MATHEMATICS OF OPERATIONS RESEARCH
Vol. 4, No. 4, November 1979
Printed in U.S.A.

INEQUALITIES FOR THE COMPLETION TIMES OF STOCHASTIC PERT NETWORKS*

LUC P. DEVROYE

McGill University

A stochastic PERT network is one in which the durations of the activities are random variables. In the absence of complete information about the distribution of these random variables or for the purpose of a fast rough analysis of the network, inequalities relating the mean $E\{T\}$ and the variance $\sigma^2\{T\}$ of the completion time T of the project (network) to the means and the variances of the individual random variables can be useful. In this note new upper bounds for $E\{T\}$ and $\sigma^2\{T\}$ are derived that are distribution free: they can be applied for all distributions of the durations of the activities.

1. Introduction. In a *PERT network* with n nodes the times needed to reach these nodes are called the *completion times* $T_1, \ldots, T_n$. The nodes are ordered such that whenever i and j are connected, then $i < j$. (Thus, $T_1 = 0$ since 1 is the starting node.) The problem addressed in this paper is the one of the relation between the completion times and the durations of the individual *activities*. When i and j are connected by an *arc* (ij) in the network then the duration of the activity (ij) is a random variable, T_{ij}. *Stochastic PERT networks* were studied by Burt and Garman [1], Charnes, Cooper and Thompson [2], Clingen [3], Elmaghraby [5], Fulkerson [7], Hartley and Wortham [10], Kleindorfer [11], MacCrimmon and Ryavec [13], Martin [14], Ringer [17] and Robillard and Trahan [18], [19]. Most of these papers are concerned with the distribution of T_n, or approximations of it, if the distributions of the T_{ij} are known. In this paper we are interested in what can be said about the distributions of the T_j if one merely knows the means

$$m_{ij} = E\{T_{ij}\}$$

and the standard deviations

$$s_{ij} = \sigma\{T_{ij}\}$$

of the durations of all the activities. We are only considering single-arc networks. If nodes i and j are connected via k arcs with corresponding duration times $T_{ij}(1), \ldots, T_{ij}(k)$, then we replace these arcs by a single arc (ij) with duration time $T_{ij} = \max(T_{ij}(1), \ldots, T_{ij}(k))$.

If B_j is the set of nodes which connect to node j via a single arc, then the T_j can be recursively defined by

$$T_1 = 0,$$
$$T_j = \max_{i \in B_j}(T_i + T_{ij}). \tag{1}$$

From this it easily follows by Jensen's inequality (Mitrinovic [15]) that

$$E\{T_j\} \geq E\left\{\max_{i \in B_j}(E\{T_i\} + T_{ij})\right\} \geq \max_{i \in B_j}(E\{T_i\} + E\{T_{ij}\}). \tag{2}$$

* Received April 24, 1978; revised September 1, 1978.
AMS 1970 subject classification. Primary: 90B35. Secondary: 90C35, 90C15.
IAOR 1973 subject classification. Main: Scheduling. Cross references: Network programming, stochastic programming.
Key words. Stochastic PERT networks, inequalities.

0364-765X/79/0404/0441$01.25

The first inequality is discussed by Fulkerson [7], Elmaghraby [5], Clingen [3], and Robillard and Trahan [18], [19] when the distributions of the T_{ij} are known.

From the second inequality in (2) it follows that if we define the numbers $\underline{E\{T_j\}}$ recursively by

$$\underline{E\{T_1\}} = 0 \quad \text{and}$$

$$\underline{E\{T_j\}} = \max_{i \in B_j}\left(\underline{E\{T_i\}} + m_{ij}\right) \tag{3}$$

then $\underline{E\{T_j\}} \leqslant E\{T_j\}$ for all j. Thus, lower bounds on the $E\{T_j\}$ that only depend upon the mean duration times m_{ij} are easy to compute. In a practical situation, however, upper bounds seem more important because they provide one with information about when a project (the complete network) is expected to be finished. In this note we present a simple recursive method for the computation of upper bounds for the $E\{T_j\}$ and $\sigma\{T_j\}$ from the m_{ij} and s_{ij} if the T_{ij} are independent random variables.

