 2001) claim that 10 activities on a path is enough for a good approximation. However, CLT presumes certain conditions in order to be safely applied and there exist several alternative conditions in theory. Unless we know a priori that distributions of the considered times satisfy them, we need to check these conditions. Violating CLT conditions can invalidate the project time estimation.

Let time of k independent tasks be random variables, i.e. we have a random vector $\boldsymbol{X} = (\mathrm{X}_1, X_2, ..., X_k)^T$ that has the size k. According to the properties of independent random variables, we can obtain expected value and variance of aggregated time of tasks:

$$\mu = E(\boldsymbol{X}) = \sum_{i=1}^{k} E[X_i] \; ; \qquad \sigma^2 = \sigma^2(\boldsymbol{X}) = \sum_{i=1}^{k} Var[X_i].$$

The CLT theorem tells us that it should be applied only if summands meet certain conditions. Then, the sum of a large number of uncorrelated random variables can converge to approximately normal distribution and we can directly obtain any quantile of time distribution:

$$\sum_{i=1}^{k} X_i \sim N\left(\sum_{i=1}^{k} \mu_i, \sum_{i=1}^{k} \sigma_i^2\right).$$

There is no discussion in project management area on whether project time data a priori satisfy the requirements. If we do not have a-priori information, we need to check the conditions. We will show that CLT applicability check is not complex. However, violating CLT conditions means that convergence of the sum to normal distribution is not guaranteed and that thus PERT estimation could not be valid.

There are a number of alternative criteria for CLT applicability for a sum of k independent random variables with finite expected values and variances.

1. Lindeberg-Feller (L-F) condition (Spanos, 1999) checks that not a single variance is greatly larger than others. From Lindeberg's condition

$$\lim_{k \to \infty} \frac{1}{\sigma^2} \sum_{i=1}^{k} \mathbb{E}\left[(X_i - \mu_i)^2 \cdot \mathbf{1}_{\{|X_i - \mu_i| > \varepsilon\sigma\}}\right] = 0,$$

Where $\mathbf{1}$ is the indicator function, follows Feller's condition:

$$\lim_{k \to \infty} \left(\max_{1 \le i \le k} \left(\frac{\sigma_i^2}{\sigma^2} \right) \right) = 0 .$$

The meaning of the conditions is that no single random variable dominates others in variance. Therefore, we can approximate: $\forall i \in [1, k]$ and given constant $\delta > 0, \max_i \frac{\sigma_i^2}{\sigma^2} < \delta$. Parameter δ is assumed to be small.

Parameter δ regulates applicability of CLT regarding the quality of convergence. One may imagine it as a ratio of known values that is easy to maintain, e.g. $\delta \geq \frac{1}{k}$. Thus, the choice of δ or comparable parameter for CLT is a measure of applicability of the method. If Lindeberg-Feller condition does not hold, we can not use CLT.

2. Another possible solution for CLT applicability is Lyapunov's condition (Spanos, 1999):

$$\forall\delta > 0,\ \frac{1}{\sigma^{2+\delta}} \sum_{i=1}^{k} E\left[|X_i - \mu_i|^{2+\delta}\right] \xrightarrow[k\to\infty]{} 0.$$

Higher central moments or their absolute values are not given for the sum of random time variables (fourth order is an efficient substitution for the modulo), but we can derive the central moments from noncentral moments:

$$E[(X-\mu)^3] = E(X^3) + 2(E(X))^3 - 3E(X)E(X^2),$$
$$E[(X-\mu)^4] = E(X^4) - 4E(X^3)E(X) + 6E(X^2)(E(X))^2 - 3(E(X))^4.$$

Noncentral moments of the third or higher order can be computed as derivatives of the respective order of known characteristic function (CF) of each time distribution at point 0. And noncentral moments of the sum of time distributions are obtained as nonlinear combination of products of its marginal non-central moments, e.g.:

