

Understanding PERT

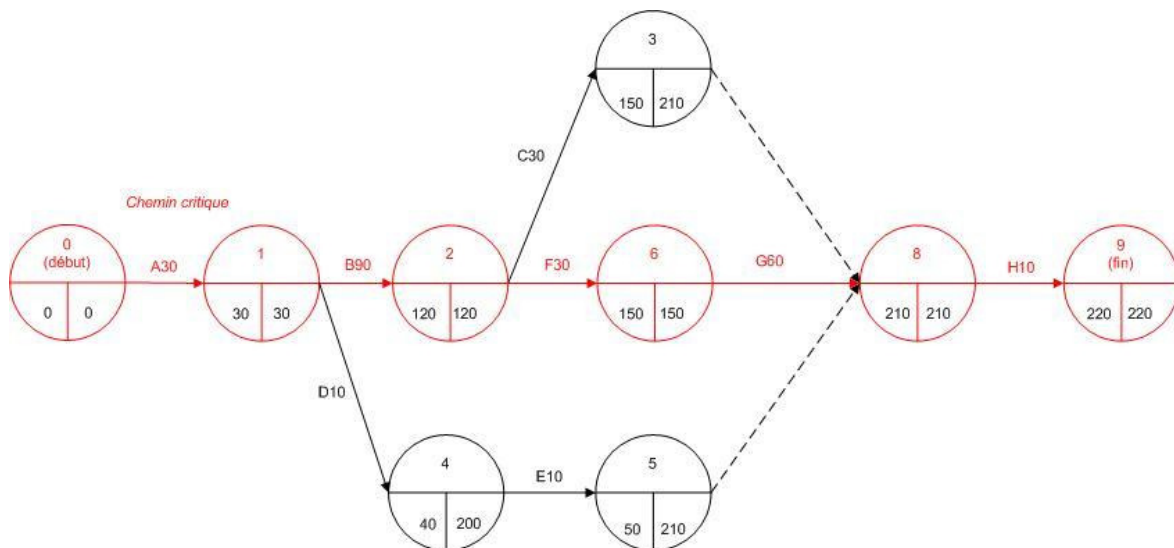
Programme Evaluation Review Technique

PERT is the oldest and arguable the least effective / least accurate way to model the uncertainty associated with every cost estimate or duration estimate used in a schedule. PERT can be described as an ‘event on node’ model; the US Navy wanted emphasis on knowing when activities would be complete (the event). The purpose of this White Paper is to define PERT and to clearly delineate what PERT is and what it is not.

Origins of PERT¹

The US Navy ‘Special Projects Office’, Bureau of Ordnance Navy (SPO) for the Fleet Ballistic Missile program (Polaris) was created on Nov. 17th 1955. During 1956 the ‘Plans and Programs Division’ of the SPO surveyed the control systems developed by other organisations used to working on large scale projects. In December 1956 a small development team comprising members of the SPO and consulting firm Booz, Allen and Hamilton was established to progress this work and develop a system for use on the program.

In January 1957, Admiral Raborn, head of the Polaris program, outlined his thoughts on the features of a control system and development work formally began on 27th January 1957, on a program called PERT, which was an acronym for ‘Programme Evaluation Research Task’. By July of the same year the name had changed to ‘Programme Evaluation Review Technique’. The team lead by Mr Wil Fazar quickly described the features of the system including ‘a precise knowledge of the sequencing of activities’ and ‘a careful time estimate for each activity, ideally with a probability estimate of the times the activity might require’. By July 1957 the first PERT procedures had been published, and by October PERT was being run on computers² and the system was applied to the Fleet Ballistic Missile programme.



The focus of the PERT analysis was to determine the probability of achieving a designated event.

¹ The **original 1959 PERT paper** can be found at: http://www.mosaicprojects.com.au/PDF/PM-History_PERT-Original_Paper.pdf.

² PERT was initially developed to run on the Navy’s NORC computer. The IBM Naval Ordnance Research Calculator (NORC) was a one-of-a-kind first-generation computer built by IBM for the United States Navy’s Bureau of Ordnance. It went into service in December 1954 and was likely the most powerful computer at the time. In 1955 NORC was moved to the Naval Proving Ground at Dahlgren, Virginia, and was their main computer until 1958, when more modern computers were acquired. It continued to be used until 1968.

Interestingly PERT was not widely used within the Polaris program but was widely publicised and used by Raborn to 'manage his external environment'. This publicity was very successful; first Polaris missile was launched in 1960, and PERT was credited with 'saving 2 years' on the program. By 1964 the PERT bibliography³ included more than 1000 books and articles!

The development of PERT closely followed the separate development of the CPM scheduling system by Du Pont. Apart from introducing uncertainty into schedule durations, the lasting contribution the PERT team have made to the business of scheduling was the invention of the 'critical path'. The Du Pont team lead by Kelly and Walker used the name 'main chain' for the longest path through their schedule⁴.