2. Main results. Let n_j be the number of arcs arriving at the jth node and let the number sequences $\overline{E\{T_j\}}$ and $\overline{\sigma^2\{T_j\}}$ be recursively defined by

$$\overline{\sigma^2\{T_1\}} = 0; \qquad \overline{\sigma^2\{T_j\}} = \sum_{i \in B_j}\left(\overline{\sigma^2\{T_i\}} + s_{ij}^2\right); \tag{4}$$

$$\overline{E\{T_1\}} = 0; \quad \text{and}$$

$$\overline{E\{T_j\}} = \max_{i \in B_j}\left(\overline{E\{T_i\}} + m_{ij}\right) + \sqrt{n_j \max_{i \in B_j}\left(\overline{\sigma^2\{T_i\}} + s_{ij}^2\right)} \tag{5}$$

or

$$\overline{E\{T_j\}} = \max_{i \in B_j}\left(\overline{E\{T_i\}} + m_{ij}\right) + \sqrt{(n_j - 1)\left(\max_{i \in B_j}\left(2\,\overline{\sigma^2\{T_i\}} + s_{ij}^2\right) + \min_{i \in B_j}\left(2\,\overline{\sigma^2\{T_i\}} + s_{ij}^2\right)\right)}. \tag{6}$$

THEOREM. *If the T_{ij} are independent random variables, then*

$$\sigma^2\{T_j\} \leqslant \overline{\sigma^2\{T_j\}}, \qquad E\{T_j\} \leqslant \overline{E\{T_j\}}, \qquad 1 \leqslant j \leqslant n.$$

REMARK 1. The bounds following from (4)–(6) are distribution free: they apply to *all* cases; in particular, the T_{ij} can be continuous or discrete random variables or mixtures of both. Without additional information about the distributions of the T_{ij} it is unreasonable to expect very tight bounds. One should notice, however, that some of the theoretical inequalities (e.g., Inequality 1 from the Appendix) used to arrive at (5) and (6) are the best possible distribution-free bounds involving means and standard deviations only. Furthermore, for deterministic PERT networks (that is, $s_{ij} = 0$ for all i, j) (3) and (4)–(6) are identical in form. This means that in the deterministic case, $\underline{E\{T_j\}} = E\{T_j\} = \overline{E\{T_j\}}$ for all j, and that one can therefore expect reasonable tightness for stochastic PERT networks in which the standard deviations s_{ij} are small compared to the means m_{ij}.

REMARK 2. If only one arc arrives at node j, say arc (kj), then (6) reduces to

$\overline{E\{T_j\}} = \overline{E\{T_k\}} + m_{kj}$. (6) can replace (5) whenever the resulting value for $\overline{E\{T_j\}}$ is lower.

REMARK 3. From the Cantelli-Frechet-Uspensky inequality (Frechet [6, p. 137]; Uspensky [21, p. 198]) it is known that for any random variable X with mean m and standard deviation s, $P\{X - m \geqslant cs\} \leqslant 1/(1 + c^2)$ and $P\{X - m \leqslant -cs\} \leqslant 1/(1 + c^2)$. We can therefore conclude from Theorem 1 that

$$\begin{aligned} &P\left\{T_j \geqslant \overline{E\{T_j\}} + c\,\overline{\sigma\{T_j\}}\right\} \leqslant 1/(1 + c^2) \quad \text{and} \\ &P\left\{T_j \leqslant \underline{E\{T_j\}} - c\,\overline{\sigma\{T_j\}}\right\} \leqslant 1/(1 + c^2) \end{aligned} \tag{7}$$

for all j and all $c > 0$.

REMARK 4. Notice that $\overline{\sigma^2\{T_j\}}$ is the sum of all the s_{ik}^2 corresponding to nodes k from which j can be reached (ancestor nodes).

