PROJECT RISK ANALYSIS

BRAD PAYNE BSc (Hons)

A thesis submitted in partial fulfilment of the requirements of The Nottingham Trent University for the degree of Doctor of Philosophy

Sponsoring Establishment: The Nottingham Trent University

November 1998

BRAD PAYNE: Project Risk Analysis

This thesis reviews the history and literature of Project Risk Analysis (PRA) and provides a conceptual model where time, cost and quality characteristics are modelled. Investigation of project managers requirements, specifically in the construction industry, reveals the need for a model incorporating the operation performance of the project outcome. The concept of project quality loss is introduced and models explored throughout. A representation of a project as a set of sequential phases is developed which provides the framework for modelling the project operation measures reliability, availability and output.

Investigation of current methods for performing a project risk analysis of a single characteristic such as time, revealed the choice of approach needs to depend on the activity distribution used, the network configuration and the desired accuracy. A Laplace transform approach is developed where the explicit probability function is obtained for project completion time assuming activity distributions are special Erlangian. An algorithm, based on the network reduction method of Dodin (1984a), is provided and implemented with the developed Laplace transform approach within the *Mathematica* system. Network examples are investigated, including the 'Forbidden Network' configuration of Ringer (1969) and comparisons made to the approaches PERT, method of moments and simulation.

Dependency relationships are investigated between the characteristics time, cost and quality. A bivariate exponential extension distribution is developed to model time-cost dependency of a project phase and a method defined for obtaining the convolution of n such densities. Time-quality loss relationships are discussed and a model developed.

A generalised model of the performance of a project is formulated as a stochastic process where the state space is partly discrete and partly continuous. Examples where the time, cost and quality characteristics are measured and discussion of possible Markov dependencies are presented in addition to a quality loss dependency model using the uniform distribution.

The dependency modelling capabilities of the generalised model are extended and a simulation program is developed to analyse the dependency between the characteristics time, cost and quality loss and the dependency between phases.

OBJECTIVES

The main objectives of the research for this thesis are:

(i) To analyse and extend the techniques available for performing a project risk analysis.

(ii) To investigate the relationship between the duration, cost and quality characteristics of performing a project and provide a general formulation to investigate the consequences of the dependencies between time, cost and quality.

(iii) To model the dependencies of project performance in terms of time, cost and quality on the operation of the project outcome.

(iv) To develop a project quality loss measure that enables assessment of operation measures.

(v) To develop, implement and test an algorithm to perform a project risk analysis of single characteristics using Laplace transforms.

(vi) To develop models relating the project characteristics; time, cost and quality.

(vii) To develop software that enables modelling of the performance of successive project phases dependent on the realisation of completed phases with predetermined targets.

ADVANCED STUDIES

The following advanced studies were undertaken in connection with the programmes of research for this thesis:

(i) Participation in research seminars at The Nottingham Trent University.

(ii) Participation in the following conference: Poster display at the Royal Statistical Society, International Conference, Newcastle Upon Tyne 1994.

(iii) Attendance at the following conference: Reliability & Quality in Design, International Society of Science and Applied Technologies, Seattle, Washington, U.S.A. 1994.

(iv) Presentation and discussion of research studies with Rolls Royce plc, Bristol - September, 1994, Derby - December 1994.

(v) Attendance of seminars of the East Midland Quality club, viz:

Rapid Prototyping, Nottingham University, 1997.

World Quality Day, Nottingham, 1996, 1997.

(vi) Professor Bendell and the author presented aspects of risk dependency modelling, specifically the time-cost model and the dependence of operating reliability on project quality concept at Luleå University of Technology, Sweden, 1996.

(vii) Appropriate reading.

Throughout the period of registration, the author has undertaken some part-time lecturing in the form of Mathematical, Statistical, Computing and Management Science lectures and tutorials for a number of courses, viz:

Quantitative methods for Occupational Health and Safety
Quantitative Methods for Business Studies I
I.T. and Communication Skills with Statistics I
Statistics for Science I
Business Information Processing and Modelling II
Nonparametric Statistics II
Management Science II & III

ACKNOWLEDGEMENTS

The author would like to thank The Nottingham Trent University for the resources provided and their assistance.

Thanks to Angus Tozer and Barry Cooper of the Strategic Planning department of Rolls-Royce plc for helpful discussion in the early stages of research, Steve Butler of VSEL for his views as a project planner of large scale development programmes, and to John Bailey of Sol construction for advice regarding the construction industry. Thanks to Anilesh Mohammi for his help with the 'forbidden network'.

The author is immensely grateful to his supervisors, Professor A. Bendell, Mark Carter and Dr. D. Wightman for their help and guidance. Special thanks to Dr. D. Wightman, a good supportive friend who sadly passed away in June 1997, and to whom this thesis is dedicated to. Thanks to Professor A. Bendell for the encouragement and support throughout the hard times and final stages of performing this thesis.

Thanks is also due to his friends, flatmates, and family for their support and belief in the author. Thanks to the special lady, Mrs Anne Marie Rose for her friendship and kindness. Lastly many thanks to Chris Wensley for her time, delicious fajitas and unfailing support.

Contents

List of Figures

List of Tables

Chapter 1

Project Risk Analysis

1.1 Introduction

The explicit treatment of Project Risk is a relatively new concept which, to many practitioners, is viewed as an extension of project planning techniques developed almost forty years ago. The first project planning method that aroused much interest within the field of Operational Research and to practitioners of project planning was the Project Evaluation and Review Technique, commonly referred to as 'PERT'. It was developed in the late 1950's as the project management aid for the development and production of the Polaris missile system, Bureau of Naval Weapons (1958).

A similar technique, applied in 1957 at the DuPont Company, focused on reducing the cost of the construction of a chemical plant, DuPont (1958). This technique was given the name of Critical Path Method which is typically referred to as CPM. It is a deterministic approach as a single time value is used for each activity, however CPM has the added advantage that trade-offs may be made between project duration and project cost. For simplicity it is often assumed that the relationship between time and cost is linear.

Both PERT and CPM assume a strict sequence of activities where the start of an activity may only commence upon the completion of all preceding activities. Also there may only be a single starting point and a single finishing point. In addition all activities are assumed known and the repetition of activities is not allowed. These assumptions are applied with the Precedence Diagramming Method, PDM which is explained in Moder et al (1983), and is typically used in conjuction with PERT and CPM.

PERT and CPM are presented in most Management Science and Operational Research textbooks. A practitioner's approach to project management covering both PERT and CPM is presented by Moder et al (1983). In most cases the PERT technique is used solely to assess the expected duration of a project, and the CPM is used for controlling project schedules.

During the 1960's material criticising the PERT technique began appearing in Operational Research journals. Particular texts of interest include those by Grubbs (1962) reporting on the inaccuracies of modelling activity times by the beta probability distribution, and Ringer (1969) who recognises that the expected project completion time using PERT is biased optimistically. This work initiated much research into modifications of the PERT technique and further methods to assess the project completion time.

The studies performed may be classified into five categories as reported by Soroush (1993). Analytical studies cover both numerical techniques and bounding approaches. Monte Carlo simulation techniques first discussed by Van Slyke (1963) and improved by Cook and Jennings (1979), Ragsdale (1989), have recently become more attractive as simulation computation time is reduced considerably due to the advances in computer technology. The third category was reported as analytical approaches to estimate the expected project completion time. A study of criticality indices for both paths and activities was developed to assess which project activities were more likely to influence the overall project completion time. Indices were first computed using simulation by Van Slyke (1963) and approximations determined by Dodin and Elmaghraby (1985). The last category concerns the direct study of the three estimate PERT method.

Due to the complexity of most of the techniques only PERT and CPM appear to have been reported as being applied in practise. The extent of the use of management methods during the 1960's and early 1970's was limited as indicated by a survey within the construction industry, Davis (1974), where only 55% of companies applied project management techniques. However due to the variety of PERT type analysis software utilising simulation procedures, analysis of these methods is now more accessible to non-experts. Areas of application include research and development programmes, production processes, construction projects, and maintenance schedules.

Variations of the PERT technique include GERT, Pritsker and Burgess (1970), PNET, Ang et al (1975), and VERT, Moeller (1972). The GERT approach, which stands for Graphic Evaluation and Review Technique, was developed to address certain projects, mainly research and development programmes where contingencies often exist through the evolution of a project, and predetermined project activities are not necessarily known. It uses simulation techniques to model the project performance. At this point in time, computer technology, even though was developing rapidly, had not advanced sufficiently to perform the analysis of large project networks without considerable expense. It was this reason that Ang et al (1975) developed the PNET algorithm which was considered an improvement over PERT. It is simple to apply and does not rely on computer technology. Unlike many techniques the PNET approach can be applied to realistic projects. Even though oversimplified, included in the article of Ang et al (1975) is a study of a road pavement project and an industrial building project.

It was at this time that the term 'risk' began appearing in the project planning literature. Although not formally defined it was commonly interchanged with the term probability and used to consider network decision problems such as the chance of completing or not completing the project within a designated period. The VERT method was developed by Moeller (1972) and has since had many modifications. The version compared by Kidd (1987) to other methods was that enhanced by Lee et al (1982).

Chris Chapman, a prominent figure in the UK Operational Research arena, has provided guidelines for the management of risks specific to large-scale projects; first reported in Chapman (1979). Since the development of the procedures in the late 1970s for North Sea oil and gas projects reported in Chapman et al (1983), Chapman has played a key role in the advancement of Project Risk Analysis. A concise report of his experiences on numerous projects and techniques developed are presented in the book covering management material for engineers, Chapman et al (1987).

In light of the evolution of management approaches during the 1980's, a fresh impetus of ideas and relationships has seen the use of quality measures incorporated into many areas of business and management tools developed. In the context of Project Risk Analysis, recent studies incorporating the concept of quality have been made by Chapman's colleagues within the Department of Accountancy and Management Science at the University of Southampton. Klein (1993) discusses the inclusion of duration, cost, and quality as key activity project components. Rather than isolating and modelling each component separately, Klein considers the relationship between the components and discusses the trade-offs that are likely to be performed in practice. This approach supports the ideas of Barnes (1988), who when referring to achieving the required performance of a project indicates that both budgetary limit and target date are important considerations. The VERT simulation technique, compared by Kidd (1987) to other project duration estimation methods, indicates that the characteristics; time, cost, and performance may be independently determined for each project activity and are essential to monitor whether boundaries of time, cost, and performance are met. Similarly Ward et al (1991) exchange the term 'quality' for the 'performance' characteristic and identify time, cost, and quality as primary performance criteria common in construction projects.

Rather than developing the relationship associated with time, cost, and quality, further work performed by Klein includes the discussion of a Project Risk Analysis approach based on the assessment of risk associated with a prototype activity, Klein et al (1994).

Many journals publish first hand experiences and the sharing of project management ideas, for example, *Project Manager Today* and *Construction Management and Economics.* Techniques are presented in Operational Research type journals, including *Management Science*, *Operations Research*, *OMEGA*, *Journal of the Operational Research Society*, and *International Journal of Project Management.*

There are consultancy groups that offer training courses in Risk Engineering and Risk Management. The Decision Support Services company EUROLOG claim to be 'Europe's leading specialist' in Integrated Risk Management. Its experiences are associated mostly with large-scale military projects, however the techniques applied originate from the PERT three time estimate approach. Once again similar to the characteristics identified by Klein, EUROLOG identifies cost, schedule, and resource as the key components for an integrated assessment. In the latter part of 1997 it provided a useful directory of Project Risk Management software on the internet for a special interest group of the Association of Project Management. The information is in zipped format, shown as filename, 'products.zip', and may be accessed from internet site *http://www.eurolog.demon.co.uk/.* Although recent updates have been made the latest version, dated June 1998, still contains incorrect contact information. The unzipped file contains a brief summary for a selection of thirty-one packages, which include the type of risk that may be assessed, the computer platform together with the price and a contact. It should be noted that the information regarding the contacts is out of date, since we wrote to all suppliers and had positive replies from only ten. In addition a fifteen page Project Risk Management bibliography is provided identifying all key texts from various sources including journals, theses, and conference proceedings.