PERT and CPM were remarkably similar, both used the arrow diagramming technique (ADM - with arrows representing the activities)⁵. The key difference was that Du Pont's business was fundamentally known (construction and maintenance of chemical plants) and activity durations could be estimated with some degree of accuracy based on known quantities and production rates. Consequently, CPM was focused on optimising costs by balancing resources. The Navy's work on Polaris was largely R&D with cost a secondary issue; durations could only be assessed and PERT was focused on determining the probability of an event happening by some future date. However, by the late 1960s both CPM and PERT had merged into a generic ADM 'network based management system'.

The Precedence Diagramming Method (PDM) commonly used today was published as a non-computer approach to scheduling in 1962 and moved to its dominant place in the market on the back of the PC revolution of the 1980s. PDM uses boxes (or Nodes) to define the activities connected by links (or Arrows) to define the logical flow of work. From a practical perspective, PERT can be applied to either PDM or ADM networks provided the PDM network only uses Finish-to-Start links with a Zero duration. There is no capability in PERT to deal with other link type or the effect of Leads or Lags for the simple reason PERT was developed before the concept of a PDM network was available in the public domain.

The Basics of PERT

As outlined above, PERT originated as an Arrow Diagramming Method (ADM) scheduling technique and assumes each activity finishes at a node from which others start – there is no concept of overlapping activities in a basic CPM network⁶. To apply PERT to a schedule the following steps were needed:

- A logic network is developed with 3 time estimates for each activity⁷.
- The 'critical path' is determined using the PERT critical path (based on t_c) a calculation similar to CPM
- This longest path is then analysed using the PERT calculations to determine the probable completion date for the schedule.

³ One of the early publications was the specification issued by the **US Navy MIL-P-23189 PERT/TIME**, to encourage industry to take up the concept. Revision A was issued on the 25 October 1962 as: PERT/TIME AND PERT/COST MANAGEMENT INFORMATION SYSTEMS FOR PLANNING AND CONTROL. The standard was cancelled on the 1 February 1980.

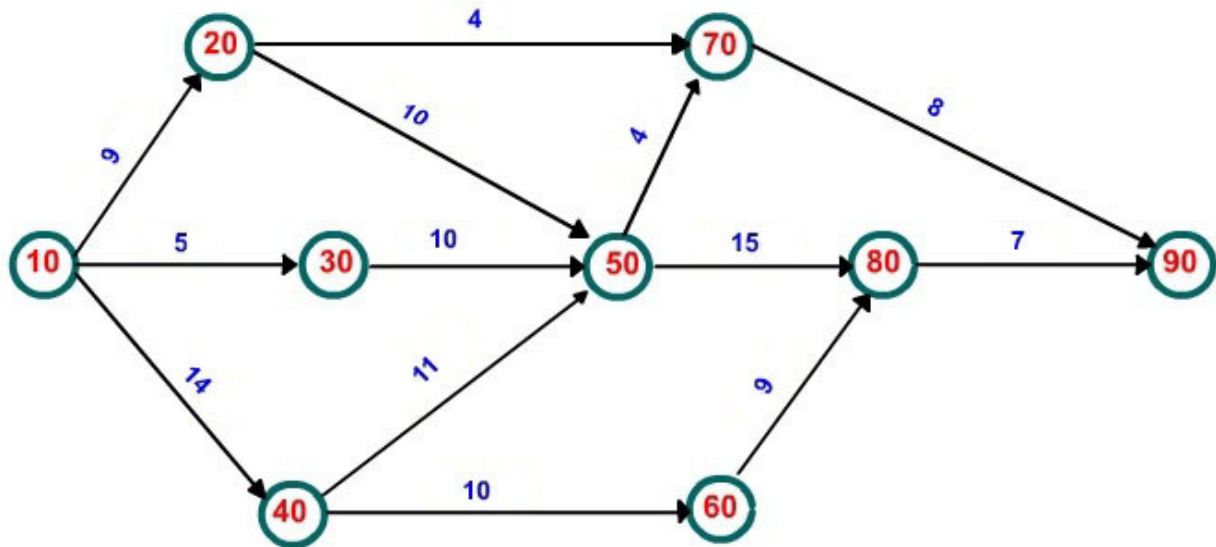
⁴ For more on the origins of CPM (including PDM and ADM) and PERT see **A Brief History of Scheduling**: http://www.mosaicprojects.com.au/Resources_Papers_042.html

⁵ For copies of the original CPM, PERT and PDM papers see: <http://www.mosaicprojects.com.au/PM-History.html#CPM>

⁶ The concept of 'Ladders' in ADM networks developed much later. For more on 'Ladders' see **Links, Lags & Ladders**: http://www.mosaicprojects.com.au/PDF/Links_Lags_Ladders.pdf

⁷ For more on **duration estimating** see: http://www.mosaicprojects.com.au/WhitePapers/WP1052_Time_Estimating.pdf





A PERT Network showing the 'Most Likely' duration for each Activity

Interestingly, no-one thought of writing activity descriptions on these early network diagrams!