REMARK 5 (multiple arcs). For the case that

$$T_{ij} = \max_{1 \leqslant l \leqslant n_{ij}} (T_{ij}(l))$$

where n_{ij} is the multiplicity of the arc (ij), and $T_{ij}(l)$ has mean $m_{ij}(l)$ and variance $s_{ij}^2(l)$, all that is said above remains true if in (3) we replace m_{ij} by

$$\max_{1 \leqslant l \leqslant n_{ij}} m_{ij}(l),$$

and in (4)–(6) we replace s_{ij}^2 by $\sum_{l=1}^{n_{ij}} s_{ij}^2(l)$, and m_{ij} by either

$$\max_{1 \leqslant l \leqslant n_{ij}} m_{ij}(l) + \sqrt{n_{ij} \max_{1 \leqslant l \leqslant n_{ij}} s_{ij}^2(l)}$$

or

$$\max_{1 \leqslant l \leqslant n_{ij}} m_{ij}(l) + \sqrt{(n_{ij} - 1)\left(\max_{1 \leqslant l \leqslant n_{ij}} s_{ij}^2(l) + \min_{1 \leqslant l \leqslant n_{ij}} s_{ij}^2(l)\right)}\ .$$

Both changes follow trivially from Jensen's inequality and Inequality 1 given in the next section.

REMARK 6 (tightness of the bounds). One may be interested in finding out when the bounds derived here are tight. For easy analysis and quick insight, consider two extreme cases:

(i) A chain-type network where $T_1 = 0$, and $T_j = T_{j-1} + T_{j-1j}$, all j.

(ii) A 2-node parallel network with $T_1 = 0$, and $T_2 = \max(T_{12}(1), \ldots, T_{12}(N))$.

All other networks can be constructed by proper combinations of networks of types (i) and (ii). It is clear that for (i) we have exact bounds:

$$\underline{E\{T_j\}} = E\{T_j\} = \overline{E\{T_j\}}, \qquad \overline{\sigma\{T_j\}} = \sigma\{T_j\}.$$

For the network (ii) the answer depends upon the tails of the distributions of the duration times. Assume for the moment that all $T_{12}(l)$ are distributed as Y with mean m and standard deviation s. Our bounds are $\underline{E\{T_2\}} = m$ and $\overline{E\{T_2\}} = m + s\sqrt{N}$ if we use (3) and (5). If Y has ess sup $Y = c < \infty$, then $E\{T_2\} \to c$ as $N \to \infty$. Thus, for bounded random variables with relatively large standard deviation s (relative to $c - m$) appearing in short heavily branched networks, our upper bound is not attractive. The situation improves with the size of the tail (here, the ratio $(c - m)/s$). Consider now an infinite-tailed distribution such as the exponential: $P\{Y > y\} = \exp(-y), y > 0$. Clearly, $m = s = 1$ and $E\{T_2\} \sim \log N$ (see Gumbel [8]). If Y has a still heavier tail,

$$P\{Y > y\} = \begin{cases} 1, & 0 \leq y \leq 1, \\ y^{-p}, & y > 1, \end{cases}$$

where $p > 2$ (otherwise, $\sigma\{Y\} = \infty$), then

$$E\{T_2\} = \int_0^1 N u^{-1/p}(1-u)^{N-1} du = O(N^{1/p}).$$

Thus the rate of increase of $E\{T_2\}$ as $N \to \infty$ can be made arbitrarily close to the rate of $\overline{E\{T_2\}} = m + s\sqrt{N}$ for p close to 2. In these cases, the classical lower bound $\underline{E\{T_2\}}$ seems almost useless. We conclude that our upper bounds seem best suited for long stretched-out networks and networks with duration times that have heavy tails (that is, in situations with high uncertainty about the duration of one or more jobs). For short fat networks with low-variance duration times another approach seems necessary. One involving Chernoff-Bernstein like exponential inequalities is currently being studied by the author.