An essential report omitted from the directory is a study titled 'Risk: analysis, perception, management', conducted by a study group of The Royal Society (1992). The report contains many aspects of risk covering issues regarding the quantification, communication, and management of risk. In addition to a review of terminology available for quantifying types of risk the report contains a variety of areas where risk management is applied. Relevant material to project risk type analysis is discussed with the treatment of engineering risks. Here emphasis is on studies to model risks at the design stage of systems that would effect the operational performance. An area of analysis that is briefly mentioned is techniques for assessment of the reliability of a system. Although examples are engineering based and the risk measured is in terms of fatalities, commonly known as 'social risk', two key points were identified applicable to all forms of risk analysis. Since in most cases uncertainty is apparent with regard to the information available it is pointless spending too much effort in the pursuit of accuracy for its own sake. It is noted that care must be taken not to make the treatment artificially rigid, hence losing the realism of the situation.

A collection of work, edited by Ansell and Wharton (1992), which is the outcome of a series of seminars held at Hull University during 1989/90, covers many aspects of risk management. The contribution of a variety of material provides a conceptual framework for a general appreciation of the ethos of risk. Relevant material includes a discussion of the basic concepts and general principles of risk management by Wharton (1992), a further account of the approach employed by Chapman in the planning of North Sea projects based on Chapman et al (1983), the perception of risk by Jackson and Carter (1992), and a discussion by Ansell (1992) of techniques for modelling reliability within the context of an industrial risk assessment.

The difference between risk assessment and risk analysis is summarised by Wharton (1992) who states 'the identification of possible outcomes of decisions is the purpose of risk analysis whilst the estimation of probabilities and size of the outcomes is the subject of risk assessment'. Also the risks that are being considered are perceived risks and are not necessarily actual risks. Arguably only the risks that may be perceived can be modelled, however it should be noted that there may be many actual risks that are overlooked and are therefore not represented in the risk assessment and risk analysis. Part of this problem is addressed by Jackson and Carter (1992) who concludes that risk can only be anticipated through the perception of causality, the problem here lies in the measurement of perception as it is not objective and is influenced by non-epistemological factors. However to progress with the information available, it may be filtered to retain only the relevant details for a decision making process at the possible expense of losing some of the realism.

1.1.1 Risks in the Construction Industry

The construction industry is a major user of project management techniques for scheduling, resource allocation, time management, and cost management. Such techniques are used to provide a means for measuring the progress of a project and aid the decision making progress. Typical concerns of a construction project management team include whether the project will be completed on time, whether the costs will remain within budget and whether the project will be a success.

Recent concerns in the construction industry, specifically quality issues, were discussed at the 1991 European Symposium on Management, Quality and Economics in Housing and other Building Sectors, presented as Proceedings by Bezelga and Brandon (1991). The book contains 189 invited papers, the majority of which are from European contributors, however material from the USA, Asia, Canada, and Australia is also presented.

Management material of interest to the current thesis includes the work by:

Baxendale (1991), who stresses the importance of the integration of time and cost and infers that *cost information cannot be meaningful unless it is related to time.* It is implied that such integration allows assistance to management in controlling projects from a better understanding.

Birrel (1991), considers the factors that govern the efficiency of performing a construction project. An important factor identified was that *the duration of activities is dependent on the quality of resource and the amount of resource used.*

Lastly, Herbsman and Ellis (1991) indicate the drawbacks associated with the low bid system for the procurement of construction contracts. *Problems have arisen where contracts are won solely on cost aspects, including extensive delays in planned schedule and quality problems such as construction failure.* An alternative system is proposed in which the characteristics, *time, cost, and quality* are accounted for in the bidding process. In addition it is suggested, depending on the type of contract, that safety, durability, security and maintenance could be included and that according to the project type weights are assigned according to the importance of each characteristic. *The quality characteristic is identified as the most complicated and difficult to quantify.* In this context the quality is how the contractor will perform the work in question. It is suggested that the contractors are assessed by their performance on previous projects . Apart from the Construction Quality Assessment System

(CONQUAS) used in Singapore to score contractors according to the quality performance of previous jobs, quality aspects are rarely considered in the bidding process. A conclusion made is that further research is necessary in the quantification of quality to assess the performance of a project.

Further relevant papers presented under the subject heading of quality include:

Abiko (1991), in the context of building quality houses, identifies that *the quality of the construction effects cost, durability and maintenance.*

Bezelga and Sousa (1991), present evaluation grids for evaluating the quality of buildings, which essentially involves a checklist for a selection of quality criteria including architectural quality, construction quality and deferred costs and durability.

As identified by Braz-Oliveira (1991), in the installation of aluminium glazing frames, *non-quality* is typically observed as a result of the complexities associated with installation from both technical difficulties and the insufficient integration of management.

In support of this, as presented by Cnudde (1991), it is suggested that the 80/20 ratio established by Deming is applicable to the building sector where 80% of the causes of lack of quality can be attributed to management. Also presented is the distribution of the causes of damage in a selection of five European countries. In Great Britain, it was revealed that 49% of the costs associated with the lack of quality can be attributed directly to the project whereas 29% of the costs are incurred from repair work during the execution. From consideration of the cost associated with the lack of quality Cnudde suggests that approximately 20% of the turnover is associated with insufficient levels of quality. Cnudde also suggests that Juran's definition of Quality, defined as Quality=Fitness for Use, is more flexible than the ISO 8402 standard which defines quality as 'the totality of characteristics' of a construction project 'that bear on its ability to satisfy stated or implied needs'.

Cornick (1991) reveals the areas of incompatibility of the typical quality management tools, presented in the ISO 9001-9004 series, when applied in the construction industry. He stresses the importance of not only managing the quality for project phases independently, but also to consider the interdependence of the various activities. An example provided is, if a project comprises of two sequential phases, namely design and construction, the architects and engineers have a product of 'design' whereas the commercial contractors have a product of 'construction'. It is stated that the quality of each of their products affects the quality of the other's product which in turn affects the quality of the building or civil engineering construction. Cornick provides a simplistic process model for building project management, however does not indicate methods for the quantification of quality or the interdependencies.

Three interesting case studies are presented by Dregner (1991) in which the failure of buildings is attributed to quality deficiencies experienced during design and construction.

Hammarlund and Josephson (1991) discuss the sources of quality failures in Swedish housing and state that 10% of production costs are due to failure costs. They state that approximately 50% of the failure costs can be influenced during the construction phases. In addition they discovered that one third of failure costs are caused by poor management. They conclude that further studies are necessary in order to gain a better understanding of the phenomena that cause failures.

Mathur and McGeorge (1991) state that the majority of techniques in the construction industry relate to the control of time, cost, and performance of individual components or project activities. In the context of providing an integrated decision making environment for cost verses quality control, they recognise that the success of a building project is attributed to low capital cost, high profits, good quality end product, and completion on time. The relationship between cost and quality is implied where cost is regarded as a synonym for quality, thus in some instances better quality is achieved with more money spent. However, in contrast, in some cases spending more money on a project may encourage waste and inefficiency rather than improve quality. In conclusion they state that there is a need for a quality management system where all quality aspects of a project may be quantified, including the quality of components and sub-systems, however to provide such a measure requires a new complex process that is multi-dimensional and multi-disciplinary. Since every building project is unique the process cannot be simplified with the aid of knowledge based systems which are currently inadequate to provide the necessary information on quality aspects.

Lastly, an interesting simplistic quantitative model for the evaluation of construction quality is presented by Nero (1991). Here the global quality, η, of a project is defined on $[0,1]$ as a function of quality, η_i evaluated for each phase and measure of influence, α_i, where $0 \leq \alpha_i \leq 1$. For n sequential phases, global quality is defined as,

$$\eta = \prod_{i=1}^{n} \eta_i^{\alpha_i}$$

and for n phases in parallel,

$$\eta = \sum_{i=1}^{n} \frac{\alpha_i \eta_i}{\sum_{i=1}^{n} \alpha_i}$$

Dependency between phases is recognised but not developed within the context of the model.

1.2 Basic Concepts of a Project Risk Analysis

Although, as justified in our study of relevant construction literature the time, cost, and quality attributes of a project are of concern, it appears that since the inception of project management techniques much effort has mainly been focused in developing methods for determining accurate results for the project completion time or the project cost, however time and cost are rarely modelled together. Research has indicated the need to incorporate other measures, however only non-quantitative approaches have been considered. Klein (1993) adopted non-quantitative approaches as he considered the modelling of numerous relationships between time, cost, and quality, was mathematically complex and in many cases intractable. We begin with a summary of common risk definitions in order to develop suitable project risk measures which will enable the general modelling of the time, cost and quality associated with a project.

1.2.1 Definitions of Risk

'Risk' as defined in the Oxford English Dictionary is 'to expose to the chance of injury or loss', however the origin of the word is unclear. Kedar (1970) suggests risk comes from either the Arabic word *risq* or the Latin word *riscum*. Alternatively Macrimmon and Wehrung (1986) suggest the origin is Italian but is unsure. As pointed out by Wharton (1992) a positive connotation of the use of the modern French *risque*, translated as 'nothing ventured nothing gained', applies best to financial type risks.

The application of risk modelling is apparent in many areas. For example in the context of financial risk, as described by Thomas (1992), areas of financial uncertainty and therefore risk include; insurance, portfolio analysis and option pricing. In a certain type of portfolio analysis it may be necessary for an investor to select a project or group of projects in order to maximise the expected utility. As indicated by Hertz (1964) this may be regarded as minimising the risk of large scale project ventures. For a full background of the development and application of risk refer to the section 'Studying Risk' of the book by MacCrimmon and Wehrung (1986), however it should be noted that Project Risk is not defined. In most cases the word risk implies the chancing of a potential loss of some measure. Thus, in general, there are numerous definitions of risk which are similar in interpretation. For example;

Rowe (1977) defines risk as 'the potential for unwanted negative consequences of an event or an activity'.

Lowrance (1976) on the other hand defines risk as 'a measure of the probability and severity of adverse effects'.

More recently, Rescher (1983) explains that 'risk is the chancing of a negative outcome'.

The definitions of both Rowe (1977) and Rescher (1983) are similar in that a single measure is used to assess the occurrence of an undesirable event. Lowrance (1976), on the other hand, considers not only the probability of the occurrence but also the severity. As described by Gratt (1987), based on statistical expectation, a combined measure referred to as expected risk is given by the product of the probability and the severity.

The risk associated with a project may be quantified by as a probability measure associated with not meeting a specified project target completion time, τ. The risk may be defined as,

$$ProjectRisk = Prob[T > \tau] \tag{1.1}$$

where T is assumed to be a continuous random variable representing project conpletion time with an associated probability density function $f(t)$, and τ is the target time.