$$\mathbf{E}[(\mathbf{X_1}+\mathbf{X_2})^\mathbf{3}] = E[X_1^3] + 3\left(E[X_1^2]E[X_2] + E[X_1]E[X_2^2]\right) + E[X_2^3],$$
$$\mathbf{E}[(\mathbf{X_1}+\mathbf{X_2})^\mathbf{4}] = E[X_1^4] + 4E[X_1^3]E[X_2] + 6E[X_1^2]E[X_2^2] + 4E[X_1]E[X_2^3] + E[X_2^4],$$
$$\mathbf{E}[(\mathbf{X_1}+\mathbf{X_2}+\mathbf{X_3})^\mathbf{2}] = E[X_1^2] + E[X_2^2] + E[X_3^2] + 2\left(E[X_1^X{}_2 + X_1X_3^2 + X_2X_3]\right),$$
$$\mathbf{E}[(\mathbf{X_1}+\mathbf{X_2}+\mathbf{X_3})^\mathbf{3}] = E[X_1^3] + E[X_2^3] + E[X_3^3] + 3(E[X_1]E[X_2^2] + E[X_1^2]E[X_2] +$$
$$+E[X_1^2]E[X_3] + E[X_2^2]E[X_3] + E[X_1]E[X_3^2] + E[X_2]E[X_3^2]) + 6E[X_1]E[X_2]E[X_3],$$
$$\mathbf{E}[(\mathbf{X_1}+\mathbf{X_2}+\mathbf{X_3})^\mathbf{4}] = [X_1^4] + E[X_2^4] + E[X_3^4] + 4(E[X_1^3]E[X_2] + E[X_1^3]E[X_3] +$$
$$+E[X_2^3]E[X_3] + E[X_2^3]E[X_1] + E[X_3^3]E[X_1] + E[X_3^3]E[X_2]) +$$
$$+6\left(E[X_3^2]E[X_1^2] + E[X_1^2]E[X_2^2] + E[X_2^2]E[X_3^2]\right) +$$
$$+12\left(E[X_2^2]E[X_1]E[X_3] + E[X_3^2]E[X_1]E[X_2] + E[X_1^2]E[X_2]E[X_3]\right).$$

Powers of expected values of orders of random variables in the last formulae, i.e. X^2, X^3, X^4, etc., are easily calculated for independent random variables. Assuming X and Y are independent random variables with DF (Spanos, 1999):

$$f_{XY}(X,Y) = f_X(X)f_Y(Y)$$
$$E[X\cdot Y] = E[X]\cdot E[Y]$$
$$Var[X\cdot Y] = Var[X]\cdot Var[Y] + (E[Y])^2Var[X] + (E[X])^2Var[Y].$$

Unfortunately, the complexity for these operations is at least polynomial in *k*, i.e., the number of summands grows very fast. Nevertheless, this operation is conceptually feasible because of existing recurrent relations for the formulae. Nevertheless, an algorithmic implementation is required.

3. Berry-Esseen (B-E) theorem with constant 0.4097 ≤ C ≤ 0.7975 (Spanos, 1999) for independent random variables provides the rate of convergence of the sum to the normal distribution and the maximal error of approximation:

$$\max_{z\in\mathbb{R}}\left|Pr\left(\frac{1}{\sigma}\sum_{i=1}^{k}(X_i-\mu)\le z\right)-\Phi(z)\right|\le C\frac{\sum_{i=1}^{k}E(|X_i-\mu_i|^3)}{\sigma^3}.$$

Application of higher moments is possible. The relation gives a bound on the maximal error of approximation between the normal distribution and the normalized distribution of the sum of random variables (measured by the Kolmogorov–Smirnov distance). Use of the upper bound of parameter C requires thousands of iid. random variables for good convergence (Spanos, 1999) because convergence rate of the difference to zero is $n^{-1/2}$. Although it is not convenient, central moments for the sum of arbitrary distributions can be calculated in the same way as for Lyapunov's CLT condition above. The third central moment of modulo can be substituted with the fourth order higher central moment.

4. Finally, it is possible to check convergence of the sum of random time variables to the standard normal distribution using the following three metrics for arriving at parameters of the normal distribution because all the parameters are constant for any normal distribution:

– variance of the sum $\sum_{i=1}^{k}\sigma_i^2\to 1$,
– skewness of the sum $E[(X-\mu)^3]/\sigma^3\to 0$,
– kurtosis of the sum $E[(X-\mu)^4]/\sigma^4\to 3$.

For the sum of arbitrary distributions, these parameters should be examined in the same way as for Lyapunov's CLT condition above.