3. Proofs. The crucial result used in the proof of the theorem is an upper bound for $E\{\max_i X_i\}$ in terms of $E\{X_i\}$ and $\sigma\{X_i\}$, $1 \leq i \leq n$, where $X_1, \ldots, X_n$ is *any* sequence of random variables (possibly dependent). For independent identically distributed random variables, bounds of this type are discussed in David's book [4]. The original inequalities are developed in a series of papers by Moriguti [16], Gumbel [8], Hartley and David [9] and Rustagi [20].

Inequality 1. If $X_1, \ldots, X_n$ is an arbitrary sequence of random variables with finite means $E\{X_1\}, \ldots, E\{X_n\}$ and finite variances $\sigma^2\{X_1\}, \ldots, \sigma^2\{X_n\}$, then

$$E\left\{\max_i X_i\right\} \leq \max_i E\{X_i\} + \sqrt{n}\, \max_i \sigma\{X_i\} \tag{8}$$

and

$$E\left\{\max_i X_i\right\} \leq \max_i E\{X_i\} + \sqrt{n-1}\, \min_j \max_{i \neq j} \sigma\{X_i - X_j\}. \tag{9}$$

PROOF. Define the events A_i, $1 \leq i \leq n$, as follows:

$$A_i = \left\{X_i > \max_{j<i} X_j\right\} \cap \left\{X_i \geq \max_{j>i} X_j\right\}.$$

Clearly, the A_i are disjoint and exhaust the entire space. Further,

$$E\left\{\max_i X_i\right\} = \sum_{i=1}^n E\{X_i I_{\{A_i\}}\} = \sum_{i=1}^n E\{E\{X_i I_{\{A_i\}} \mid X_i\}\}$$

$$= \sum_{i=1}^n E\{X_i E\{I_{\{A_i\}} \mid X_i\}\} \leq \sum_{i=1}^n \left(E\{X_i^2\} E\left\{(E\{I_{\{A_i\}} \mid X_i\})^2\right\}\right)^{1/2}.$$

Here we used standard properties of conditional expectations [12]. The symbol I denotes the indicator function of an event. The last inequality in the chain is a form of the Cauchy-Schwartz inequality (see Loeve [12, p. 156]). Let us assume that $E\{X_1\} = E\{X_2\} = \cdots = E\{X_n\} = 0$. Then

$$E\left\{X_i^2\right\} \leq \max_i \sigma^2\{X_i\}.$$

Regardless of the zero mean assumption we always have

$$E\left\{(E\{I_{\{A_i\}} \mid X_i\})^2\right\} \leq P\{A_i\}.$$

Thus,

$$E\left\{\max_i X_i\right\} \leqslant \max_i \sigma\{X_i\} \sum_{i=1}^{n} \sqrt{P\{A_i\}}$$

$$\leqslant \max_i \sigma\{X_i\} n \sqrt{\sum_{i=1}^{n} P\{A_i\}/n} = \sqrt{n}\,\max_i \sigma\{X_i\}, \qquad (10)$$

by Jensen's inequality and the fact that $P\{A_1\} + \cdots + P\{A_n\} = 1$. Next, since $E\{\max_i X_i\} = E\{\max_i X_i - X_1\} + E\{X_1\} = E\{\max_i X_i - X_1\} = E\{\max_{i \geqslant 2}(X_i - X_1)_+\}$ where $u_+ = \max(u, 0)$, and since $E\{(X_i - X_1)_+^2\} \leqslant E\{(X_i - X_1)^2\}$ we have

$$E\left\{\max_i X_i\right\} \leqslant \sqrt{n-1}\,\max_{i \geqslant 2} \sigma\{(X_i - X_1)\}.$$

The choice of X_1 was arbitrary. Hence, the following is also true:

$$E\left\{\max_i X_i\right\} \leqslant \sqrt{n-1}\,\min_j \max_{i \neq j} \sigma\{(X_i - X_j)\}.$$