Project Risk may equivalently be represented by,

$$ProjectRisk = \int_{\tau}^{\infty} f(t)\, dt \tag{1.2}$$

Alternatively a common measure is the probability of completing the project before the specified target time, as reported in the project management software package *Primavera.*

With additional information regarding the consequences of not completing on time, which may typically be in terms of a liquidated and ascertained damage clause, Smith and Keenan (1979), the Project Risk may be quantified similar to either of the definitions stated by Rowe, Lowrance or Rescher. In the case where the enforcement of the clause is considered an unwanted negative consequence or a negative outcome then Project Risk may be derived from the definitions of Rowe or Rescher. On the other hand, if the penalty associated with the clause is quantified, possibly in terms of the clients loss in revenue, suppose an amount $M(t)$, then Project Risk may be defined in terms of the expected cost from failing to complete on time. In this case,

$$ProjectRisk = \int_{\tau}^{\infty} M(t)\, f(t)\, dt \tag{1.3}$$

A logical extension is to incorporate project costs. In this case, if ignoring the possibility of a penalty type clause, Project Risk may be defined in terms of the chance of not completing the project by time τ and within a cost κ. Thus where the time and cost to complete the project are represented by the joint probability density functions $f(t, c)$, Project Risk may be determined by,

$$ProjectRisk = \int_{\tau}^{\infty} \int_{\kappa}^{\infty} f(t, c)\, dc\, dt \tag{1.4}$$

Where the distributions for time and cost are independent the joint distribution is replaced by the product $f(t)f(c)$. The concept of the risk associated with quality specific to projects has not been developed.

1.2.2 Risk Analysis using a simplified FMECA

An approach sometimes applied in practice to identify and prioritise the risks associated with a project is based on a Failure Mode, Effects and Criticality Analysis, FMECA. The procedures for performing a FMECA are oulined in US MIL-STD-1629, where two basic methods are provided, (Method 101, Method 102). As indicated by Bendell (1998) Project Risk Analysis is typically performed using a method based on both the non-quantitative Method 101 and criticality analysis Method 102. The risks associated with the success of the project are treated as *threats*, where for each threat identified, an incident rate I, a severity rate, S and a detectability rate D are assessed and the project stage noted. The risk priority is quantified as the product ISO and provides a means of identifying the risks of high priority. The versatility of the FMECA approach enables risks specific to given performance criteria to be assessed, such as reliability, availability and output. The use of FMECA in the context of a project, provides a structure for the identification of risks and possible consequences. Such an approach may provide the necessary detail to model the complex dependency relationships within a project.

1.2.3 Risk Analysis of Investment Projects

The technique to quantify the risk of an investment project was originally developed by Hertz (1964) and is discussed in detail in Hertz and Thomas (1983). In essence cash flows may be modelled on input variables where typically cash flows begin negative and hopefully after a period positive cash flows follow. From modelling the cash flow over time, the net present value, NPV of the investment may be determined. The risk of an investment project is thus assessed based on the NPV. The input variables which may include production costs, product price and market share are typically modelled using probability distributions such as the triangular or beta, however to simplify calculations the variables are assumed independent.

1.3 Project Risk Analysis Software

The majority of risk management software involves the use of Monte Carlo simulation where spreadsheets are often used to simulate risk activities. For a summary of common risk analysis software with comments, see Appendix 1.

1.4 Findings from Literature Review and Initial Investigations

Based on the above literature search, we conclude that there is a gap in the literature concerning quantitative models incorporating all three project characteristics time, cost and quality where dependency exists between them.

There have been few advances in developing a Project Risk Analysis method suitable for modelling and assessing the risk associated with project time, project cost, and project quality. We now develop a conceptual framework and exploit attributes common to all projects .

1.5 Characteristics of a Project

In most situations a project is thought of as an arrangement of tasks to achieve some predetermined end. It exists to develop a product, system, or in general it is created in response to a problem. Nearly all projects can be divided into phases to indicate the type of tasks or activities to be conducted in some logical sequence. As indicated by Nicholas (1990) projects may involve plans and undertakings in the areas of Design, Research & Development, Engineering programmes, Construction or Production planning. Although projects are different in nature there are common properties that we may exploit to formulate a simplified model. As previously stated most Project Risk Analysis techniques and indeed most risk software applications solely model either the duration or the cost of a project, based on activity uncertainty specified by a probability distribution. A restricting assumption is that all activities are independent. This is clearly not the case in real projects. A broad definition of a project is given in BS6046. Based on this British standard an initial formulation was presented by Payne et al (1994) which provides a general framework for modelling the risks associated with time, cost, and quality. Four statements presented in BS6046 that we elaborated upon in our initial formulation are summarised below:

No two projects are the same

Projects between different business sectors will obviously be different, but it is also unlikely that any two projects performed in the same business area will be identical. For example a construction firm will experience different constraints for all projects even though the required building is the same. Everything that could affect the outcome of the project has to be considered. This may include the working conditions, the expertise, and what can go wrong. All of these factors will vary for different projects. The identification of what can go wrong is commonly performed using hazard identification techniques, for example, HAZOP, FMEA.

A project is made up of phases

For planning and control of a project, phases are identified which also enable monitoring of progress towards an often predetermined target. The completion of certain phases may be considered as project milestones. Within a phase, activities may be specified which may be in series and/or in parallel.

A project contains uncertainty and risk

Because of the one-off nature of the project and not being able to exactly forecast future occurrences a project will inevitably contain uncertainty.

The defining, developing, and manufacturing phases often affect the subsequent life of the product

We may modify the above statement to apply to all projects rather than those associated with manufacturing. If we replace manufacturing with achieving and substitute project outcome for product we have a common terminology for all project types. Often the project is considered complete when its outcome is achieved and released to a customer of some form. The project outcome is then put into operation. For some project types this is easier to perceive. Consider the construction of a power plant for example, where the both the cost of building the plant and the cost of maintaining are incurred by the owner. Typical management concerns will include many issues including the time and cost to build the plant. Uncertainty is also apparent in other areas. Will the plant succeed in performing as required? What are the maintenance costs? What is the reliability of the plant? Could the plant be built more effectively to reduce maintenance costs? Is there a way of controlling the construction of the plant in order to minimise the risk of the plant being a failure in terms of operation output and maintenance? It is these concerns that we focus upon in the development of a Project Risk Analysis formulation which may be applied to projects where it is crucial to incorporate relationships between project performance and the consequences experienced after the traditional time point of project completion. In support of this are our findings concerning the risks regarding quality within the construction industry. We begin with the concept of a Project Life Cycle.

1.6 Project Life Cycle

Life Cycle modelling as applied in the manufacturing industry, Blanchard (1978) and in the management of reliability is not a new approach. Much of the reliability studies were instigated by the United States where engineering reliability was a key concern in the development of high technology equipment. Studies revealed cost saving benefits from expending much effort and resources early on in the engineering programme to maximise the reliability of the system engineered. This was only achieved by considering the whole life costs of the reliability programme which included, definition, design and development, production and operation. Many of the guidelines are presented as US Military Standards, however two key British standards are Defence Standard 00-40 issued in 1981 and BS5760 which covers the management of reliability. Similar to the concept of life cycles associated within manufacturing industry, we define a Project Life Cycle to include a phase representing the operation of the project outcome in addition to all phases prior to the typical project completion event of handover. To avoid confusion, from now on we shall refer to all phases prior to the operation phase as the project phases.

Based on the characteristics of a project, the risk measures of concern and the concept of a Project Life cycle, we now provide an initial formulation to represent essential performance criteria of projects. The formulation stated here will be the basis of discussion and development throughout the thesis.

1.7 Formulation

Based on the definition of a project given in BS6046 we assume that a project can be divided into serial phases to indicate the types of tasks or activities to be conducted in some logical sequence. A phase may therefore constitute a single activity or represent a collection of activities which may be in series or parallel. In addition, for convenience we assume that the phases are sequential and consist of activities which, when completed, denote a specific milestone or event of the project. For an indication of the overall project performance, characteristic measures of interest need to be quantified for every phase.

In general suppose a project consists of n phases where for each phase we have an indication of the characteristics, time, cost, and quality. For a given phase i let T_i denote the time, C_i the cost, and Q_i the quality. Thus the total time, T, and total cost, C, of the project are given by,

$$C = \sum_{i=1}^{n} C_i \tag{1.5}$$

$$T = \sum_{i=1}^{n} T_i \tag{1.6}$$

However, in order to quantify the quality associated with performance of a project we first need to develop the concept of the measure of quality. At this stage there are many variations on how a model may be developed. Our consideration of possibilities is by no means exhaustive, however we develop a formulation that enables intuitive relationships, that exist in practice but are seldom realistically quantified, to be incorporated. Such relationships are discussed with examples in Chapters 4, 5 and 6.

1.7.1 The Concept of Project Quality

We define Project Quality, PQ, to quantify how well the project is performed. If the entire project life cycle is considered, the operation of the project outcome may be related to efforts expended during the project phases. For convenience we define a Project Quality function to take values in the interval $[0, 1]$, where 0 is defined as no quality and 1 to be perfect. Intuitively, achieved quality changes throughout progress of the project. We assume that the final level of quality is determined by how well the project is performed overall.

If at the start of the project, Project Quality is initially 1, then it is how the project is conducted that will determine the final level of quality. Rather than monitoring how much is done correctly throughout the project we consider the loss in quality that is incurred up to the operation phase. We refer to this measure as the *quality loss.* For consistency with the characteristics time and cost, the measure quality loss is also defined in the range $[0, \infty)$.

1.7.2 Taguchi's Quality Loss Function

The term 'quality loss' is common to the application of Taguchi methods for the control of quality in product. In this context 'quality' is defined by Taguchi as 'the (minimum) loss imparted by the product to society from the time the product is shipped', Taguchi and Wu (1979). In the use of such quality control methods a target is set for a product characteristic and the loss quantified in costs due to deviations from the target value is represented by a simple quadratic loss function. The loss function indicates that a reduction in variability about the target leads to a decrease in loss and a subsequent increase in quality. As well as symmetrical quality loss functions, a half-parabola may be used where a quality characteristic is to be maximised or minimised. Such measures and methods of quality control may be applied in the context of a project, for example in assessing the quality of specific products of the project, however are not suitable for providing a general approach for determining how well a project is performed.

1.7.3 Phase Quality Loss

We thus proceed in developing a general model of quality loss relevant to the context of our project formulation. Unlike Taguchi's symmetric quality loss function, the amount of wrong doing experienced in a given project phase, modelled as the level of quality loss is unlikely to be a symmetric function dependent on a project characteristic, such as time or cost. Our argument for this type of dependency is supported by Mathur and McGeorge (1991) who discussed the possible relationship between quality and cost. In addition we suppose the nature of the dependency is affected by the nature of the project phase.

1.7.4 Project Quality Loss

With quality defined on $[0, 1]$, it is apparent that associated with the cumulation of quality loss on $[0, \infty)$ there will be a reduction in the level of quality of the unit with magnitude in $[0, 1]$. Assuming that for a given project a unique level of quality is associated with a specified level of quality loss, we can thus search for an appropriate transformation from cumulative quality loss to project quality.

Since the transformation is onto the range $[0, 1]$, one set of transformations which are appropriate are the probability integral transformations given by,

$$Q = \int_{ql}^{\infty} h(u)\, du = H(ql) \tag{1.7}$$

where $h(ql)$ is some specified probability density function for the level of quality loss defined on $[0, \infty)$, and $H(ql)$ is the corresponding survivor function. Transformations of this type are monotonically decreasing and map infinite quality loss onto level of quality zero, and zero quality loss onto level of quality unity. Alternatively quality may be related to quality loss by an exponential relationship defined by,

$$Q = e^{-ql} \tag{1.8}$$

This also gives a one to one transformation where $Q = 1$ for $ql = 0$, and $Q = 0$ for $ql = \infty$.

The PQ may be controlled once influencing factors may be determined and relationships specified. To provide a workable structure to our problem type we limit the influencing factors to represent the performance criteria revealed in the literature review. Klein refers to time, cost and quality as *components* of the project risk, however to avoid confusion with mechanical components we refer to the influencing factors as *characteristics*. Therefore in addition to the time and cost characteristics, the quality loss characteristic is incorporated. To include such characteristic measures throughout the thesis we consider their effects on each other.

