A NOTE ABOUT THE PERT CONSTANT VARIANCE ASSUMPTION

José García Pérez – Catalina García García – María del Mar López Martín

Abstract

The main goal of Project Evaluation Review Technique is to estimate the expected value and variance for the `total time to finish a task', from the classical values provided by an expert in relation to the optimistic or least possible time (a), more likely or modal time (m), and pessimistic or greater possible time (b) values. The beta distribution has been traditionally proposed as the underlying distribution of this methodology. However, the original expressions proposed by the creators of PERT are only obtained from the beta distribution if a constant variance is considered. This paper analyses, from a critical point of view, the hypothesis of constant variance in PERT methodology. The main conclusion is that if, as it is usual, the beta is the underlying distribution, the hypothesis of constant variance is more reasonable than the variance depending on the most likely value, *m*. Alternatively, other distributions are presented where the hypothesis of constant variance will not seem adequate. This fact suggests that despite of its historical application, the beta distribution should not be an appropriate distribution to be applied in PERT.

Key words: Robust project management, Activity times, Beta distribution, Expert Judgment, uncertainty.

JEL Code: C44, C46, D81.

Introduction

The main goal of Project Evaluation Review Technique (PERT), proposed by Malcolm et al (1959), is to estimate the expected value and variance for the `total time to finish a task', which is a continuous random variable elicited from the classical values provided by an expert in relation to the optimistic or least possible time (a), more likely or modal time (m), and pessimistic or greater possible time (b) values. The beta distribution has been traditionally proposed as the underlying distribution of this methodology.

$$E(t) = \frac{1+4M}{6} \quad (1)$$

$$var(t) = \frac{1}{36} \quad (2)$$

where $t = \frac{x-a}{b-a}$ and $M = \frac{m-a}{b-a}$. The hypothesis of constant variance is clearly appreciated in expression (2). Grubbs (1962) pointed out ``the lack of theoretical justification and the unavoidable defects of the PERT statements, since estimates (1) and (2) are, indeed, 'rough' and cannot be obtained from the beta distribution on the basis of values a, m and b determined by the analyst". Sasieni (1986) proposed the following alternative expression for the expected value:

$$E(t) = \frac{1+kM}{k+2}, \quad (3)$$

Supposing the beta as underlying distribution, Golenko-Ginzburg (1988) presented the following expression for the variance:

$$var(t) = \frac{k^2M(1-M)+(k+1)}{(k+3)(k+2)^2} \quad (4)$$

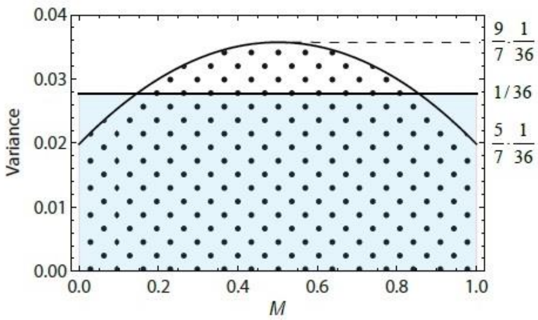
Note that (3) and (4) are identical to (1) and (2) when $k = 4$. This fact led Sasieni to wonder why the creator of PERT selected the value of k equal to 4. Littlefield and Randolph (1987) provided some of the first answers to these questions, thus obtaining the classical PERT expressions, (1) and (2) from the following assumptions: (i) the activity duration is beta distributed; (ii) the three point estimates a, m and b are good; (iii) the variance is expressed as (2) and does not depend on m; and (iv) a linear approximation is possible between M and k to solve the cubic equations obtained from the previous assumptions. Littlefield and Randolph's answer (1987) focused on the constant variance assumption but they did not justify the reason. In this paper, we present different reasons to justify the assumption of the hypothesis of constant variance in PERT. Firstly, Section 1 reviews the recent literature about the justification and criticism of the constant variance assumption, Section 2 provides a possible justification for the constant variance assumption and finally main conclusions are summarized in section 3.