Finally, if the X_i's are arbitrary, we can use

$$E\left\{\max_i X_i\right\} \leqslant E\left\{\max_i (X_i - E\{X_i\})\right\} + \max_i E\{X_i\}. \quad ∎$$

Note. If $X_1, \ldots, X_n$ are also independent then $\sigma^2\{X_i - X_j\} = \sigma^2\{X_i\} + \sigma^2\{X_j\}$, and (9) reduces to

$$E\left\{\max_i X_i\right\} \leqslant \max_i E\{X_i\} + \sqrt{n-1}\sqrt{\max_i \sigma^2\{X_i\} + \min_i \sigma^2\{X_i\}} \qquad (11)$$

which is at least as sharp as (8) whenever

$$\max_i \sigma^2\{X_i\} / \min_i \sigma^2\{X_i\} \geqslant n - 1.$$

Inequality 2. If $X_1, \ldots, X_n$ are random variables (possibly dependent) then

$$\sigma^2\left\{\max_i X_i\right\} \leqslant \sum_{i=1}^{n} \sigma^2\{X_i\}. \qquad (12)$$

PROOF.

$$\sigma^2\left\{\max_i X_i\right\} = E\left\{\left(\max_i X_i - E\left\{\max_i X_i\right\}\right)^2\right\}$$

$$\leqslant E\left\{\left(\max_i X_i - \max_i E\{X_i\}\right)^2\right\}$$

$$\leqslant E\left\{\max_i (X_i - E\{X_i\})^2\right\}$$

$$\leqslant E\left\{\sum_{i=1}^{n} (X_i - E\{X_i\})^2\right\} = \sum_{i=1}^{n} \sigma^2\{X_i\}. \quad ∎$$

Note. One would expect a tighter bound of the type

$$\sigma^2\left\{\max_i X_i\right\} \leqslant c \max_i \sigma^2\{X_i\}, \quad \text{some } c > 0. \qquad (13)$$

However, even if the X_i's are independent, the best possible c cannot be smaller than n, which makes (13) weaker than (12). Indeed, consider n binomial $\{0, 1\}$-valued random variables with equal mean p. Since $\max_i X_i$ is a $\{0, 1\}$-valued binomial random variable with mean $1-(1-p)^n$, it follows that

$$\frac{\max_i \sigma^2\{X_i\}}{\sigma^2\{\max_i X_i\}} = \frac{p(1-p)}{(1-p)^n(1-(1-p)^n)} \overset{p\to 0}{\to} \frac{1}{n}$$

and thus that $c \geqslant n$.

PROOF OF THE THEOREM. From Inequality 2 and the independence of T_i and T_{ij} for all i, it follows that

$$\sigma^2\{T_1\} = 0$$

and

$$\sigma^2\{T_j\} \leqslant \sum_{i\in B_j} \sigma^2\{(T_i + T_{ij})\} = \sum_{i\in B_j}(\sigma^2\{T_i\} + s_{ij}^2).$$

Also, Inequality 1 implies that $E\{T_1\} = 0$ and that

$$E\{T_j\} \leqslant \max_{i\in B_j} E\{T_i + T_{ij}\} + \sqrt{n_j}\,\max_{i\in B_j}\sigma\{T_i + T_{ij}\}$$

$$= \max_{i\in B_j}(E\{T_i\} + m_{ij}) + \sqrt{n_j \max_{i\in B_j}(\sigma^2\{T_i\} + s_{ij}^2)}\ .$$

This shows the validity of the theorem for the inequalities (4) and (5). At all times (5) can be replaced by (6) because from (9) we have

$$E\{T_j\} \leqslant \max_{i\in B_j}(E\{T_i\} + m_{ij}) + \sqrt{n_j - 1}\ \min_{k\in B_j}\ \max_{\substack{i\neq k\\ i\in B_j}}\sigma\{T_i + T_{ij} - T_k - T_{kj}\}.$$