1.7.5 Relationship between Time, Cost, and Quality Loss

The time, cost, and quality loss experienced within an activity, within a phase, and between activities and phases are not independent. A possible relationship for time and cost was established with the familiar CPM technique where the cost has a perfect negative linear relationship with time. This enabled analysis of the cost of reducing the completion time under normal operation for selected activities, from adding resources, in an attempt to reduce the project completion time. The limitation of such a relationship applied as a CPM approach is that uncertainty is not represented. Time and cost are obviously related, however the nature of the dependency at this stage is undefined. Also the quality of a phase, i.e. how well is the phase performed or how much quality loss is experienced, is related to the amount of time and cost expended.

Due to the very nature of the characteristics it may be argued that each of the characteristics can be affected or affect any of the others in different ways. The extent to which each characteristic affects or is affected by the others is debatable, however at this stage we recognise that two way influences between characteristics may exist as shown in Figure 1.1. See Chapter 4 for the development of a time-cost and time-quality loss relationship and Chapter 6 for further details of characteristic relationships.

Qualitative Dependency Examples

Suppose that for a given phase the level of quality monotonically increases with time spent. Possible effects will include a lower level of quality loss and may include an increase in costs. Alternatively, once again for a given phase, if quality monotonically decreases with time, which is possible if it is perceived that further work contributes negative value, or that degradation occurs when a delay is encountered before continuing with successive phases. Also extreme weather conditions may affect the quality. For example the overrun of a UK based construction project ideally performed in the summer months, may result in an increase in quality loss and possibly costs as the winter months set in. The importance is to understand the driving forces that affect the characteristics. The modelling of such detail provides two distinct advantages in the approach to project management. Firstly, as expected, an understanding of the mechanics of a project is augmented with more relevant information. This in turn will enable control to ensure targets are met for time, cost, and quality criteria set prior to the operation of the project outcome. Secondly, and possibly of the greatest benefit in our formulation, is that the Project Outcome performance may be related to how the project is conducted prior to operation. Such a study not only allows assessment of the risk associated within the scope of the project, but also risks associated with meeting operation criteria.

1.7.6 A Conceptual Model

Based on previous discussion we now present a conceptual model which encaptures the necessary detail for performing a Project Risk Analysis that allows relationships between project performance and operation to be incorporated.

Suppose a project comprises of n sequential phases, where for each phase, i, the time, cost and quality loss experienced are T_i, C_i, and QL_i respectively. If each characteristic is a random variable, then realisations may be denoted by t_i, c_i, and ql_i. For convenience to represent the time, cost, and quality loss experienced in each of the n phases we adopt vector notation, thus $\underline{t}$, $\underline{c}$, and $\underline{ql}$ refer to all the characteristics realised for the entire project.

Immediately succeeding the project phases is the operation phase where the project outcome, PO, is put to use. Measures of interest which represent operation performance criteria include the reliability, output, and availability.

1.7.7 Project Outcome Reliability

Typically reliability is a time dependent probability function denoted by $R(t)$, representing the probability that no failure occurs in the interval $(0, t)$, where in our formulation t is measured from the start of the operation phase. Of equal meaning $R(t)$ can represent the probability of survival past age t, Ascher and Feingold (1984). We shall define the reliability of the PO, in use for time t, denoted by $R(t)$ as the probability that the project outcome is performing as required at that age, and $R(t)$ is taken to be monotonically non-increasing with $R(0) \equiv \alpha$, $R(\infty) \equiv 0$. The level α falls in the range $[0, 1]$ and allows for all possible initial reliability values at project completion. Analogous to component reliability functions we may define similar functions for hazard and related functions. Consider a PO which has not yet failed by time t and let $h(t)$ be the limit of the ratio to Δt of the probability of failure in $(t, t + \Delta t)$. That is,

$$h(t) = \lim_{\Delta t \to 0^+} \frac{prob(t < T \leq t + \Delta t | t < T)}{\Delta t} \tag{1.9}$$

Thus roughly speaking, $h(t)$ gives the probability of immediate failure of the PO known to be in use for time t. The hazard function (1.9) in terms of related functions is,

$$h(t) = \frac{f(t)}{R(t)} \tag{1.10}$$

where $f(t)$ is the probability density function of time to PO failure.

Clearly,

$$R(t) = exp\left[-\int_0^t h(x)\,dx\right] \tag{1.11}$$

and

$$f(t) = h(t)\,exp\left[-\int_0^t h(x)\,dx\right]. \tag{1.12}$$

In addition to considering the dichotomic case where it is assumed the *PO* is either operating as required or is not, as covered in Chapter 5, we may use the partial operation formulation of Bendell and Humble (1985) to model in-between states.

1.7.8 Analysis with Covariates

As identified by Ansell and Phillips (1994), since the formulation of the proportional hazards model as suggested by Cox (1972), much interest has been focused in modelling lifetimes with covariate data. Other models exist, however our objective is to indicate *PO* reliabilty can be modelled in terms of the covariates $\underline{t}$, $\underline{c}$, and $\underline{ql}$. For illustration suppose that the time to *PO* failure is a random variable which has an exponential distribution, with expectation given by a function, $\frac{1}{\lambda(\underline{t},\underline{c},\underline{ql})}$. In this case the p.d.f of the time to failure is given by,

$$f(t) = \lambda(\underline{t}, \underline{c}, \underline{ql})\exp(-\lambda(\underline{t}, \underline{c}, \underline{ql})t)$$

A further example will be considered in Chapter 4 where the proportional hazards model is considered.

Comments

We have provided a general definition of project outcome reliability which provides us with the time to failure of the project outcome. The definition of failure in this application is debatable and project dependent and may result in various alternative risk measures. For example, how well a bridge is constructed will affect the performance of the bridge. In this case particular interest is focused in assessing the reliability of the bridge. In the extreme case, performance failure may result in the collapse of the bridge and loss of life. Projects with possible similar concerns include water barriers, tunnels, and power stations. Other considerations are also important. Since many projects are performed as profit making business ventures, Mathur and McGeorge (1991) the availability and output of the PO are of great importance. For example, availability of the function of a bridge such as the 'Severn Bridge' is essential in order for people to be charged for utilising the service. In addition all lanes must be in use in order to maximise possible output. Availability in this sense may be regarded in terms of a partial operational model where the level of operation directly affects the output. However if we consider the repair of PO failures it is possible to model the PO performance as a renewal process.

1.7.9 Project Outcome Availability

The concept of availability assumes that on failure, repairs are possible. Availability as defined in MIL-STD-721C (1981) is, 'A measure of the degree to which an item is in the operable and committable state at the start of the mission, when the mission is called for at an unknown (random) time'. However as pointed out by Ascher and Feingold (1984) and covered in O'Connor (1991), availability is commonly expressed as the steady state measure, $\frac{MTBF}{MTBF+MTTR}$ where the mean time before failure, $MTBF$ is the value of each of the (IID) uptimes and $MTTR$ is the expected value of each of the (IID) downtimes. In our fomulation we are concerned with the point availability, i.e. the probability that the PO is available for use at time t in operation. Also of interest as indicated by Ansell and Phillips (1994) is the interval availability which is the expected availability during a time period (t_1, t_2) which may be denoted by $Av(t_1, t_2)$. For period $(0, \infty)$, asymptotic availability is given by the steady state measure of above.

1.7.10 Project Outcome Output

In the case where output may be specified as a proportion of total operating capacity, $O(t)$ may be defined on $[0, 1]$ where $O(t) = 1$ denotes maximum possible output at time t, and $O(t) = 0$ indicates no output at time t. Alternatively the actual units of output may be specified at time t. Depending on the nature of the PO and possible 'burn in' periods or initialisation stages the function $O(t)$ may vary between various project types. Suppose the maximum output level possible is denoted by θ. A possible scenario is when output is initially at a maximum, thus $O(0) = \theta$ and the function $O(t)$ is non-increasing. Where a set-up or 'burn in' period is needed an initialisation process may be necessary. This is common in engineering projects where system checks are made to ensure the correct functioning of the product. Examples include the testing of jet engines and power plants. In these cases $O(0) = \alpha$ where $\alpha \geq 0$ and is the initial output level. Various possibilities exist regarding functions for $O(t)$ and the possible relationship between $R(t)$, $A(t)$, and $O(t)$. Depending on output requirements output level θ may not be required and thus reduces the work load of the PO. This in turn may improve the reliability of the PO. On the other hand if $O(t) = \theta$ for all t the reliability may be reduced.

1.8 Summary

In summary we define $R(t)$, $A(t)$, and $O(t)$ as the reliability, availability and output of the Project Outcome for time, t, spent in operation. Furthermore to investigate the relationship of operation performance to project conduct we shall define $\underline{t}$, $\underline{c}$, and $\underline{ql}$ as covariates. Incorporating the covariates we may model operation reliability by $R(t;\underline{t},\underline{c},\underline{ql})$, operation availability by $A(t;\underline{t},\underline{c},\underline{ql})$, and operation output as $O(t;\underline{t},\underline{c},\underline{ql})$.

In addition, assuming a project comprises of n phases, using the quality one-to-one transformation of equation (1.8), the quality of the project outcome is given by,

$$Q_n = \exp^{-\sum_{i=1}^{n} ql_i}$$

where ql_i is the quality loss experienced in phase i.

For an illustration of the conceptual model, see Figure 1.2.