1. Literature review

Chuen-Tao (1974) suggested that the creators of PERT selected the beta distribution for its capacity to be asymmetric but, at the same time, they pretended that the selected beta distribution were similar to the normal distribution. For this assumption, since 99.9% of the range of a normal variable is between 3σ and -3σ it is possible to obtain the expression (2) by assuming that the total range is $(b-a)=6\sigma$ and solving for σ. Following the hypothesis of a beta underlying distribution, Chae (1990) obtained from expression (1) the following expression for the variance:

$$\sigma^2(M)=\left(\frac{5}{7}+\frac{16}{7}M(1-M)\right)\frac{1}{36}. \tag{5}$$

Fig. 1: Representation of variance expression by Chae (1990).



Source: own computation.

Note that this expression (represented in Figure 1) does not lead to the constant variance assumption. That is, following Chae (1990) and considering the beta distribution, it is possible to obtain expression (1) from the assumption required in expression (2), but from the assumption required in expression (1) it is not possible to obtain expression (2). The interpretation of this condition is shown in Figure 1. Golenko and Ginburzg (1988) obtained these different expressions of PERT by imposing the condition $\int_0^1 \sigma^2 dM=\frac{1}{36}$:

$$E(t)=\frac{9M+2}{13} \tag{6}$$

$$\sigma^2(t)=\frac{1}{1268}(22+81M-81M^2) \tag{7}$$

On the other hand, the PERT expressions have been defended by authors such as Kamburowski (1997) and Herrerías et al. (2003) who focused on the similarity between the beta and the Gaussian distributions, not only in relation to the variance as proposed by Chuen-Tao (1974), but also in relation to kurtosis. That is, it is possible to obtain the exact classical

PERT expression by requiring the beta to have the constant variance $\frac{1}{36}$ and a kurtosis equal to 3. In spite of these justifications, the constant variance assumption has been widely criticized due to the fact that it ignores the most likely value which could be understood as the most committed value provided by the expert. Hahn (2008) even stated that the constant variance assumption "may be in conflict with reality" and proposed an alternative distribution called the beta rectangular.

In a recent work, Herrerías-Velasco et al. (2011) criticized the constant variance assumption and considered the following assumptions: (i) the activity duration is beta distributed and (ii) the expected value expression is (1). From these assumptions, they proposed an alternative beta distribution with a `PERT variance adjustment factor' that allows the mode M to be present in the calculation of the variance. They also defended the notion that the expert provides less information when placing the mode in the central point of the interval (a, b). Summing up, Herrerías-Velasco et al. (2011) provided an answer contrary to the one given by Littlefield and Randolph (1987) and obtained a beta distribution with a variance that depends on M from expression (1). This beta distribution had already been proposed by Chae (1990), who presented a graph similar to the that of Herrerías-Velasco et al. (2011) (see Figure 1).

2.- Some justifications for the hypothesis of constant variance

Based on the results of Chae (1990) and Herrerías-Velasco et al. (2011), we obtain a possible justification for the hypothesis of constant variance considering the following assumptions: i) the activity duration is beta distributed and ii) the expected value expression is given by expression (3). Firstly, from expression (4) but taking $k = 4$, it is possible to obtain an expression of the variance depending on M similar to the one obtained by Chae (1990).

$$\sigma^2(M) = \left(\frac{5}{7} + \frac{16}{7} M(1-M)\right)\frac{1}{36} = C(M)\frac{1}{36} \tag{8}$$

Note that Herrerías –Velasco (2011) obtained this same expression called C(M) the `PERT variance adjustment factor'. Expression (8) is represented as a truncated parabola with a maximum value of $\frac{9}{7}\frac{1}{36}$ at point $M = \frac{1}{2}$ and a minimum value of $\frac{5}{7}\frac{1}{36}$ at the extreme values 0 and 1. That is, the variance is maximum when the expert provides the most likely value in the center of the interval (a,b) and minimum when the expert places the most likely value at any extreme. Although this situation is consistent with the proposal of Herrerías –