The Basic PERT Calculation

The objective of the PERT calculations is to estimate an elapsed duration from start to end of an activity or project and establish the probability of achieving that end date. To achieve this three durations are estimated for each activity in the network; they are:

- **Optimistic** – the least time needed for the work (called either t_o or **a**)
- **Most Likely** – the expected time needed for the work (called either t_m or **b**)
- **Pessimistic** – the maximum time needed for the work (called either t_p or **c**)

The optimistic and pessimistic durations are based on normal working (ie, excluding defined risk events) and are nominally expected to have a 1:100 or 1% chance of being exceeded (pessimistic) or bettered (optimistic). Remember, it is almost impossible to estimate a duration that has no possibility of being exceeded (or bettered). Both PERT and Monte Carlo use the 1% rule to keep the data range reasonable⁸.

From this data, we can see the longest path through the above network, the 'CPM critical path' is 10 – 40 – 50 – 80 – 90. The most likely duration for this chain of activities is: $14 + 11 + 15 + 7 = 47$ but this CPM⁹ duration is optimistic!!

The 'expected' duration for each activity (t_e) is the **mean** (or average) duration which has a 50% chance of being achieved or bettered and a 50% chance of being exceeded. Using PERT, this is calculated as:

$$\text{Mean} = \frac{\text{optimistic} + (4 \times \text{most likely}) + \text{pessimistic}}{6}$$

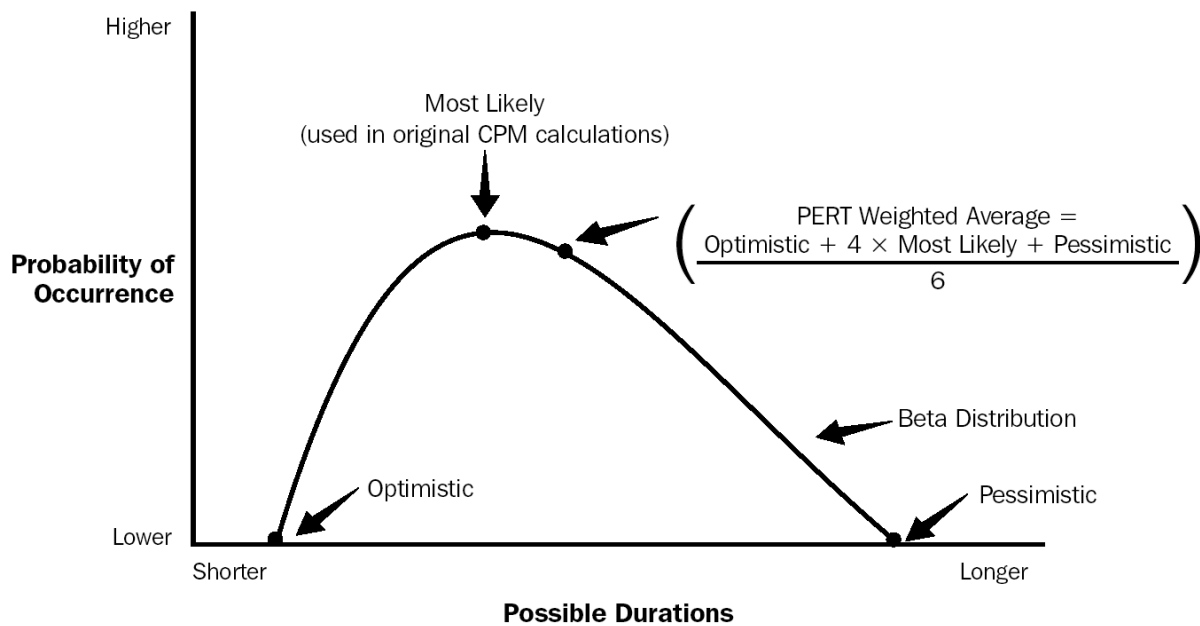
⁸ Early PERT models set the probability of achieving a duration less than the optimistic estimate at 1%, and the probability of achieving a duration greater than the pessimistic estimate at 1% (1% tails - but there were many other variations, 5% and 10% tails being common); but estimating these durations accurately proved nearly impossible resulting in 'garbage in - garbage-out'. Accurately assessing the 'range ends' remains a significant problem in the absence of reliable data.