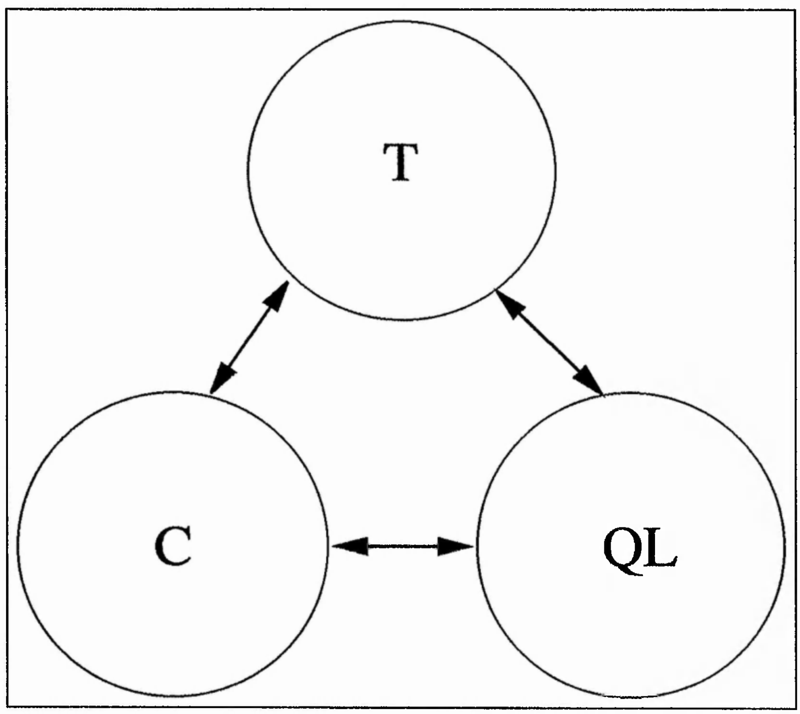


Figure 1.1: Characteristic Dependency

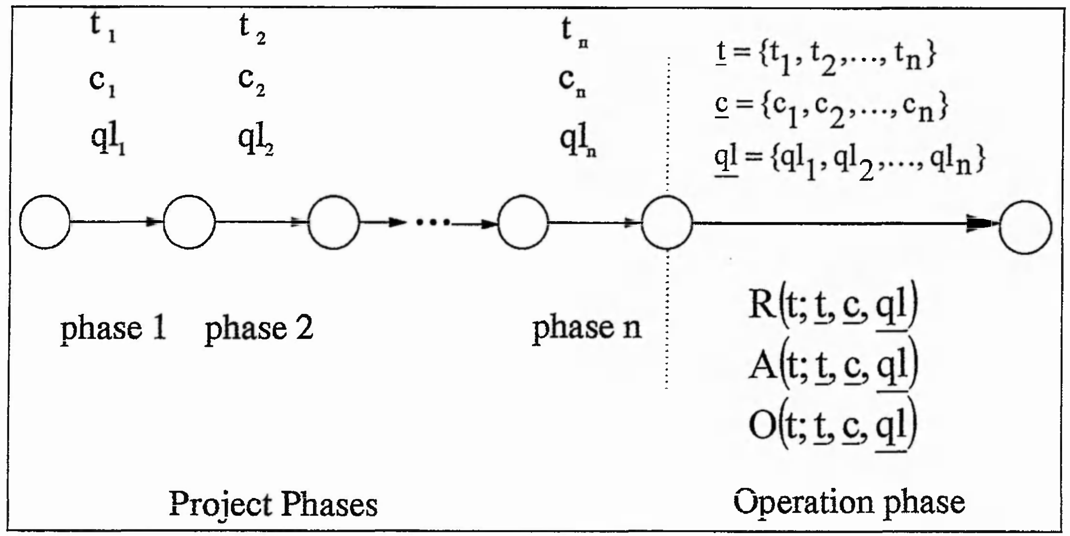


Figure 1.2: Project Formulation

Chapter 2

Methods for Univariate Metrics

Introduction

In order to establish appropriate techniques to facilitate the analysis of our project risk formulation, involving the project characteristics of time, cost, and quality loss, we begin with a review of relevant techniques. Such techniques are commonly applied in the analysis of the project completion time. Since our risk measure requires a probability element we shall concentrate efforts on representing uncertainty as a probability function. In the case of a single characteristic, the use of probability distributions will enable the risk of not meeting the characteristic target to be quantified. The probability operations for dealing with time distributions are established, however different operations may be necessary for the characteristics cost and quality loss. Also of interest are the moments of the distribution discussed by Bendell, Solomon (formerly Jafaar), and Carter (1995) and criticality measures which are commonly reported in the simulation of project networks. For information on criticality measures, see Dodin (1984b), Dodin and Elmaghraby (1985), and Williams (1992).

The two most documented methodologies are PERT and CPM, however it is well known that these methods have limitations. As we are concerned with a probabilistic type measure of risk we shall concentrate efforts on the PERT type approach. The standard PERT analysis requires three time estimates, namely the optimistic, the most likely, and the pessimistic to be specified for each activity within the project. Results produced using the PERT technique may be biased as the possibility of a change in critical path is not accounted for, Fulkerson (1962), MacCrimmon and Ryavec (1964). This has resulted in much research in modifying the PERT method and seeking alternative approaches.

The four main approaches based on a PERT network can be classified as analytical, numerical approximation, use of moments and simulation. An alternative non-probabilistic approach based on the application fuzzy number theory of Zadeh (1965) has also been developed. A brief review of these procedures together with examples is presented and limitations and advantages highlighted. For a summary of Monte Carlo simulation techniques of PERT networks, see Van Slyke (1963), and Burt and Garman (1971).

2.1 PERT

This approach may be adapted when either the characteristics, time, cost and quality are independent or when the Project Outcome, *PO* is dependent on a single characteristic. As there is an extensive amount of literature on the subject of PERT networks only a brief account will be provided here. A good source of information is however provided by Moder et al (1983). Throughout this thesis we adopt the conventions of event labelling, activity numbering, and the use of dummy variables as documented in the majority of Operational Research and Management Science text books, for example, Harper and Lim (1982), Anderson et al (1991).

2.1.1 Project Completion Time

The structure of a project may be represented as a network of independent activities and events. To complete the project all paths shown on the network must be completed. To demonstrate this technique the network configuration of Figure 2.1 will be used where activities are labelled $a_1, \ldots a_5$ and the project starts at event one and event five denotes project completion. It is convention that time flows from left to right and an activity cannot be reached until all activities into it are complete. There are only two paths through the network, a_1, a_3, a_5 and a_2, a_4, a_5. Thus, since simultaneous activities are represented on different paths, it is possible to determine the project completion time by identifying the paths and computing the length of each by summing its activity times. The path with the longest completion time, called the *critical path*, gives the project completion time.

2.1.2 Activity Time

The original PERT technique is described as a stochastic model as it allows uncertainty in the activity completion times. It is assumed that each activity time is randomly distributed with a beta probability distribution of the form,

$$f(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(t-a)^{\alpha-1}(b-t)^{\beta-1}}{(b-a)^{\alpha+\beta-1}} \tag{2.1}$$

where $a < t < b$ and $\alpha,\ \beta > 0$.

By taking the transformation,

$$X = \frac{T-a}{b-a} \tag{2.2}$$

we obtain the standard form of the beta distribution,

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1},\ 0 \le x \le 1. \tag{2.3}$$

The c.d.f is not of a closed form and the integral of distribution (2.3) up to x is commonly known as the *incomplete beta function ratio* and is denoted by $I_x(\alpha, \beta)$ therefore,

$$I_x(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x y^{\alpha-1}(1-y)^{\beta-1}\, dy \tag{2.4}$$

where $B(\alpha, \beta)$ is the *beta function* defined as,

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

2.1.3 Discussion of PERT Estimates

The beta distribution was originally selected due to its modelling properties. It is unimodal, has finite and non-negative end points, and is not necessarily symmetric therefore allowing many duration scenarios to be represented. Determining an activity time distribution from the three time estimates is outlined by Clark et al (1959) and Clark (1961). In brief, for each activity, three subjective values, obtained from expert opinion, are used to paramatise the model: an optimistic time (a), a most likely time (m), and a pessimistic time (b). As noted by Grubbs (1962) the estimates a, m and b are not obtained in the random sampling sense, and since they are determined from expert opinion, they may not have connections with the true population variance of the activity time variance.

It should be noted that the explicit beta probability distribution function of t is not applied in the PERT technique as only estimates for the mean and variance are employed given by equations (2.5) and (2.6) respectively,

$$\hat{\mu_t} = \frac{(a + 4m + b)}{6} \tag{2.5}$$

$$\hat{\sigma_t^2} = \frac{(b-a)^2}{36} \tag{2.6}$$

The estimates of the mean and variance may be transformed using (2.2) giving $\hat{\mu_x}$ and $\hat{\sigma_x^2}$, which using the identities (2.7) and (2.8), may be used to determine the beta distribution parameters, α and β,

$$\mu_x = \frac{\alpha}{\alpha + \beta} \tag{2.7}$$

$$\sigma_x^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{2.8}$$

The justification and derivation of the estimates for μ_t and σ_t^2 has aroused much interest since the inception of PERT in 1958. A summary of key derivations and refinements are presented by Golenko-Ginzburg (1989). In summary the refinements are based on the location of the modal estimate (most likely time). As indicated by Golenko-Ginzburg (1988) when the estimated value of m is located in the tails of the distribution, improved estimates for the mean and variance are,

$$\hat{\mu} = \frac{2a + 9m + 2b}{13} \tag{2.9}$$

$$\hat{\sigma}^2 = \frac{(b-a)^2}{1268}\left[22 + \frac{81(m-a)}{b-a} - 81\left(\frac{m-a}{b-a}\right)^2\right] \tag{2.10}$$

An alternative adjustment is provided by Farnum and Stanton (1987) where if the modal value of the standardised beta distribution, denoted by m_x, is less than 0.13 or greater than 0.87, adjustments are necessary in the estimates for of μ_x and σ_x^2. For $0.13 \leq m_x \leq 0.87$ the unmodified estimates associated with (2.5) and (2.6) may be used. The improved standardised estimates are as follows:

For $m_x < 0.13$,

$$\hat{\hat{\mu}}_x = \frac{2}{2 + \frac{1}{m_x}}$$
$$\hat{\sigma}^2_x = \frac{m_x^2(1 - m_x)}{1 + m_x}$$

and for $m_x > 0.87$,

$$\hat{\hat{\mu}}_x = \frac{1}{3 - 2m_x}$$
$$\hat{\sigma}^2_x = \frac{m_x(1 - m_x)^2}{2 - m_x}$$

In light of the problems associated with the beta estimates both Golenko-Ginzburg (1989) and Berny (1989) developed their own distributions. Others such as Williams (1992) have considered different distribution types, such as the triangular. It is noted that in the case of the Berny distribution,where the c.d.f is given by,

$$Q(x) = 1 - P_M exp\left[(1 - \frac{1}{m})(1 - \frac{x}{x_M})^m\right]$$

where $P_M = \frac{1}{exp(1-\frac{1}{m})}$ and $m > 1$, $x_M > 0$. If we let

$$b = \frac{x_M}{(1 - \frac{1}{m})^{\frac{1}{m}}}$$

and $c = m$ then a Weibull distribution is obtained with shape parameter c and scale parameter b, see Payne (1993).

As discussed by MacCrimmon and Ryavec (1964), the PERT literature is inconsistent as it is stated in the original PERT report, Bureau of Naval Weapons (1958), that 'each activity has a *time*. The time is stochastic and normally distributed ...'. In a later appendix the beta distribution is implied as the time distribution to use. Indeed even though much effort has been focused in illustrating the limitations of the PERT technique it still remains a common project management tool.

With an appreciation of how little the beta distribution is actually applied in the PERT approach we shall now continue with the PERT example by summing the activity mean times for each path of Figure 2.1,

$$\Pi_1 = \mu_1 + \mu_3 + \mu_5$$
$$\Pi_2 = \mu_2 + \mu_4 + \mu_5$$

The critical path is the path with the greatest total mean path time. Thus the estimate of the project mean completion time is the critical path mean time,

$$\mu_{cp} = max\{\Pi_1, \Pi_2\}$$

Similarly the variance of the project completion time, denoted by σ_{cp}^2 is given by summing the variance times of activities on the critical path.

Using the mean and variance of the critical path, μ_{cp}, and variance σ_{cp}^2, an estimate for the probability of completing the project by a target time, t, referred to as *project success* is given by,

$$\begin{aligned} \text{Project Success} &= P(T \leq t) \\ &= P(Z \leq \frac{t - \mu_{cp}}{\sigma_{cp}}) \end{aligned} \qquad (2.11)$$

Alternatively we may define *project risk* as the probability of not completing within a target time t. Thus,

$$\begin{aligned} \text{Project Risk} &= P(T > t) \\ &= P(Z > \frac{t - \mu_{cp}}{\sigma_{cp}}) \end{aligned}$$

Comments

The assumption that the project completion time is normally distributed is questionable. If we assume a sufficiently large number of (IID) activities in the critical path, use of the central limit theorem as discussed in Johnson, Kotz and Balakrishman (1994), implies the completion time is normally distributed. However since activity times are assumed beta distributed, typically with different parameters, it is unlikely that the completion time can be assumed normal and therefore the PERT technique is questionable.

2.1.4 Using PERT to determine Project Cost and Project Quality Loss

In order to incorporate measures for both cost and quality loss we extend the model employed by Dodin and Elmaghraby (1985) for determining the project duration. Let $A\psi_i$ denote the random variable associated with a risk characteristic of activity i and let ψ_j denote the random variable associated with a risk characteristic from the project start to event j where in both cases ψ is a given risk characteristic such that $\psi \epsilon \{T, C, QL\}$. The subscript on ψ is dropped when it is clear we are dealing with measures associated with the entire project. We shall use P to denote the set of paths through the project network, where Π_h is the hth path and $\Pi_h \epsilon P$. We shall use $\psi(\Pi_h)$ to denote either total time, cost or quality loss associated with path Π_h.

The mean and variance of a risk characteristic of a given activity is denoted by $\mu_{A\psi_i}$ and $\sigma_{A\psi_i}$ respectively.