Velasco (2011), we believe that it *may be in conflict with reality*, as stated by Hahn (2008) with regard to the constant variance assumption. We will defend later that from a PERT perspective there is no justification for giving maximum variance when $m = \frac{a+b}{2}$. This may explain why the creators of PERT selected a constant variance model that can be obtained simply with the average of the minimum and maximum variance values as already noted by Chae (1990) and shown in Figure 1.

$$\sigma^2 = \left(\frac{5}{7} + \frac{9}{7}\right)\frac{1}{2}\frac{1}{36} = \frac{1}{36} \tag{9}$$

Chae (1990) obtained the expression (2) from (1) by assuming a beta distribution but without the need for any similitude with the Gaussian distribution. However, he did not provide any justification for the constant variance assumption. Indeed, Chae's proposal (1990) in relation to the constant variance is, in our opinion, a timid proposal that must be read between the lines and does not even appear in the conclusions of the article.

We now provide arguments to defend that, if the underlying distribution of the total activity time is a beta and accepting that the expected value is given by (3), the most reasonable will be that the variance should be estimated by following the hypothesis of constant variance by using expression (2). In general, if the expected value is given by (3), the variance should be estimated from the following expression:

$$\sigma^2(t) = \frac{k^2 + 8k + 8}{8(k+3)(k+2)^2} \tag{10}$$

Note that this expression does not depend on M and becomes (2) when $k = 4$.

2.1. Reliability of expert

By following Hammond and Bickel (2013), the PERT question can be reduced to approximate the probability density function (PDF) or the cumulative distribution function (CDF) of a continuous distribution from a set of values x_i with probabilities p_i where $i = 1, 2, ..., n$. Returning to the discussion on whether it is coherent that the variance of the underlying distribution of PERT reaches a maximum when the expert provides the medium value of the interval (a, b) for the modal value (m), Herrerías-Velasco et al. (2011) stated that "maximum variance is obtained when the most likely value is centered at $\frac{a+b}{2}$. This is consistent with the mode location being least informative in that case". In contrast, we will defend a different line of thinking. From the assumptions that the activity duration is beta distributed and the expected value expression is given by expression (3) the following results are obtained:

Proposition 1: When the expert is questioned about the values of a, m and b, the answer provides a discrete distribution. Thus, by parting from expression (3), the discrete distribution is given in Table 1. Then, it is possible to analyse the discrete distribution provided by the expert and not only the values. Remember that in PERT the expert is asked about a, b and m values, or by standardizing $(0, M, 1)$. In any case, the expert provides three values of a discrete distribution and the expected value can be obtained from expression (3). Table 1 shows the different probabilities assigned to every value of the discrete distribution provided by the expert.

Tab. 1: Discrete distribution in PERT by following Sasieni (1986).

x_i	0	M	1
p_i	$\frac{1}{k+2}$	$\frac{k}{k+2}$	$\frac{1}{k+2}$

Source: own elaboration.

Proposition 2: The variance of the discrete standardized distribution provided by the expert depends on M and k and for any value of k, it is a truncated parabola with a minimum when $M = \frac{1}{2}$. The variance is constant for $k = 0$ and decreasing in k for any value of M.

Thus, from Table 1 it is possible to obtain the expected value of the discrete distribution and the corresponding variance:

$$var(t) = \frac{1 + k(1 - 2(1 - M)M)}{(k+2)^2} \tag{11}$$

Table 2 shows some values for the variance for different values of k. Figure 2 supports Proposition 2, displaying the representation of expression (11) according to the different values of M and k.