There is multivariate CLT for random vectors that can be applied to joint distributions. For a sequence of iid. random vectors $\boldsymbol{X}_i$ with $E(\boldsymbol{X}_i)=\boldsymbol{\mu}$ and $Cov(\boldsymbol{X}_i)=\boldsymbol{\Sigma}$ under the restriction that no random vector dominates, it converges in order α to the following multivariate normal distribution (Spanos, 1999):

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{X}_i-\boldsymbol{\mu}\right)\sim_\alpha N(\boldsymbol{0},\boldsymbol{\Sigma}).$$

However, we do not go beyond the requirements of the original PERT method. Multivariate (conditional) time random variables need a better model that includes assumption of task dependence.

Thus, applicability of CLT is easy to verify and in the following we will consider two examples. It is obvious that Lindeberg-Feller condition is easier to verify because it needs only a ratio of each variance to the sum of all variances and a threshold.

1. Examples of data verification for CLT applicability

We will consider two examples of sample problems where we check applicability of CLT. We assume different distributions (Beta, Normal, Triangular and Uniform) in the examples. We apply the simplest Lindeberg-Feller condition. A parameter $\delta = 0.05$ is used. This is very rough and conservative approximation proposal: It does not approach to zero well when applied with Feller condition. Nevertheless, even with such a favourable parameter, CLT conditions may not hold. Results of evaluation of CLT conditions for the single original PERT's critical path are given in the Tables 1 and 2.

Table 1

Critical path of 10 identical tasks, δ= 0.05

№	Distribu-tion	Opti-mistic	Most likely	Pessi-mistic	E(X)	VAR (X)	Sum VAR	L-F cond.	B-E cond.
Part 1: Ten original random variables									
1	Beta-10%	54	90	135	91.5	182.3	1822.5	0.1	0.152
2	Beta	60	100	150	101.7	225.0	2250	0.1	0.152
3	Beta+10%	66	110	165	111.8	272.3	2722.5	0.1	0.155
4	Triangular	60	100	150	103.3	338.9	3388.9	0.1	0.074
5	Uniform	60	100	150	105.0	675.0	6750	0.1	0.098
Part 2: Ten modified random variables (basis)									
1	Beta a	29.83	38.7	47.57	38.7	9.7	87.4	0.1	0.130
2	Beta	38	43	48	43	2.8	27.8	0.1	0.125
3	Beta b	37.57	47.3	57.03	49.3	10.5	105.2	0.1	0.129
4	Triangular	38	43	48	43	4.2	41.7	0.1	0.074
5	Uniform	38	43	48	43	8.3	83.3	0.1	0.098
Part 3: Single additional random variable q (smaller variance)									
1	Beta	41	43	45	43	0.4	87.9	0.0995	0.129
2	Beta	41	43	45	43	0.4	28.1	0.0984	0.121
3	Beta	41	43	45	43	0.4	105.6	0.0996	0.124
4	Triangular	41	43	45	43	0.7	42.3	0.0984	0.072
5	Uniform	41	43	45	43	1.3	84.7	0.0984	0.095
Part 4: Single additional random variable q (larger variance)									
1	Beta	33	43	53	43	11.1	98.5	0.1128	0.119
2	Beta	33	43	53	43	11.1	38.9	0.2857	0.170
3	Beta	33	43	53	43	11.1	116.3	0.0955	0.117
4	Triangular	33	43	53	43	16.7	58.3	0.2857	0.098
5	Uniform	33	43	53	43	33.3	116.7	0.2857	0.130

Source: Author's calculations based on Hajdu and Bokor (2014)

The first example in Table 1 was taken from sample 1 in (Hajdu and Bokor, 2014). Part 1 is simply a sequential set of 10 independent tasks with identical DF. Part 2 contains changed data: Derived from the most likely value (m.l.v.) of "Beta" were most likely values of "Beta a" m.l.v. ("Beta a")=0.9*m.l.v.("Beta") and m.l.v.("Beta b")=1.1*m.l.v.("Beta"). Optimistic (pessimistic) values of distributions "Beta a" are obtained as m.l.v.*0.9-5 (m.l.v.*1.1+5). Finally, CLT is not applicable according to Lindeberg-Feller condition. The third and forth examples add an eleventh random variable of the same distribution as other 10 variables, but with slightly smaller variance in example 3 and larger one in example 4. In examples 3 and 4 optimistic, most likely, and pessimistic estimations of the additional random variable with its expected value and variance are given in Table 1. The variable is added to the 10 basis variables in the respective lines from example 2 to obtain the sum of variances and Feller's condition for the sum.