⁹ For more on **CPM calculations** see: http://www.mosaicprojects.com.au/PDF/Schedule_Calculations.pdf

The traditional way the formula is expressed is: $\text{Mean} = \frac{a + 4b + c}{6}$

This PERT formula is based on the 'Modified Beta' distribution shown below¹⁰. What this distribution of possible durations for each activity shows is that whilst the most likely duration has a higher probability of occurring than the calculated Mean, it has a lower probability of being achieved. In project management, no one really cares what value is actually achieved – the object is to estimate an elapsed duration from start to end of an activity (or project) and establish the probability of achieving that end date. The Mean is at a point where 50% of the possible occurrences are earlier then the date, and 50% are later.

The PERT critical path is based on a similar concept to CPM - it is the longest sequential path through the network, but is calculated using the 'expected' or 'mean' (t_e) durations rather than the t_m durations - this means the PERT critical path may follow a different sequence of activities to the standard CPM critical path and will inevitably produce a longer overall schedule duration.



Based on the PERT formula, the calculations for the 'PERT critical path'¹¹ to determine the likely duration for event '90' (project completion) that has a 50% probability of being achieved is:

ID	A (t_o)	B (t_m)	C (t_p)	Expected (t_e)	SD	Var
10 - 40	12	14	18	14.33		
40 - 50	9	11	17	11.66		
50 - 80	11	15	19	15.00		
80 - 90	6	7	12	7.66		
		47		48.65		

¹⁰ This distribution (and a normal Beta Distribution) places a greater emphasis on the 'most likely' duration than the triangular distribution more commonly used today, which means PERT tends to produce a more optimistic average than a triangular distribution due to the pessimistic 'skew' in typical project data.

¹¹ On more complex networks the **PERT critical path** may not be the same as the CPM critical path, all of the T_e durations need to be calculated to determine the longest sequence of activities that constitute the **PERT critical path**.

Note, most of the activities have a ‘pessimistic shift’¹², this is because on most project activities the number of things that may cause the actual duration to be longer than expected are significantly greater than the number of things that may result in a shorter than expected duration. In this example the project has a 50% probability of completing during day 49 or earlier and a 50% probability of finishing later¹³.

Probability¹⁴ within PERT

652

Malcolm, Roseboom, Clark, and Fazar

tivity time, it was possible to develop equations for calculating t_e and $\sigma^2(t_e)$

$$t_e = \frac{1}{3} [2m + \frac{1}{2}(a+b)], \tag{1}$$

$$\sigma^2(t_e) = [\frac{1}{6}(b-a)]^2 \tag{2}$$

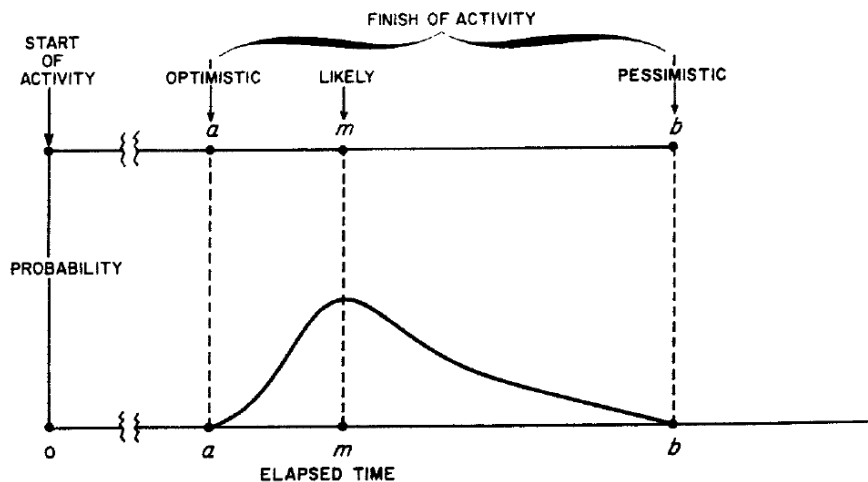


Fig 3. Estimating the elapsed-time distribution

Extracted from the 1959 PERT paper. If you multiply equation (1) above by 2 you get the modern version of $[a + 4m + b] / 6$ although the letters are used differently in the modern convention.

¹² In that original paper by Malcolm et al (at Footnote 1) they promise a follow up *paper by Clark* on how to do the analysis properly. That paper can be downloaded from: http://www.mosaicprojects.com.au/PDF/PM-History_PERT-Original_Paper.pdf.

To make the mathematics work, Clark has to assume a normal distribution, in contradiction to the beta assumption of the original paper, and tries to justify this approximation despite the ‘bell curve’ being uniform about the Mode. The math is quite intense and as he says it becomes totally intractable with more than two merging paths. (You can always treat a merging of more than two paths as a succession of merges of two, but this compounds the error due to the normal approximation, since even if the original paths have normally distributed lengths, the merged path won't.)