Suppose that the measure of interest is the project cost, which is assumed independent of both time and quality loss. Since all activities are assumed to be performed the cost to complete the project is simply the sum of all the activity costs. Applying the PERT type approach the estimate of the mean project completion cost, μ_c, for a project of n activities is given by,

$$\mu_C = \sum_{i=1}^{n} \mu_{AC_i} \tag{2.12}$$

and an estimate of the project completion cost variance, σ_C^2 is,

$$\sigma_C^2 = \sum_{i=1}^{n} \sigma_{AC_i}^2 \tag{2.13}$$

Estimating statistics for the project quality loss measure requires consideration of how quality loss is affected throughout the activities of a project. So far, in Chapter 1, we have quantified quality loss as a random variable on $[0, \infty)$, however we have not considered exactly how quality loss evolves within the context of a project network. If, once again, it is assumed that all activities within a project network are independent, we may establish rules specific to combining quality loss uncertainties.

In accordance to the PERT type approach we assume that for each activity, quality loss may be represented by an appropriate beta probability distribution and that for each activity i, both μ_{Aql_i} and $\sigma_{Aql_i}^2$ have been estimated. Unlike both the approximation methods for determining time and cost, it may be argued that there are numerous possibilities when calculating the level of total quality loss. If the quality loss measure, or indeed the amount of wrong doing of each activity, will contribute wholly to the overall project quality loss then for a project of n activities the approximation may be determined in similar way to the project cost, thus,

$$\mu_{QL} = \sum_{i=1}^{n} \mu_{AQL_i} \tag{2.14}$$

and

$$\sigma_{QL}^2 = \sum_{i=1}^{n} \sigma_{ql_i}^2 \tag{2.15}$$

If quality loss is considered to accumulate in a similar fashion to time where at any node the maximum level is of concern, then an approximation of the total quality loss may be determined from identification of the critical path associated with the measure of quality loss. It should be noted that if analyses of both time and quality loss are performed the critical paths for each analysis may not necessarily be the same.

Alternatively suppose that quality loss only accumulates by summation for activities in series and that a different rule is required when parallel activities meet. In order to avoid computational difficulties, similar to the PERT approach path, independence may be assumed. Consider the case where there are q paths through the project network. Suppose that the mean and variance of total quality loss experienced for each path, h is denoted by $\mu_{\psi(\Pi_h)}$ and $\sigma^2_{\psi(\Pi_h)}$. A possible approach for approximating the quality loss includes averaging the quality loss contribution from each path. If any of the q paths is considered to contribute greater then weighted averages may be used. In the simplest case,

$$\mu_{QL} = \frac{1}{n}\sum_{q} \mu_{\psi(\Pi_h)} \tag{2.16}$$

and

$$\sigma^2_{QL_j} = \frac{1}{n}\sum_{q} \sigma^2_{\psi(\Pi_h)} \tag{2.17}$$

Example

Suppose that there are k elements of the PO and that each element is independent yet performs a similar task. An applied example may be in the context of providing k specialised power generators in which either the utilisation time is shared or the generators are run in parallel. From our formulation discussed in Chapter 1, it is reasonable to assume that the performance of the PO may be related to how well the project is undertaken to produce the generators. Suppose that each path of the project network corresponds to the production of one of the k generators. It follows that a suitable measure of the quality loss associated with the PO may be estimated from averaging the measures of all paths using (2.16) and (2.17).

Remarks

There has been much criticism in the literature concerning the use of PERT as a suitable method. Doubts were first viewed by Grubbs (1962) who questioned the lack of theoretical justification in the technique. This initiated much research into the accuracy of the results. It has been shown that where there are more than one critical path, or paths near the critical path, errors in the mean as much as 33% of the range may be obtained, Welsh (1965), MacCrimmon and Ryavec (1964), and Lukaszewicz (1965). Even though improvements and modifications have been suggested by Ang et al (1975), Murray (1962), Donaldson (1965), Coon (1965), and Farnum and Stanton (1987) no single method has been shown to generate accurate time completion distributions for all project and activity time possibilities. Biasness of the expected project completion time will always be the main source of inaccuracy when applying the PERT type techniques, however the technique is still widely applied in practice due to its ease of use.

Analysis of the project network where the activity times are beta distributed is generally regarded as too complex, therefore in the literature only approximations are obtained. In many cases since the true distribution of an activity time is unlikely to be known, any distribution that has the properties of unimodality, continuity and allows for skewed representation may be used, such as the triangular distribution as investigated by Williams (1992).

A possible modification to the PERT approach is to consider a different type of activity time distribution which like the beta distribution will offer a rich family of distribution shapes. Experiences from the application of the PERT technique for modelling time have revealed many deficiencies which could also apply in the modelling of project quality loss. The approximation of cost, since all activities are considered independent, involves a simple calculation.

2.2 Analytical Approach

Literature on analytical procedures to determine the project duration distribution, $F(t)$, can be classified into two groups. The first is concerned with computing upper and lower bound estimates of $F(t)$. The second deals with approaches to approximate the completion time distribution.

The procedure of Anklesoria and Drezner (1986) enables path dependency to be modelled using a multivariate normal distribution for activity durations. It uses the expected time to identify the K paths with the largest expected duration. Each of these paths is then considered critical and their completion time probabilities are used as an estimate for the probability of project completion.

Charnes et al (1964) present an analytical approach assuming activity distributions are of the exponential type, each characterised by the same parameter q. Although effective, it is unlikely that all activities in a project network will have identical time distributions.

Dodin (1980) approximates the completion time by discretising the activity distribution functions.

Ringer (1969) extended the algorithm developed by Hartley and Wortham (1966), providing a network reduction technique for developing an integral operator for any PERT network. Hartley and Wortham considered the cumulative distribution of completion time for the following network structures (a) two activities in series, (b) several activities in parallel, (c) five activities arranged in a *Wheatstone Bridge* configuration. Ringer extended on these to include, (d) the *Double Wheatstone Bridge* configuration, and (e) the 'crisscross'.

A continuous Markov chain where activity times are independent exponentially distributed has been developed by Kulkarni and Adlakha (1986) which accurately models project duration. Calculations are extensive and requires the use of a computers to store state space values.

Martin (1965) uses a computational method for evaluating the project time distributions under the assumption that activity density functions are polynomials. This is manageable for small networks, however the number of required polynomial coefficients increases exponentially with the number of network activities.

The approach adopted by Chapman et al (1983) involves the convolution of discrete random variables specified for project activity times. It is indicated that where information is minimal simple calculations may be performed to determine a discrete distribution of the overall project duration. The precision of such a method is limited however increasing the number of activity time classes improves the accuracy but will often require the use of Controlled Interval and Memory computer software (CIM).

In order to appreciate the difficulty in determining the exact distribution of certain project networks where path dependency exists we begin with a summary of analytical operators and demonstrate their use with examples.

2.3 Operators

We define the operators in the context of project time and elaborate on these where necessary for the characteristics of cost and quality loss. We also consider the method specified by Ringer (1969) and Hartley and Wortham (1966) for analysis of a *Wheatstone Bridge* network configuration.

2.3.1 Activities in Series

In general, the time density function of n activities in series involves a convolution of the form,

$$\int_{-\infty}^{\infty} \cdots \int f_{T_1}(u_1) f_{T_2}(u_2 - u_1) \ldots f_{T_n}(x - u_{n-1})\, du_1 \ldots du_{n-1}, \tag{2.18}$$

where $f_{T_i}(x)$ is defined as zero for negative x.

2.3.2 Activities in Parallel

In general the p.d.f. of n activities in parallel is given by,

$$f_z = \sum_{i=1}^{n} \prod_{j \neq i} f_i F_j \tag{2.19}$$

Alternatively the c.d.f. of Z is,

$$F_Z = \prod_{i=1}^{n} F_i \tag{2.20}$$

It should be noted that performing convolutions can be easier for specific probability densities types, such as the normal, however the desirable properties associated with the beta distribution may not be available. For example, it is well recognised that the family of normal densities is closed under convolutions. However the feasibility of normally distributed activity times is questionable since a desirable property is to allow for asymmetric activity time distributions.

The negative exponential is a plausible activity time density where convolutions are elementary, however limitations include the requirement that the optimistic activity time, a, is equal to the most likely activity time, m, and that the coefficient of skewness is constant for all exponential densities.

The gamma distribution is unimodal, continuous, and may be asymmetric and thus is similar in properties attributed to the beta distribution for modelling activity duration. As indicated in Johnson, Kotz and Balakrishnan (1994), the gamma distribution has received special attention in recent literature. Results of interest include the exact distribution of $Y = \sum_{i=1}^{n} X_i$ by Sim (1992) where $X_i \sim \text{gamma}(\alpha_i, \beta_i)$ and α_i and β_i are the shape and scale parameters respectively, where $\alpha, \beta > 0$. An alternative derivation for the same result, based on the inversion of the moment generating function is given by Moschopoulos (1985). A general expression for the property that $f_{\alpha,\beta_i} * f_{\alpha,\beta_j} = f_{\alpha,\beta_i+\beta_j}$, as stated in Feller (1971), is given by Mathai (1982) together with an expression where all α_i's are integer.

Other methods for convoluting random variables include the use of Laplace transforms similar to the use of moment generating functions, however their application in the context of convoluting project activity times is not apparent in the literature. We shall explore the use of Laplace transforms further in Chapters 3 and 4.

2.4 Computational Difficulties

Depending on the network structure computational difficulties may occur in the exact solution of $F(t)$ where path independence is not assumed. This will also apply to the exact solution of $H(ql)$ if the parallel operation defined for the combining of quality loss uncertainties involves taking the maximum or an average.

$$F(t) = Pr(T_N \leq t) = Pr(Z(\pi) \leq t \text{ for all } \pi \in P)$$

where T_N is the completion time of the project, P is the set of all paths and $Z(\pi)$ is defined as the duration of path $\pi \in P$. If the project network can be reduced to a single, equivalent activity starting at node 1 and finishing at node

n, T_N or a similarly defined QL_N can be computed with ease. The reducability procedure is outlined as follows.

2.5 Reducibility of PERT Networks

The two operations available to perform the reduction are convolution (in series) and multiplication (in parallel). These are repeatedly applied until either a single arc is obtained or no further reduction is possible.

Operation 1

Condition: There is at least one node with only one arc directed into it and only one arc emanating from it. Action: The two arcs are in series and are reduced to a single arc by performing a convolution.

Operation 2

Condition: There exist at least two arcs with the same starting and ending nodes. Action: The two arcs are in parallel and may be reduced to a single arc by performing a multiplication of the distribution function.

As indicated by Dodin (1984a) using a reducible network it will take $N-2$ convolutions and $A-N+1$ multiplication operations to obtain a single activity where A and N are the number of activities and nodes respectively. The difficulty arises when there is an interdependency between the paths, i.e. when at least two paths have at least one activity in common. This is best explained with the aid of an example.