It is easy to see from Table 1 that the original data from the paper (Hajdu and Bokor, 2014) does not satisfy CLT conditions. Parts 2-4 show that adding a single random value with the same mean value but smaller variance does not change a lot, while on the opposite, adding an extra variable with relatively larger variance deteriorates the condition dramatically. It is worth to note that "relatively large" starts only at 15% larger variance in the example 4 and the difference of the criteria are evident.

Berry-Esseen condition for all the cases was computed using *fourth central moments*. It shows the maximum difference of cumulative DF of the sum and the respective normal distribution cumulative DF. It is clear to see that 10 iid. distributions are not enough to achieve a sufficient similarity to normal distribution. The difference is over 10% most of the time. Although a DM should decide on whether this is acceptable, the popular values of statistical significance today are 5% or less. In Example 4, it is clear to see that adding only one random value with larger variance increases the difference in all cases. Both L-F and B-E conditions behave consistently.

Table 2

Critical path of 10 tasks, δ = 0.05

Parameter	Beta (Mean/Var)	Triangular (Mean/Var)	Uniform (Mean/Var)
Task 2	19/9	16.67/15.50	23/27
Task 5	2.33/0.44	2.67/0.72	3.00/1.33
Task 8	0.63/0.01	0.67/0.02	0.70/0.03
Task 15	3.17/0.25	3.33/0.39	3.50/0.75
Task 16	4.00/0.11	4.00/0.17	4.00/0.33
Task 17	10.00/0.44	10.00/0.67	10.00/1.33
Task 18	2.00/0.11	2.00/0.17	2.00/0.33
Task 27	0.52/0.025	0.53/0.03	0.55/0.07
Task 28	3.83/0.257	3.67/0.39	3.50/0.75
Task 29	2.17/0.25	2.33/0.39	2.50/0.75
Sum VAR	10.89	18.44	32.68
Max L-F cond.	0.83	0.84	0.83

Source: Author's calculations based on Birge and Maddox (1995)

The second example was taken from (Birge and Maddox, 1995). Assuming minimum duration as pessimistic, maximum as optimistic, mean as most likely from table 8 in (Birge and Maddox, 1995), we produced beta distribution parameters of task duration according to the original PERT method. Critical path was also determined according to the original PERT approach: 2-5-8-15-16-17-18-27-28-29. We verified L-F CLT condition for the critical path and summarized in Table 2. We then experimented with triangular and uniform distributions based on the same initial three point estimates. Although the variance of these distributions is larger, the effect on CLT condition is limited.

We can see from Table 2 that in this case CLT is not applicable for 10 tasks in the path because variance of task 2 is dominating.

To conclude, reliance on CLT for good estimations is very poor. As it follows from theory, CLT is a tool for *large* number of summands of relatively close (preferred small) variance. Outliers in sense of variance will be dominating when sample size is close to infinity. Small samples may not assure sustainable and reliable approximation. The rule of thumb of 30 samples for CLT application does not hold either. From our further experiments with example 1 in Table 1, sample sizes should be about one hundred to get the Kholmogorov-Smirnov difference less than 1 % with conservative conditions and data. This conforms to the theory of CLT. We considered very simple case of iid. variables with only one outlier and it was hard to obtain close approximation of normal distribution. Results will clearly depend on data. Therefore, for project estimation, we suggest verification of the conditions, not relying on chance.

Discussion

We found that the research question can be answered affirmatively: By verifying CLT conditions we can assure that original PERT gives appropriate answer for the given critical path. Unfortunately, this is a fixture for only one PERT assumption. Moreover, we should consider next, what to do in case when CLT is not applicable.

We believe that further research activities should be applied to definition of multiple critical paths and usage of bounding methods to derive time bounds for each task. Search for uncertain duration presumes aggregation of time DFs along paths in the project stochastic network. A universal approach should presume arbitrary distributions and dependencies of tasks. We consider that the uncertain nature of the problem imposes little credit to very precise time estimations. Therefore, methods that can compute reasonable bounds of time distributions of task durations promise more practical benefits to DM. These reasonable methods should include DM's risk perception in some way.