As Tony Welsh commented when passing on the information about this paper, “All in all it is an eloquent argument for Monte Carlo!”. Tony develops risk analysis software - <http://www.barbecana.com>

¹³ Probabilistic calculations effectively remove the option **to calculate ‘Float’**. Float is the difference between the earliest and latest dates on an activity – establishing ‘the’ early date and ‘the’ late date for this calculation requires a deterministic calculation (ie, CPM), probabilistic calculations by definition only provide a degree of certainty about any date. PERT and Monte Carlo calculate the probability of achieving a particular date (see below).

The PERT concept of ‘Slack’ was similar to ‘Float’ but applied to an event at the end of a merging path. The presence of ‘Slack’ meant the path was no considered in the PERT calculations leading to the issues of ‘PERT Merge Bias’. The calculation of slack is based on the t_e durations.

For more on calculating Float see: http://www.mosaicprojects.com.au/PDF/Schedule_Float.pdf

¹⁴ For more on **probability** see: http://www.mosaicprojects.com.au/WhitePapers/WP1037_Probability.pdf



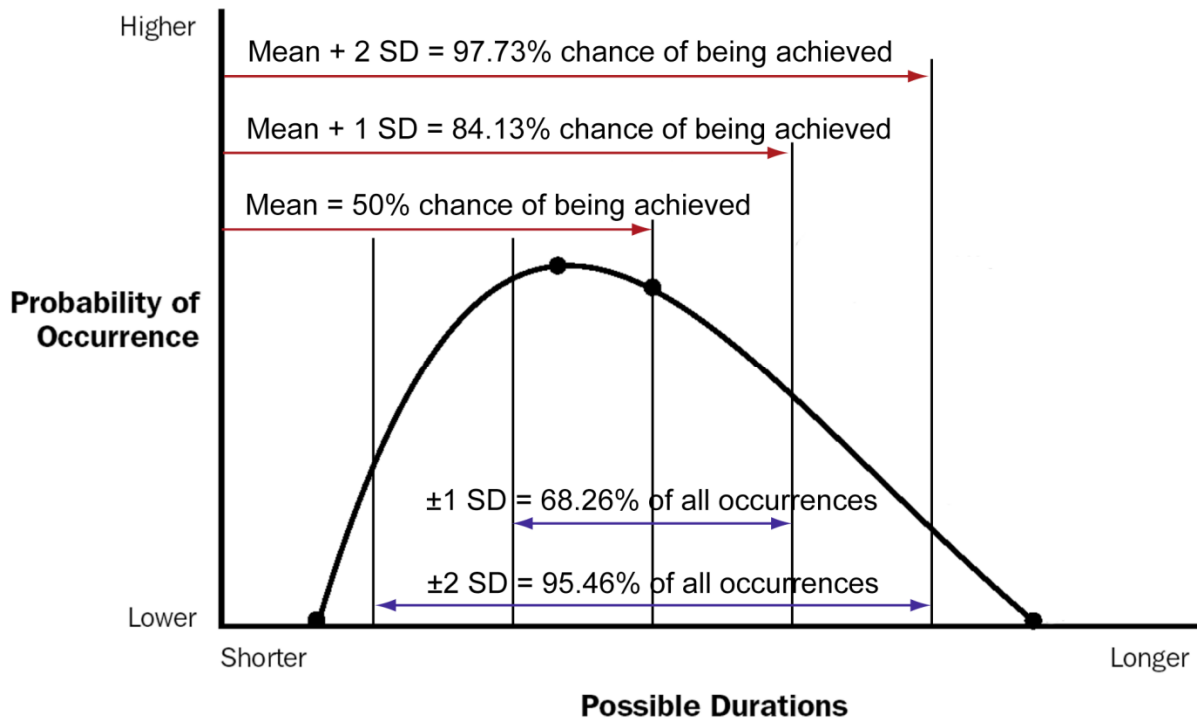
The PERT formula outlined above was a pragmatic simplification designed to facilitate reasonable calculation times. The system developers in their 1959 paper presented to the Institute for Operations Research (INFORMS) stated: *“This simplification gives biased estimates such that the estimated expected time of events are always too small. It was felt that, considering the nature of the input data, the utility of the other outputs possible from the data and the need for speedy, economical implementation the method [is] completely satisfactory.”* The primary simplification was to only consider the ‘controlling path’ – the path with the greatest value when analysing the variance at any particular node (typically the completion of the schedule).

Having established the ‘expected duration, to estimate a safer completion date with a better than 50% probability of being achieved (or to calculate the probability of achieving a target date), the variability in the data has to be considered; equation (2) above. This variability is measured by calculating the **Standard Deviation** within the data set. The size of ‘one standard deviation’ increases as the variability in the data increases, so that ‘one standard deviation’ will always have the same percentage of occurrences, or for our situation the same probability of completing.