By the independence of $T_i - T_k$, T_{ij} and T_{kj} we have

$$\sigma^2\{T_i - T_k + T_{ij} - T_{kj}\} = \sigma^2\{T_i - T_k\} + s_{ij}^2 + s_{kj}^2 \leqslant 2\sigma^2\{T_i\} + s_{ij}^2 + 2\sigma^2\{T_k\} + s_{kj}^2.$$

(Here we use the fact that for any random variables X and Y, $\sigma^2\{X+Y\} \leqslant 2\sigma^2\{X\} + 2\sigma^2\{Y\}$.) The applicability of (6) now follows trivially. ∎

Acknowledgement. The author is indebted to Professor Tim Merrett for bringing the stochastic PERT network problem to his attention and for the subsequent discussions that led to the present results.

References

[1] Burt, J. M. and Garman, M. B. (1971). Monte Carlo Techniques for Stochastic PERT Network Analysis. *Information* **9** 248–262.

[2] Charnes, A., Cooper, W. W. and Thompson, A. L. (1964). Critical Path Analyses via Chance Constrained and Stochastic Programming. *Operations Res.* **12** 460–470.

[3] Clingen, C. T. (1964). A Modification of Fulkerson's Algorithm for Expected Duration of a PERT Project When Activities Have Continuous d.f. *Operations Res.* **12** 629–632.

[4] David, H. A. (1970). *Order Statistics*. John Wiley, New York.

[5] Elmaghraby, S. E. (1967). On the Expected Duration of PERT Type Networks. *Management Sci.* **13** 299–306.

[6] Frechet, M. (1950). *Generalites sur les Probabilities. Elements Aleatoires*. Gauthier-Villars, Paris.

[7] Fulkerson, D. R. (1962). Expected Critical Path Lengths in PERT Networks. *Operations Res.* **10** 808–817.

[8] Gumbel, E. J. (1954). The Maxima of the Mean Largest Value and of the Range. *Ann. Math. Statist.* **25** 76–84.

[9] Hartley, H. O. and David, H. A. (1954). Universal Bounds for Mean Range and Extreme Observation. *Ann. Math. Statist.* **25** 85–99.

[10] ——— and Wortham, A. W. (1966). A Statistical Theory for PERT Critical Path Analysis. *Management Sci.* **12** B469–B481.

[11] Kleindorfer, G. B. (1971). Bounding Distributions for a Stochastic Acyclic Network. *Operations Res.* **19** 1586–1601.
[12] Loeve, M. (1964). *Probability Theory*. Van Nostrand, Princeton, New Jersey.
[13] MacCrimmon, K. R. and Ryavec, C. A. (1964). An Analytical Study of the PERT Assumptions. *Operations Res.* **12** 16–37.
[14] Martin, J. J. (1965). Distribution of Time through a Directed, Acyclic Network. *Operations Res.* **13** 46–66.
[15] Mitrinovic, D. S. (1970). *Analytic Inequalities*. Springer-Verlag, Berlin.
[16] Moriguti, S. (1951). Extremal Properties of Extreme Value Distributions. *Ann. Math. Statist.* **22** 523–536.
[17] Ringer, L. J. (1969). Numerical Operators for Statistical PERT Critical Path Analysis. *Management Sci.* **16** B136–B143.
[18] Robillard, P. and Trahan, M. (1976). Expected Completion Time in PERT Networks. *Operations Res.* **24** 177–183.
[19] ——— and ———. (1977). The Completion Time of PERT Networks. *Operations Res.* **25** 15–29.
[20] Rustagi, J. S. (1957). On Minimizing and Maximizing a Certain Integral with Statistical Applications. *Ann. Math. Statist.* **28** 309–328.
[21] Uspensky, J. V. (1937). *Introduction to Mathematical Probability*. McGraw-Hill, New York.

SCHOOL OF COMPUTER SCIENCE, BURNSIDE HALL, MCGILL UNIVERSITY, 805 SHERBROOKE STREET WEST, MONTREAL, QUEBEC, CANADA H3A 2K6