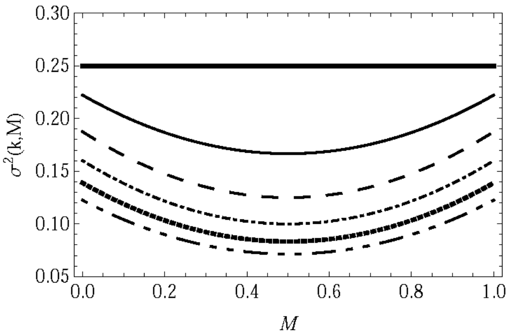
Tab. 2: Variance expressions for $k = 0$, $k = 1$, $k = 2$, $k = 3$, $k = 4$, $k = 5$ together the values of variance for $M = 0$, $M = 1$ and $M = 1/2$.

k	Variance Expression	Value of variance for $M = 0$ and $M = 1$	Value of variance for $M = \frac{1}{2}$
0	¼	¼	¼

1	$\frac{2}{9}(1+(M-1)M)$	2/9	0.1667
2	$\frac{1}{16}(3-4(1-M)M)$	3/16	0.125
3	$\frac{2}{25}(2-3(1-M)M)$	4/25	0.1
4	$\frac{1}{36}(5-8(1-M)M)$	5/36	0.0833
5	$\frac{2}{49}(3-5(1-M)M)$	6/49	0.0714

Source: own elaboration.

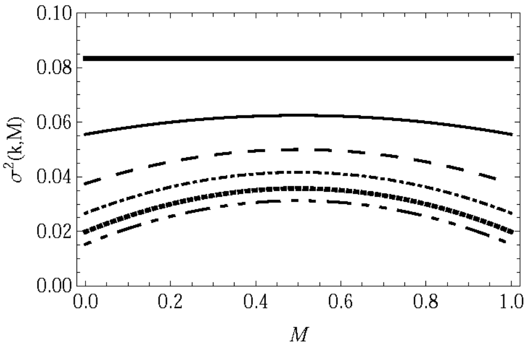
Fig. 2. Representation of variance expression (11) for k =0 (solid thick line); k =1 (solid thin line); k =2 (dashed line); k =3 (dotdashed line); k =4 (dotted line); k =5 (dot-dot-dashed line).



Source: own computation.

Focusing on the case k=4, note that when the standardized mode M is closer to the center point $M=\frac{1}{2}$, the variance is smaller and the reliability of the expert will therefore be higher. We consider that when the dispersion of the discrete distribution is high, the expert will provide less information. This is consistent with the fact that when the dispersion is maximum ($k=0$), the elicitation leads directly to the uniform distribution, which is the continuous distribution with the highest variance. However, when $a=m=b$ the variance is zero. Figure 3 displays the variance of the beta distribution applied in PERT (expression (4)) for different values of k and $M\in[0,1]$. In all cases, except for $k=0$, the representation is an inverted and truncated parabola. Contrary to what happens for the discrete distribution, all of them have a maximum for $M=\frac{1}{2}$. For all $k>0$, the variances increase as M is closer to $\frac{1}{2}$. Thus, if we use a tetraparametric beta distribution, the variance of the discrete distribution is minimum when the variance of the corresponding continuous distribution is maximum. It does not seem logical to estimate a higher variance when the expert gives more reliability to the provided values.

Fig. 3. Representation of variance expression (4) for k =0 (solid thick line); k =1 (solid thin line); k =2 (dashed line); k =3 (dotdashed line); k =4 (dotted line); k =5 (dot-dot-dashed line).



Source: own computation.

By considering that the expert's reliability is inversely related to the variance of the discrete distribution that is being provided, then the expert is acting more reliably when the variance of the discrete distribution is minimum, that is, when M is centered. On the other hand, the greater the reliability of the expert, the lower the variance of the underlying continuous distribution and this distribution must also have a minimum variance in $M=\frac{1}{2}$. As shown, this property is not verified by the beta distribution. In the opinion of Miller and Rice (1983), *few people would accept an approximation that did not have roughly the same mean, variance and skew as the original distribution*. This problem is only mitigated assuming the constant variance (averaging its minimum and maximum values) or changing the underlying distribution.