Calculating the correct Standard Deviation (SD) for a data set is extremely complex. PERT uses a very simplistic approximation¹⁵:

$$\text{Standard Deviation} = \frac{\text{pessimistic} - \text{optimistic}}{6}$$

Once the Standard Deviation is known, the probabilities of completing are as follows:



The **standard deviation** (SD) calculation provides a range indicator for assessing the likelihood of achieving completion by a particular date (or more accurately, what the variability of the data set is). For each activity along the Critical Path, the calculation of its SD is: **SD = (c - a)/6**

¹⁵ The computing power available during the development of PERT was extremely limited. The developers used this simple approximation to allow their computations to run in a reasonable time.



However, to calculate the SD for a series of tasks, you cannot simply add up the individual SDs, you need to add their variances and then take the square root of the total; where the variance is the square of the individual standard deviations. For each activity: $Var = \{(c - a)/6\}^2$

Therefore, to calculate the duration that has an 84% probability of being achieved or bettered for our network, the calculation is:

ID	A	B	C	Expected	SD	Var
10 - 40	12	14	18	14.33	1.000	1.000
40 - 50	9	11	17	11.66	1.333	1.777
50 - 80	11	15	19	15.00	1.333	1.777
80 - 90	6	7	12	7.66	1.000	1.000
		47		48.65		5.444

The Square Root of 5.444 is **2.357**. This is the Standard Deviation for the whole series of activities constituting the ‘critical path’, add this to the Mean of 48.65 and you get: $48.65 + 2.357 = 51$ as the duration that has an approximately 80% chance of being achieved or bettered (the p80 duration)¹⁶.

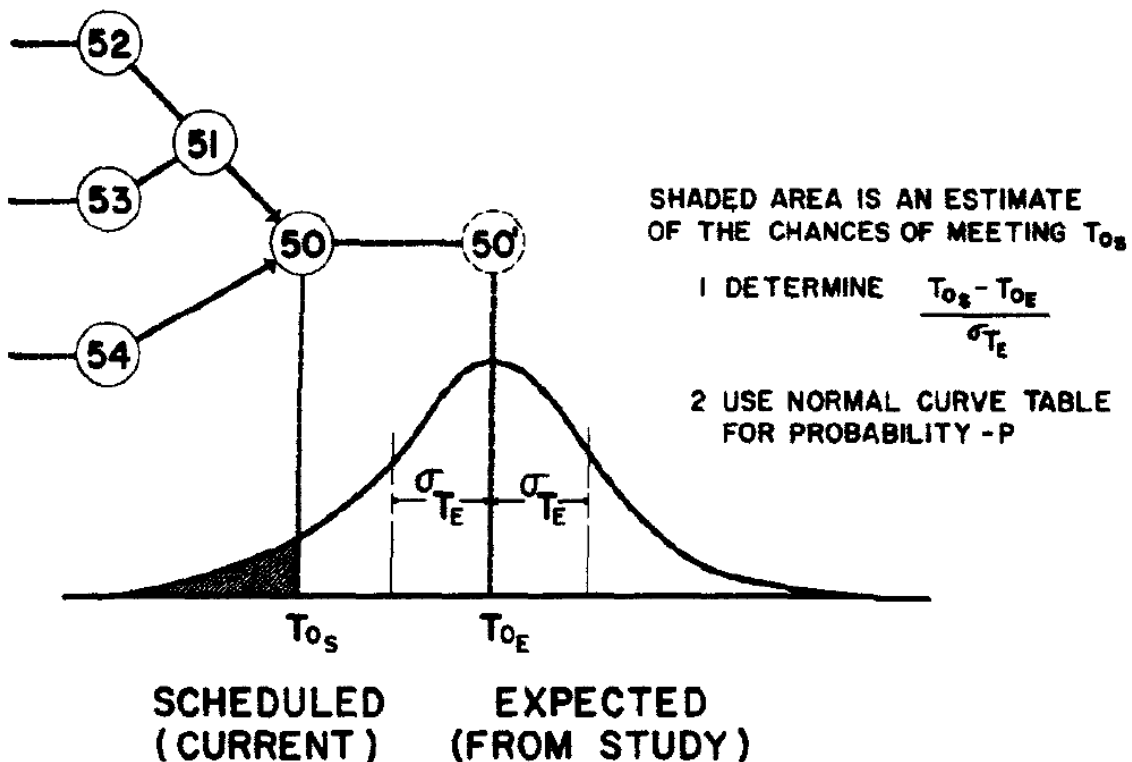


Fig. 9. Estimate of probability of meeting scheduled date, T_{Os}

The calculation of the probability of meeting a target or scheduled date the PERT developers used the central-limit theorem¹⁷ and an assumed probability distributing approximating a ‘normal distribution

¹⁶ For more on the importance of this see *Why Critical Path Scheduling (CPM) is Wildly Optimistic:* http://www.mosaicprojects.com.au/Resources_Papers_117.html

density'. In the extract from the 1959 paper above, the probability of meeting the scheduled date for event 50 at week 82 is calculated at 0.05, ie, very unlikely. The expected completion is week 92 and the Variance 38 weeks (giving a Standard Deviation for the controlling path of 6.164 weeks). This data was used to identify element of the program where the schedule was in jeopardy.