Based on the current study, we propose a six-value-set that may describe every time distribution. These are expected value, variance, upper and lower bounds of DFs and corresponding probabilities for the upper and the lower bounds. In case of a lopsided DF, only information about variance can not reveal how much the distribution is skewed, but the bounds will always show this information. This gives efficient information about time DFs to judge about their criticality for DM. Next, attention to problems of conditional uncertainty and cases of complex multivariate distributions is also needed. Finally, project management is a continuous activity. After the project start we observe realizations of time of tasks in the network. This is vital information about real performance and hence, it should be used to update time estimations of remaining tasks and even for corrections of the model. This update is needed according to Deming Cycle and a possible solution for that is Bayesian approach.

Many researchers employ DFs that have nice mathematical properties in order to simplify formulas and computations. Ability to easily derive the resulting distribution of the convolution is vital for mathematical manipulations. But there are issues that have not been fully understood. For instance, if one applies normally distributed times, to what extent an infinite tail is realistic or what is the meaning of negative time in the left tail? It is not feasible to consider an infinite

ISSN 2256-0513, e-ISSN 2501-0263

delay of tasks: The tasks or the whole project will be terminated relatively shortly due to resource constraints and management intervention. Truncated random variables lose some of their nice properties and their addition is as hard as for uniformly distributed variables. Finally, how to build and verify normally distributed times based on the current state of knowledge and uncertainty about a particular project?

We believe that realistic DFs are important. Consideration of uniform or triangular distributions instead of beta distribution is new in this context since they were mostly not considered in the last 30 years. Nevertheless, these simple distributions are easier to build with an expert and understand their meaning. Johnson (Johnson, 1997) has shown that more simple intuitively obvious triangular distribution can be very close to beta DF and proposed a procedure for estimating the necessary parameters.

Chance constraint optimization model

We attempt to provide a model for prospective improvement of original PERT. An improved mathematical model should have realistic assumptions about distributions used and the notion of stochastic critical path. We aim at creation of *probabilistic* PERT with outcomes similar to CPM method, i.e. estimation of earliest and latest start and end times for every task. It should be not significantly more complex in use for practitioners than current deterministic methods, yet be based on well developed mathematical theory. Most often, the main goal is estimation of the maximum time of a project. However, resource and cash flow uncertainties are also important and should be represented in the model. There are models that consider costs or cash flows along with time, e.g. (Benati, 2006). However, to the best of our knowledge there are no fully stochastic models that combine all aspects of time, cost and resource availability, i.e. current solutions are stochastic only in one of the two aspects: time or cash flow.

One of the first chance constraint programming models for SN was proposed by (Charnes et al., 1964; Bruni et al., 2009). In their stochastic optimization problem activity times had exponential and discrete DFs respectively. The former model was shown to have errors (Bruni et al., 2009) and used a chance constraint programming method that turned out to be problematic (Elmaghraby et al., 2001).

A universal management approach should account for time and resources. The only assumption is that DFs of the activities' durations are known. We do not assume any specific DF types or independence. DM's risk aversion can be expressed with reliability thresholds α_i for different constraints. The model is as follows:

$$
\begin{aligned}
\min\ & Ub(\pi_q) && //\text{upper time bound of critical path}\\
& (-Lb(\pi_q)) && //\text{lower time bound of critical path}\\
& C_{\mathbf{r}}(\pi_q) && //\text{cost of resources}\\
\text{s.t.}\ & Pr\{\pi_q \geq \pi_p\} \geq \alpha,\ \forall q,p \in \Pi,\ q <> p,\ q = 1,\ldots,Q && //\text{a) set of Q critical paths}\\
& Pr\{X_j^s \leq t_j^s\} \geq \beta_j^s && //\text{b) specific start time constraints}\\
& Pr\{X_j^f \leq t_j^f\} \geq \beta_j^f && //\text{c) specific finish time constraints}\\
& Pr\{f_{\mathbf{r}}^t(X_j) \leq \mathbf{r}_j^t\} \geq \rho && //\text{d) resource constraints at certain times}\\
& X_j^s \geq \max_{i \in Prec.} X_i^f && //\text{e) start time of the current task}\\
& X_j^f \leq \min_{i \in Succ.} X_i^s && //\text{f) finish time of the current task}\\
& \pi_p = \Sigma X_j | j \in \text{path p} && //\text{g) time of a path p}
\end{aligned}
$$

where

$X_j \in \mathbb{R}$ random duration of task *j*;