2.5.1 Path Dependency

We shall consider the smallest irreducible network (Figure 2.2) which has initiated much research and has been referred to as a 'crossed network' by Hartley and Wortham (1966), a 'Wheatstone Bridge' by Ringer (1969), and a 'forbidden network' by Dodin (1984a). We denote the completion times for each activity by $t_1, \ldots, t_5$ and their respective c.d.f.'s by $F_i(t_i)$. Given the realisations of t_2 and t_5, the three possible paths through the network are,

$$\begin{aligned} \Pi_1 &= t_1 + t_5 \\ \Pi_3 &= t_2 + t_3 + t_5 \end{aligned}$$

and

$$\Pi_4 = t_2 + t_4.$$

Transforming t_i to Π_i for $i = 1, 3, 4$, the inequalities $\Pi_i \leq t$ become,

$$\begin{aligned} t_1 &\leq t - t_5 \\ t_3 &\leq t - t_2 - t_5 \\ t_4 &\leq t - t_2 \end{aligned}$$

Defining $F_i(t_i) = 0$ for $t_i < 0$, it follows that given t_2 and t_5,

$$\begin{aligned} Pr(\max \Pi_i \leq t) &= F(t) \\ &= \int_0^t \int_0^{t-t_5} F_1(t - t_5) F_3(t - t_2 - t_5) F_4(t - t_2)\, dF_2\, dF_5 \end{aligned}$$

Example

Suppose that each of the activities shown in the forbidden network have independent exponentially distributed completion times with common parameter λ.

$$f(t) = \lambda e^{-\lambda t}$$

The exact distribution function of the project completion time is given by,

$$F_x(t) = 1 - e^{-\lambda t}[0.5(\lambda t)^2 + 3\lambda t - 3] - e^{-2\lambda t}[3 + 3\lambda t - 0.5(\lambda t)^2] - e^{-3\lambda t} \quad (2.21)$$

If all paths are considered independent, an approximation to the project completion time is the maximum of all paths through the network and is given by,

$$\begin{aligned} F_a(t) \quad = \quad & 1 - e^{-\lambda t}[3 + 3\lambda t + 0.5(\lambda t)^2] + \\ & e^{-2\lambda t}[3 + 6\lambda t + 4(\lambda t)^2 + (\lambda t)^3] - \\ & e^{-3\lambda t}[1 + 3\lambda t + 3.5(\lambda t)^2 + 2(\lambda t)^3 + \\ & 0.5(\lambda t)^4] \end{aligned}$$

Comparing the above two expressions for $t \geq 0$ and $\lambda > 0$ gives the same conclusion. Table 2.1 shows a sample of the values of the two expressions for $\lambda = 2.0$.

It can be seen that, in this case, the distribution obtained from considering the maximum of all paths is a lower bound approximation to the exact solution for all values of t.

$$F_a(t) \leq F_x(t)$$

Further consideration of the approximation technique will be considered in Chapter 3 where Erlangian activity times are modelled.

2.6 Method of Moments

The method of moments approach avoids the complexity of using the functional form of each activity distribution. Like the analytical approach, convolution and multiplication operators are used to express the sum and maximum of random variables. For simplicity it is assumed that the network paths are independent, therefore enabling a fully reducible network to be obtained. From considering the first four moments of the reduced single activity distribution Pearson curve parameters, $\sqrt{\beta_1}$ and β_2 may be calculated enabling probability points to be read off from Pearson curves, Pearson and Hartley (1976). This enables the determination of $P(T \leq t)$ for given t values. Alternatively if the single reduced activity is assumed to be Gaussian distributed then only the first two moments are required to determine probability points of the normal distribution.

Sculli (1983) proposed an approximation for project completion time mean and variance in which the activity durations are normally and independently distributed. Compared to the PERT calculated approach Sculli's approach provides a closer estimate of the project completion time, however the measure of variation was less accurate. A general method enabling any input distribution to be used has been discussed by Kottas and Lau (1978). Only the first four moments of the input distributions are required to enable the sum or maximum distribution moments to be generated. The moments about the origin, μ' can be used which are related to the central moments as follows,

$$\begin{aligned}
\mu_1 &= \mu \\
\mu_2 &= \mu_2' - \mu^2 \\
\mu_3 &= \mu_3' - 3\mu_2'\mu + 2\mu^3 \\
\mu_4 &= \mu_4' - 4\mu_3'\mu + 6\mu_1^2\mu_2' - 3\mu^4
\end{aligned} \tag{2.22}$$

2.6.1 Activities in Series

The method in general is based on the computation of the first four moments of the distributions. Using the following equations, the first four moments of the single reduced activity time distribution can be generated.

For n activities in series let $W_n = \sum_{i=1}^{n} t_i$ and μ_{ij} denote the ith moment of the jth activity, then,

$$\begin{aligned} \mu_i(W_n) &= \textstyle\sum_{i=1}^{n} \mu_{ij} && \text{for } i = 1, 2, 3 \\ \mu_4(W_n) &= \mu_4(W_{n-1}) + 6\mu_2(t_n)\textstyle\sum_{i=1}^{n-1} \mu_2(W_i) + \mu_4(t_n) && \text{for } n \geq 2 \end{aligned}$$

Example

To illustrate the technique we shall consider two activities in series where the completion time of both activities is of interest. The times to complete the activities are defined as follows,

t_1 has a normal distribution, $N(10, 2)$.

t_2 has a gamma distribution, $r = 11, \lambda = 5$.

The first four central moments for t_1 are given by using the moment generating function,

$$M(t) = e^{\mu t + (\sigma^2 t^2/2)}$$

and the identities of (2.22) giving, $\mu_1(t_1) = 10$, $\mu_2(t_1) = 2$, $\mu_3(t_1) = 0$ and $\mu_4(t_1) = 12$.

Similarly the first four central moments of t_2 are given by the moment generating function,

$$M(t) = \left(1 - \frac{t}{\lambda}\right)^{-r}$$

giving, $\mu_1(t_2) = 0.2$, $\mu_2(t_2) = 0.44$, $\mu_3(t_2) = 0.176$ and $\mu_4(t_2) = 0.6864$.

Therefore the first four central moments of the combined distribution of $w = t_1 + t_2$ are, $\mu_1(w) = 12.2$, $\mu_2(w) = 2.44$, $\mu_3(w) = 0.176$, and $\mu_4(w) = 17.9664$.

The Pearson curve parameters are calculated as follows,

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} = 0.04615$$
$$\beta_2 = \frac{\mu_4}{\sigma^4} = 3.0177$$

The distribution of the above example is close to a normal distribution which has $\beta_1 = 0$ and $\beta_2 = 3$

Pearson Curve Approach

The use of non-normal frequency curves was originally devised to graduate observational data, however a common application has been in approximating sampling distributions where moments are known but whose p.d.f's are unknown or if the explicit distribution is difficult to handle. As outlined and tabulated in Pearson and Hartley (1976), if the variate of interest is standardised as $X = \frac{x-\mu}{\sigma}$, percentage points are provided for distributions specified in terms of the two moments ratios, $\sqrt{\beta_1}$ and, β_2. For given values of the moment ratios, six lower and six upper standardised percentage points are presented. A limitation in using such tables is in the accuracy achieved in determining the tail probability for given standardised values. For example for $\sqrt{\beta_1} = 0.04615$ and $\beta_2 = 3.0177$ the probability of completing both activities within 14 units of time, as identified from Pearson Curves, is between 0.75 and 0.90. This is not as accurate as required, however interpolation is possible and, as indicated by Bendell et al (1995), computer software is available.

Standardisation Approach

Assuming that the combined distribution is of the normal type then the following standardisation can be used,

$$P(W \leq t) = P(Z \leq \left(\frac{t - \mu_1}{\sqrt{\mu_2}}\right))$$

Using the previous example,

$$\begin{aligned} P(w \leq 14) &= P(Z \leq 1.152) \\ &= 0.875 \end{aligned}$$

2.6.2 Activities in Parallel

There are two problems associated with computing the moments when activities are performed in parallel. The first is that the single reduced activity distribution is often not of the same distribution type as the input distributions. This may not be of concern if the single reduced activity distribution is not required to be in functional format. The second is the difficulty in evaluating the moments integral,

$$\mu'_r = \int_0^\infty (\sum_{i=1}^{n} t^r f_i \prod_{j \neq i} F_j) dt$$

Example

Suppose that we have two activities, with p.d.f's f_1 and f_2, in parallel where the completion time of each is exponentially distributed. Let $f_1 = \frac{1}{3}e^{-\frac{t}{3}}$ and $f_2 = \frac{1}{5}e^{-\frac{t}{5}}$ with respective c.d.f's $F_1 = 1 - e^{-\frac{t}{3}}$ and $F_2 = 1 - e^{-\frac{t}{5}}$. The moments about the origin are given by,

$$\mu'_r = \int_0^\infty (t^r f_1 F_2 + t^r f_2 F_1) dt$$

Thus, $\mu_1' = \frac{49}{8}$, $\mu_2' = \frac{1951}{32}$, $\mu_3' = \frac{223347}{256}$, and $\mu_4' = \frac{8523453}{512}$. Using the moments relationships of (2.22), gives $\mu_1 = 6.125$, $\mu_2 = 23.453$, $\mu_3 = 211.715$, and $\mu_4 = 4773.78$.

If the single reduced activity is approximated by an exponential distribution the first moment may be matched to the distribution parameter λ using $\lambda = \frac{1}{\mu_1}$. Thus an approximation denoted by, $f_a(t)$, to the distribution is,

$$f_a(t) = 0.163265\, e^{-0.163265t} \tag{2.23}$$

We may compare this to the exact distribution function of,

$$f_e(t) = \frac{5e^{\frac{t}{5}} + 3e^{\frac{t}{3}} - 8}{15e^{\frac{8t}{15}}} \tag{2.24}$$

A plot of both functions is presented in Figure 2.3. Other than a visual comparison which indicates a poor fit we may perform a goodness of fit test. The error involved in approximating the maximum of two exponentials by an exponential distribution will depend upon the exponential parameters, however we define the error measure as,

$$D = \sup_t |S(t) - F_0(t)| \tag{2.25}$$

where, $S(\cdot)$ is the approximating cumulative exponential distribution and $F_0(\cdot)$ is the exact cumulative distribution.

Similar to the approach by Sculli and Wong (1984), who considered the approximation of the maximum of two beta variables by a further beta variable, we treat the values of the actual cumulative function as random values. In this sense we may perform a Kolmogorov-Smirnov goodness of fit. As we are not applying the technique with sample values, we consider a number of classes of equal width in the interval in which $F_0(\cdot)$ is between 0 and 1. In all of the cases considered by Sculli and Wong (1984) forty observations were taken. For convenience we also use this number of observations. In order to compute the test statistics we first state the corresponding c.d.f.'s of equations (2.23) and (2.24). Thus,

$$S(t) = 1 - e^{-0.163265t} \tag{2.26}$$

and

$$F_0(t) = 1 + e^{-\frac{8t}{15}} - e^{-\frac{t}{3}} - e^{-\frac{t}{5}} \tag{2.27}$$

From statistical tables, for example see Murdoch and Barnes (1986), we see that for $n > 35$, critical values for significance level $\alpha = 20$, $\alpha = 15$, $\alpha = 10$, and $\alpha = 5$ may be approximated using, $\frac{1.07}{\sqrt{n}}$, $\frac{1.14}{\sqrt{n}}$, $\frac{1.22}{\sqrt{n}}$, and, $\frac{1.36}{\sqrt{n}}$ respectively. We shall regard D values greater than the critical value as justification for rejecting the suitability of the fit. By assigning forty classes each of width 1, the test statistic is found to be 0.118161. For $\alpha = 20$ the critical value is 0.17, therefore we may suppose that the approximation is suitable in this example. For a stronger belief supporting the goodness of fit we suggest using a higher significance level.

The exact density function from performing the maximum of exponential densities has a modal value greater than zero which is a property that cannot be achieved by an exponential density unless thresholds are incorporated. However, based on the Kolmogorov goodness of fit, it appears that an exponential may be a suitable approximation to the maximum of exponential densities.

In light of these results we suggest that the approximation should be performed with caution. We also realise that the exponential density is a density commonly applied in modelling times as it's use often simplifies the mathematical content of a model. Examples include applications in reliability studies where time to failure of a component is assumed to follow an exponential distribution and in queueing theory where service times are commonly assumed to be exponential.

The approximation method for beta variables, as suggested by Sculli and Wong (1984), can provide suitable approximations, although it is more involved since determining the c.d.f of beta variables requires the use of numerical integration. In agreement with Sculli and Wong we suggest that if an approximation is to be used, each case must be judged on its own merits.