2.2. Other underlying distributions with variance depending on M

Hahn (2008) discussed the mean criticism of the constant variance, expression (2), and presented the mixture between the uniform and the beta distribution known as the beta rectangular distribution. The variance of the beta rectangular distribution depends on the mixture parameter θ and M. For this reason, we consider that it improves the estimations of the traditional PERT by providing a non-constant variance expression but with a minimum when $M = \frac{1}{2}$, contrary to the beta distribution proposed by Herrerías et al. (2011). The rectangular beta distribution is not the only distribution that presents a minimum when $M = \frac{1}{2}$. This property is also found in other simpler distributions, such as the two-sided power distribution, (van Dorp and Kotz, 2003) or the biparabolic distribution (García et al, 2010), among others. These distributions may be more appropriate as an underlying distribution in PERT than other distributions without this property.

3. - Conclusion

The variance of the beta distribution takes a maximum when $M = \frac{1}{2}$ and is not coherent with the general principles on which the PERT method is based since the expert's reliability is maximum at that point and the dispersion of the elicited distribution should therefore be minimum. This may explain why the creators of PERT selected a constant variance whose expression can be obtained simply by averaging the maximum and minimumvalues of the variance. Thus, the main conclusion is that if the beta distribution is the selected underlying distribution, then the more reasonable option is to consider the constant variance assumption since the variance will otherwise be maximum when the expert is given more information which is not consistent with the PERT principles.

References

Chae, K. C. (1990). A geometric interpretation of the PERT assumptions on the activity time. *International Journal of Mathematical Education in Science and Technology*, *21*(2), 283-288.

Chuen-Tao, Y., & Chuen-Tao, L. L. Y. (1974). *Aplicaciones prácticas del PERT y CPM: nuevas métodos de dirección para planificación, programación y control de proyectos*. Ediciones Deusto.

García, C. B. G., Pérez, J. G., & Rambaud, S. C. (2010). Proposal of a new distribution in PERT methodology. *Annals of Operations Research*, *181*(1), 515-538.

Golenko-Ginzburg, D. (1988). On the distribution of activity time in PERT. *Journal of the Operational Research Society*, *39*(8), 767-771.

Grubbs, F. E. (1962). Letter to the Editor—Attempts to Validate Certain PERT Statistics or "Picking on PERT". *Operations Research*, *10*(6), 912-915.

Hahn, E. D. (2008). Mixture densities for project management activity times: A robust approach to PERT. *European Journal of Operational Research*, *188*(2), 450-459.

Hammond, R. K., & Bickel, J. E. (2013). Reexamining discrete approximations to continuous distributions. *Decision Analysis*, *10*(1), 6-25.

Herrerías, R., García-Pérez, J., & Cruz Rambaud, S. (2003). A note on the reasonableness of PERT hypotheses. *Operations Research Letters*, *31*(1), 60-62.

Herrerías-Velasco, J. M., Herrerıás, R., & Van Dorp, J. R. (2011). Revisiting the PERT mean and variance. *European Journal of Operational Research*, *210*(2), 448-451.

Kamburowski, J. (1997). New validations of PERT times. *Omega*, *25*(3), 323-328.

Littlefield Jr, T. K., & Randolph, P. H. (1987). Reply—An Answer to Sasieni's Question on PERT Times. *Management Science*, *33*(10), 1357-1359.

Malcolm, D. G., Roseboom, J. H., Clark, C. E., & Fazar, W. (1959). Application of a technique for research and development program evaluation. *Operations research*, *7*(5), 646-669.

Miller III, A. C., & Rice, T. R. (1983). Discrete approximations of probability distributions. *Management science*, *29*(3), 352-362.

Sasieni, M. W. (1986). Note—A Note on Pert Times. *Management Science*, *32*(12), 1652-1653.

van Dorp, J., & Kotz, S. (2003). Generalizations of two-sided power distributions and their convolution. *Communications in Statistics-Theory and Methods*, *32*(9), 1703-1723.

Contact

José García Pérez

University of Almería

Ctra. Sacramento s/n, La Cañada de San Urbano, 04120 Almería

jgarcia@ual.es

Catalina García García

University of Granada

Campus Universitario de La Cartuja 18071 Granada

cbgarcia@ugr.es

María del Mar López Martín
University of Granada
Campus Universitario de La Cartuja 18071 Granada
mariadelmarlopez@ugr.es