X_j^s, X_j^f start and finish time of task *j*;

$\pi \in \mathbb{R}$ duration of a path;

t_j^s, t_j^f predefined time constants;

$C_{\mathbf{r}}$ cost function of resource vector ***r***;

$f_{\mathbf{r}}^t(X_j)$ function of resource requirements for task *j* at time *t* depending on task duration;

ρ reliability of meeting resource chance constraints for task *j* at time *t*;

Prec. is a set of preceding tasks to the current task;

Succ. is a set of following tasks after the current task;

Π set of all paths;

$\alpha \in [0,1]$ is a reliability threshold for probabilistic critical path selection;

$\beta \in [0,1]$ is a reliability threshold for task time constraints.

In general, this is a multicriteria stochastic optimization problem. Most potential deterministic and stochastic issues for a project can be represented with the model, therefore all constraints are optional. Maximum time bounds of the project and cost of used resources are minimized. The bounds depend on *K* longest probabilistic paths. They are determined by chance constraint a). It is possible to have deterministic time thresholds t_j^s, t_j^f in chance constraints b) and c) that mean desired time of begin or end of certain task, e.g. a milestones or explicit dependency of project operations. Constraint vector d) defines probabilistic compliance with task resource requirements during task execution time. Constraints e) and f) are deterministic, but their variables are random time variables. Together, they represent the sequential task execution strategy in PERT networks. One of our goals should be the determination of time bounds of start and finish times of each task like in deterministic CPM method. The meaning of e) is the starting time of the current task is the longest finish time of all preceding tasks. The same is valid for f), the finish time of the current task is the earliest starting time of all following tasks. Finally, constraint g) accounts for stochastic time of a path.

We do not suggest any specific method to solve the model in the current paper. The model provides a uniform representation of uncertainty of project management problems. Cost functions of resources are reserved for future research possibilities. Stochastic task scheduling can be a natural extension of this model.

Conclusions

According to original PERT, CLT theorem approximates path duration times, regardless of DFs of individual activity times. However, the theory presumes that CLT is to be applied under predefined conditions. We considered verification of the conditions. While scholars usually omit CLT conditions or artificial data in experiments seem to converge with CLT to normal distribution, e.g. (Ludwig et al., 2001), this is an unproven heuristic. On two typical examples from literature, we showed that it may not be always true and data can contradict the theory of CLT. Regrettably, we did not find empirical data from a real project in literature for testing.

We considered the conditions of CLT and aspects of their practical application in order to verify the admissibility of CLT for given activity time distributions. These simple techniques allow verification of validity of classical PERT for specified activity time estimations. Every practitioner can use the verification to assure validity of PERT assumptions and feasibility of the results of PERT analysis, even using spreadsheet software. Illustrative examples were provided. We believe that this check should improve the degree of trust in PERT results.

While criteria for CLT applicability are well known in probability theory, our conclusion is that they may have restricted application. The constraints of CLT applicability are usually omitted in papers. We do not know whether CLT conditions are assumed in existing software. We consider that CLT can be applied to sequential critical path in classical PERT, but only after verification of CLT applicability. Aggregation of times of several parallel activities is to be done with other methods.

We can conclude that in order to improve original PERT, we should develop a better model with realistic assumptions and loopback control after the beginning of the project. We provide such a model in the paper. Our following research will be directed at obtaining time bounds for project tasks with the aforementioned model.

Finally, a project is not a static thing. After the start, time of finished activities become realizations of random variables. This is valuable information related to the current project with unique ambiance. In conformance with Deming principle, we should update our time estimations based on this recent information. Bayesian method is a natural tool for that. There are already attempts at this, e.g. (Cho, 2009; Gardoni et al., 2007; Kim and Reinschmidt, 2009). Thus, model assumptions can be amended and estimates of the following activities improved based on revealed times of previous ones. Apparently, the model should be updated every time new information is received using Bayes approach.

References

Benati, S., 2006. An Optimization Model for Stochastic Project Networks with Cash Flows. *Computational Management Science*, 3(4), pp.271–284.