For its time PERT was a remarkably sophisticated tool. However, the compromises needed to make it into a useful tool that could run efficiently on the NORC computer created a number of problems and limitations.

Known Problems and Limitations within PERT

PERT has a number of limitations and known problems. The most significant are:

Time is the only variable

The effects of resource and technical performance changes enter into the analysis only in regard to their effect on elapsed time for the affected activity (Monte Carlo has a similar limitation).

Only One Distribution

The potential range of outcomes for any activity can vary depending on the type of work contained in the activity. PERT is limited to the single 'Modified Beta' distribution, modern systems allow a range of different distributions to more accurately reflect the variability inherent in each activity. Modern tools use: Triangular, Normal (Gaussian), Beta, Flat (no preference for any value between the max & min) and customisable distributions to name a few. This limitation in PERT reduces the accuracy of the calculation.

The Calculation for SD is an Approximation

The PERT calculation for the value of the SD is a significant source of error. Modern computers can easily manage the more complex calculations needed to accurately assess the actual SD for a data set.

The Single Pass Calculation

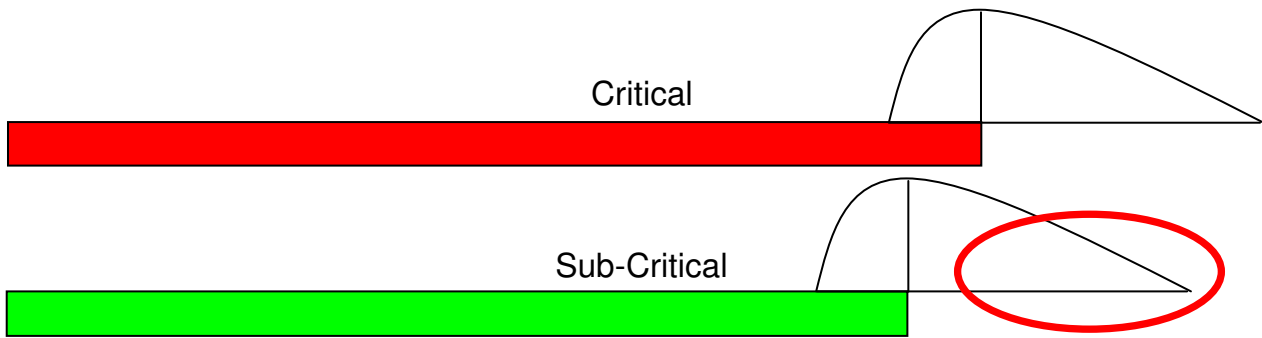
The single pass PERT calculation generates a single value. Techniques such as Monte Carlo run many hundreds of simulations that generate significantly richer and more accurate information.

The PERT Merge Bias

The single pass PERT calculation based on the 'critical path' ignores the possibility that other sub-critical paths may take longer to complete under certain circumstances. In the sketch above, if the critical path is achieving an outcome close to its 'most likely' date, but the sub-critical path is heading towards a pessimistic outcome, the project completion will be delayed, but the PERT calculations fail to recognise this¹⁸.

¹⁷ **Note:** the calculation of the 'expected duration' used the Modified Beta distribution, but the calculation of the probability of completing uses a Normal Distribution curve. This compromise was another pragmatic decision needed to make the calculations feasible. The calculated value was assumed to be sufficiently accurate to be useful for management decision making – in this example the real message is there is a greater than 90% probability of being late!

¹⁸ This problem was recognised by the original PERT developers (but thought acceptable) and publicly identified as the '**PERT Merge Bias**' in the early 1960s (MacCrimmon and Ryavec 1962; Archibald and Villoria 1967). The weakness of the PERT process is that it recognises only the PERT critical path and does not recognise that parallel slack paths can contribute to risk at every merge point. By ignoring the contribution of those slack paths PERT underestimates schedule risk at each merge point, as well as the overall schedule completion. However we cannot determine the extent of this bias without doing a Monte Carlo simulation. The 'merge bias effect' is greatest if a detailed schedule network is used for the Monte Carlo analysis, the effect is reduced if Monte Carlo is applied to a summary network.



As float is reduced on sub-critical paths, and/or the number of sub-critical paths with a duration close to the critical path increase, the probability of completing the overall project on time reduces but effect of the PERT Merge Bias hides this problem; the consequence is PERT predictions are always optimistic. Techniques such as Monte Carlo calculate the entire network each time and produce a more realistic projection of the likely range of completion dates¹⁹.