Birge, J.R. and Maddox, M.J., 1995. Bounds on Expected Project Tardiness. *Operations Research*, 43(5), pp.838–850.

Bruni, M.E., Guerriero, F. and Pinto, E., 2009. Evaluating Project Completion Time in Project Networks with Discrete Random Activity Durations. *Computers & Operations Research*, 36(10), pp.2716–2722.

Charnes, A., Cooper, W.W. and Thompson, G.L., 1964. Critical Path Analyses via Chance Constrained and Stochastic Programming. *Operations Research*, 12(3), pp.460–470.

Cho, S., 2009. A linear Bayesian stochastic approximation to update project duration estimates. *European Journal of Operational Research*, 196(2), pp.585–593.

Elmaghraby, S.E., 1989. The Estimation of Some Network Parameters in the PERT Model of Activity Networks: Review and Critique. *Advances in Project Scheduling*. Elsevier, pp.371–432.

Elmaghraby, S.E., Soewandi, H. and Yao, M.-J., 2001. Chance-constrained Programming in Activity Networks: A Critical Evaluation. In: *European Journal of Operational Research*, 131(2), pp.440–458.

Gardoni, P., Reinschmidt, K.F. and Kumar, R., 2007. A Probabilistic Framework for Bayesian Adaptive Forecasting of Project Progress. *Computer-Aided Civil and Infrastructure Engineering*, 22(3), pp.182–196.

Hajdu, M. and Bokor, O., 2014. The Effects of Different Activity Distributions on Project Duration in PERT Networks. *Procedia - Social and Behavioral Sciences*, 119, pp.766–775.

Hartley, H.O. and Wortham, A.W., 1966. A Statistical Theory for Pert Critical Path Analysis. *Management Science*, 12(10,), pp.B469–B481.

Johnson, D., 1997. The Triangular Distribution as a Proxy for The Beta Distribution in Risk Analysis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 46(3), pp.387–398.

Kim, B. and Reinschmidt, K.F., 2009. Probabilistic Forecasting of Project Duration Using Bayesian Inference and the Beta Distribution. *Journal of Construction Engineering and Management*, 135(3), pp.178–186.

Leemis, L.M., Duggan, M.J., Drew, J.H., Mallozzi, J.A. and Connell, K.W., 2006. Algorithms To Calculate The Distribution Of The Longest Path Length Of A Stochastic Activity Network With Continuous Activity Durations. *Networks*, 48(3), pp.143–165.

Ludwig, A., Möhring, R.H. and Stork, F., 2001. A Computational Study on Bounding the Makespan Distribution in Stochastic Project Networks. *Annals of Operations Research* 102(1-4), pp.49–64.

MacCrimmon, K.R. and Ryavec, C.A., 1964. An Analytical Study of the PERT Assumptions. *Operations Research*, 12(1), pp.16–37.

Malcolm, D.G., Roseboom, J.H., Clark, C.E. and Fazar, W., 1959. Application of a Technique for Research and Development Program Evaluation. *Operations Research*, 7(5), pp.646–669.

Monhor, D., 2011. A New Probabilistic Approach to the Path Criticality in Stochastic PERT. *Central European Journal of Operations Research*, 19(4), pp.615–633.

Prékopa, A., Long, J. and Szántai, T., 2004. New Bounds and Approximations for the Probability Distribution of the Length of the Critical Path. In: K. Marti, Y. Ermoliev and G. Pflug, eds.,

ISSN 2256-0513, e-ISSN 2501-0263

Dynamic Stochastic Optimization, Lecture Notes in Economics and Mathematical Systems. Springer Berlin Heidelberg, pp.293–320.

Roy, D. and Roy, R., 2013. Distribution of the Activity Time in Network Analysis: A Critical Revisit with a Gamma Alternative. *Communications in Statistics - Simulation and Computation*, 42(6), pp.1288–1297.

Spanos, A., 1999. *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge University Press.

Udoumoh, E.F. and Ebong, D.W., 2017. A Review of Activity Time Distributions in Risk Analysis. *American Journal of Operations Research*, 07(06), pp.356–371.

Yao, M.-J. and Chu, W.-M., 2007. A New Approximation Algorithm for Obtaining the Probability Distribution Function for Project Completion Time. *Computers & Mathematics with Applications*, 54(2), pp.282–295.