Some of the Things PERT Is NOT

The term PERT is widely used as a descriptor for many things that have nothing to do with the venerable technique outlined above²⁰. The practice of misusing terms to describe other things to look ‘clever’ or because the name is seen to be ‘sexy’ or commercially valuable is both dishonest and dangerous. In the modern age when someone says they are using PERT, you have absolutely no idea what they actually mean! Some of the major misuses are outlined below:

PERT ≠ a Network Diagram

This misnomer came from the early days of scheduling when PERT and CPM (both Activity-on-Arrow techniques) merged into a common ADM scheduling system. PERT requires a logic diagram to apply the PERT process and calculations but PERT is not synonymous with drawing a logic diagram; CPM, Critical Chain, RDM, MPM, and many other scheduling methodologies use network diagrams and are not PERT.

The use of the term ‘PERT Chart’ by Microsoft Project to describe its Precedence Diagram Network²¹ (PDM) has compounded the problem – PERT was developed before the paper on PDM was even published!

PERT ≠ Three Point Estimating

PERT uses three point estimating (optimistic, most likely, pessimistic) but again PERT is not synonymous with ‘three point estimating’. Any simulation or optimisation process needs a range statement; three point estimating is one way to generate this range. Once you have the ‘three points’ they can be used in Monte Carlo, optimisation systems, simulation systems and/or PERT.

¹⁹ Monte Carlo was first applied to schedules in 1963 by the Rand Corporation using the same data collected for the then ‘normal PERT’ analysis.

²⁰ The probable source of the misuse of the names **PERT Chart** and **Gantt Chart** is identified in our research paper: **Where did the Misuse of the names Gantt and PERT Originate?** A copy of this paper can be downloaded from: http://www.mosaicprojects.com.au/PDF_Papers/P173_misuse-terms-gantt-PERT.pdf

²¹ For more on the difference between ADM and PDM see **A Brief History of Scheduling:** http://www.mosaicprojects.com.au/Resources_Papers_042.html

PERT ≠ a Probability Calculation based on the Whole Network

The PERT calculation is based on the ‘critical path’ only. There is no process within PERT to calculate probabilities based on a whole network. Modern systems such as Monte Carlo, Latin Hypercube and others calculate probabilities based on the whole network. The techniques are different and should be called by their correct names.

In the 1970s and 80s there were numerous advances made on the basic PERT calculations using the more powerful mainframe computers then available. These tools were ‘PERT’ based, using a single Modified Beta distribution but looked at the whole network to reduce the effect of the PERT Merge Bias, ICL VME PERT was one such software tool. These proprietary systems disappeared with the mainframe computers and were always called by their commercial name. Since the 1980s Monte Carlo has been the dominant tool for analysing probabilities.

Conclusions

PERT is a venerable technique that made a major contribution to the development of project controls. The PERT process is described in the first part of this paper and in more detail in the original publication²². However, as described in this White Paper and the original 1959 presentation, the PERT developers made a range of pragmatic decision to develop a tool that could be used at the time. Because of its simplicity PERT is still a very useful teaching tool. Students can do the calculations and see the effect, even if the answers are less than accurate.

From a historical perspective, the limitations inherent in PERT were well understood by the mid 1970s and since that time more effective and accurate tools have been developed, most based on Monte Carlo and using the rapidly increasing power of contemporary computers. In the commercial world, PERT was effectively dead by the end of the 1970s for the simple reason more accurate tools such as Monte Carlo produce better answers.

Unfortunately predisposition of some people and organisations to use the word ‘PERT’ to describe any simulation and/or modelling technique and/or any network drawing technique either through a lack of awareness, or in an attempt to look clever, is a major concern²³. When someone says they are doing or using PERT you no longer know if they mean they are drawing a network diagram, using some form of modelling or simulation, using a full blown Monte Carlo analysis or actually doing PERT the way the developers of PERT intended. This confusion kills effective communication²⁴ and should be discouraged.

This White Paper is part of Mosaic’s **Project Knowledge Index** to view and download a wide range of published papers and articles see:
http://www.mosaicprojects.com.au/PM-Knowledge_Index.html

Mosaic’s complete PM History resource is available from:
<http://www.mosaicprojects.com.au/PM-History.html>

²² The **original 1959 PERT paper** can be found at:
http://www.mosaicprojects.com.au/PDF/PM-History_PERT-Original_Paper.pdf.

²³ For more on the communication problems caused by the **misuse of the name PERT** see our Blog:
<http://mosaicprojects.wordpress.com/2013/03/18/pert-whats-in-a-name/>

²⁴ For more on **communication theory** see:
http://www.mosaicprojects.com.au/WhitePapers/WP1066_Communcation_Theory.pdf