2.6.3 Moments Method using Erlangian Variables

The moments method recently applied by Bendell et al (1995) assumes Erlang distributed activity times. Results of interest include moments formulae for two parallel activities and the application of the moments method to bimodal activity time distributions. We state the moments formulae specific to the maximum of two Erlang densities and shall apply them in Chapter 3 to compare results with our Laplace Transform approach.

Let t_1 and t_2 be independent Erlang distributed times with means μ_1 and μ_2 respectively and shape parameters c_1 and c_2. If p and q are defined as follows,

$$p = \frac{c_1\mu_2}{(c_1\mu_2 + c_2\mu_1)} \tag{2.28}$$

and

$$q = \frac{c_2\mu_1}{(c_1 + \mu_2 + c_2\mu_1)} = 1 - p \tag{2.29}$$

the first four central moments of the distribution of $max(t_1, t_2)$ are,

$$
\begin{aligned}
\mu_{jM} &= \mu'_{j,1}\left[1 - p^{c_1+j}\sum_{i=0}^{c_2-1} C_i^{c1+i+j-1} q^i\right] \\
&+\mu'_{j,2}\left[1 - q^{c_2+j}\sum_{i=0}^{c_1-1} C_i^{c_2+i+j-1} p^i\right] - A_j \text{ for } j\text{=1, 2, 3, 4} \quad (2.30)
\end{aligned}
$$

where,

$$\mu'_{j,1} = \mu_1^j \prod_{i=0}^{j-1} \frac{c_1 + i}{c_1}$$

and

$$\mu'_{j,2} = \mu_2^j \prod_{i=0}^{j-1} \frac{c_2 + i}{c_2}$$

give the jth moment about the origin of each respective distribution. In addition A_j for j=1, 2, 3, 4 is defined as follows:

$$A_1 = 0 \quad (2.31)$$

$$A_2 = \mu_{1M}^2 \quad (2.32)$$

$$A_3 = 3\mu_{1M}\mu_{2M} + \mu_{1M}^3 \quad (2.33)$$

$$A_4 = 4\mu_{1M}\mu_{3M} + 6\mu_{1M}^2\mu_{2M} + \mu_{1M}^4. \quad (2.34)$$

2.7 Characteristic Uncertainty Modelling using Fuzzy Numbers

Following the increase in interest in the application of fuzzy set theory created by Zadeh (1965) in modelling uncertainty, techniques have been derived for assessing project completion time for project activity networks. In this application a special class of fuzzy sets are used that are defined on the set $\Re$ of real numbers and are commonly referred to as *fuzzy numbers*.

Application of fuzzy numbers analogous to the PERT technique have been presented by Buckley (1989), Chanas and Kamburowski (1981), Dubois and

Prade (1988), Mares (1989), Mares and Horak (1983). In brief, the justification presented by the above authors for using a fuzzy set to model the uncertainty is due to the following circumstances:

- The subjective nature of experts' opinions under which the activity time is determined. For new, unique projects experts are used to evaluate the project actions. Their opinions may be best reflected in a possibility distribution obtained from a fuzzy evaluation of activity duration time.
- Lack of repeatability in activity and project realisations causes the notion of probability and expectation to be meaningless in the sense of classical probability theory.
- Computational difficulties using probabilistic methods.

A direct analogy to stochastic PERT where activity times are continuous is to consider fuzzy defined activity times where the membership function is also continuous. Similarly where activity times are discrete random variables, this corresponds to using a discrete fuzzy set.

We found a variety of methods using fuzzy numbers for performing a PERT analysis however most methods differed in definitions and use of fuzzy operators. To compare with the common probabilistic approach we shall briefly cover the essence of fuzzy sets and relevant operations which may be applied in the context of a project network.

2.7.1 Fuzzy Sets

There are two basic types of fuzzy sets, *ordinary fuzzy sets* and *interval fuzzy sets*. Both allow the modelling of uncertainty and in particular enable the representation of vague concepts of the linguistic nature such as judgements and opinions. Both types of fuzzy sets are uniquely defined by a membership function. In general a membership function assigns values within a given range to the elements of a specified universal set such that the values indicate the membership grade of the elements. It is typical that values of the membership functions lie in the unit interval $[0,1]$. A universal set denoted by X is analogous to the domain of a probability function, where each element of X may be mapped by a membership function into real numbers in $[0,1]$. For example the membership function of a fuzzy set A is denoted by,

$$A : X \rightarrow [0,1] \tag{2.35}$$

It should be noted that modelling uncertainty is not a straightforward task, which is a common problem experienced in all areas of uncertainty modelling. Selecting a suitable parametric function to accurately model the uncertainty of a particular event is subjective as often the exact nature of the uncertainty is not known. Fuzzy numbers in many cases are often precisely modelled by simple functions, such as triangular, trapezoidal with a fuzzy interval, or symmetric bell shaped functions which are similar if not the same to those applied in probability modelling. Fuzzy sets of this nature are referred to as the *ordinary* type.

The *interval* type of fuzzy sets are not so precisely defined. Suppose that for each element of the universal set a single value cannot be assigned, however instead a lower bound and upper bound of the membership value may be determined. In this case the membership function is given by,

$$A : X \rightarrow \varepsilon([0,1]) \tag{2.36}$$

where $\varepsilon([0,1])$ is the set of all closed intervals of real numbers in the unit interval. This is computationally more demanding as additional information to explicitly define two functions to represent the upper and lower bounds may not be readily available. We shall therefore not pursue this type of fuzzy set as accurate information is often limited and continue our discussion of fuzzy numbers.

To enable network analysis the operations of interest include *addition* and *maximum* of fuzzy numbers. These may be performed with an understanding of α-cuts. Given a fuzzy number A defined on X then for any number $\alpha\epsilon[0,1]$ the α-cut is defined as,

$$^{\alpha}A = \{x|A(x) \geq \alpha\} \tag{2.37}$$

A variation of this is the strong α-cut,

$$^{\alpha+}A = \{x|A(x) > \alpha\} \tag{2.38}$$

This enables the identification of elements of X whose membership values in A exceed a specified level of value of α. Thus we may identify specific intervals of belief by decomposition of the fuzzy number and perform arithmetic operations on these intervals. Interval analysis in the usual mathematical sense involves arithmetic operations on closed intervals. A result of interest is that for closed intervals $[a,b]$ and $[c,d]$,

$$[a,b] + [c,d] = [a+c, b+d] \tag{2.39}$$

Suppose A and B are fuzzy numbers, both modelled by continuous membership functions then $A+B$ is defined as,

$$A + B = \bigcup_{\alpha\epsilon[0,1]} {}_{\alpha}(A+B) \tag{2.40}$$

where ${}_{\alpha}A(x) = \alpha \cdot {}^{\alpha}A(x)$

We shall use (2.40) to illustrate the addition of two fuzzy numbers. Suppose,

$$A(x) = \begin{cases} 0 & \text{for } x \leq 1 \text{ and } x > 5 \\ \frac{x-1}{2} & \text{for } 1 < x \leq 3 \\ \frac{5-x}{2} & \text{for } 3 < x \leq 5 \end{cases}$$

$$B(x) = \begin{cases} 0 & \text{for } x \leq 2 \text{ and } x > 5 \\ x-2 & \text{for } 2 < x \leq 3 \\ \frac{5-x}{2} & \text{for } 3 < x \leq 5 \end{cases}$$

The α-cuts are,

$$^{\alpha}A = [2\alpha + 1, 5 - 2\alpha] \tag{2.41}$$

and

$$^{\alpha}B = [\alpha + 2, 5 - 2\alpha] \tag{2.42}$$

Applying (2.39) to the α-cut intervals of fuzzy numbers gives,

$$^{\alpha}(A + B) = [3\alpha + 3, 10 - 4\alpha] \text{ for } \alpha\epsilon(0, 1] \tag{2.43}$$

The resulting fuzzy number is thus given by,

$$(A + B)(x) = \begin{cases} 0 & \text{for } x \leq 3 \text{ and } x > 10 \\ \frac{x-3}{3} & \text{for } 3 < x \leq 6 \\ \frac{10-x}{4} & \text{for } 6 < x \leq 10 \end{cases} \tag{2.44}$$

indicating an asymmetric distribution in the belief of the duration of both activities. In this example we conclude that the strongest belief is that both activities will take six time units, however it is also believed that both activities will take between three and ten time units. For increasing levels of α, the interval width associated with the α-cut decreases monotonically. Thus for $\alpha = 0$, the α-cut specifies the entire range of possible completion times, for $\alpha = 1$ the time with the strongest belief is given and for $0 < \alpha < 1$ the time interval values are given for which level α is exceeded.

The second approach of interest is the maximum of two fuzzy numbers. This is analogous to establishing the maximum of two p.d.f.'s. The maximum of fuzzy numbers is rather a nebulous concept as we do not have the rigour of probability techniques and therefore the maximum operation may change depending on the context of the application.

The maximum operator introduced by Chanas and Kamburowski (1981) in the context of PERT networks was stated without explanation and may cause a misunderstanding of use. To facilitate the use of the maximum operator we first state the lattice operations for *min* and *max* on real numbers which are linearly ordered. Thus,

$$min(x, y) = \begin{cases} x & \text{if } x \leq y \\ y & \text{if } y \leq x \end{cases}$$

$$max(x, y) = \begin{cases} y & \text{if } x \leq y \\ x & \text{if } y \leq x \end{cases}$$

where $x, y \epsilon \Re$.

These lattice operations are easily extended to corresponding lattice operations on fuzzy numbers. Thus as defined by Chanas and Kamburowski (1981) and explained by Klir and Yuan (1995), the maximum operator is given by

$$MAX(A,B)(z) = \sup_{z=max(x,y)} min[A(x), B(y)] \quad (2.45)$$

For example, for $A(x)$ and $B(x)$ previously defined, since $B(x)$ is a subset of $A(x)$, the maximum, $MAX(A,B)(x)$, is given by the function defined for $B(x)$.

Comments

Based on available literature, we have described two essential operators required for performing a typical network analysis where for each activity a membership function is identified which best represents the believed time to complete the activity. A point of concern is how a risk measure may be established from activity times quantified by a fuzzy number with this approach. A possibility may be to consider the upper interval value of an α cut, however in this context, interpretation of such a measure is unclear in terms of typical risk measures. Alternatively, if instead, for each activity we quantify a membership function such that the belief that the activity is completed by time t is captured, then we may define a suitable fuzzy risk value measure which is analogous to the probabilistic risk measure. For risk level α, where $\alpha\epsilon[0,1]$ we may define the fuzzy risk value, FRV_α in terms of the α-cut operation. Suppose that $Z(t)$ is a fuzzy number denoting the completion time of a project by time t then, $FRV_\alpha =^{(1-\alpha)} Z(t)$.

In summary although we believe that fuzzy numbers may be used to quantify the uncertainty and risks associated with project characteristics, no real advantage over the use of applying probability theory is obvious. In addition since the inception of PERT, excluding CPM techniques, the majority of project time analyses have employed probabilistic methods which based on reported techniques and current available risk software, may be regarded as the standard approach. In light of this we shall not pursue the use of fuzzy numbers any further in our discussion.

2.8 A Parametric Quality Loss Model

As outlined in Chapter 1, we stated the existence of a relationship between how well a project is performed and the effect it has on Project Outcome performance measures, typically, reliability, availability, and output. At this point, we demonstrate not only the effect of the amount of wrong doing, but introduce a model that indicates how the level of wrong doing accumulates throughout the project phases and tackle important issues arising in the control of such wrong doings and the implications on cost. As previously discussed we denote the level of wrong doing by ql.

2.8.1 The Cumulation of Quality Loss

Consider a project of n sequential phases and suppose that no preventative or corrective measures are in place to ensure the project requirements are met. Due to the dependency of project activities and thus the outcomes of project phases, if an earlier phase is performed badly this may contribute to further levels of wrong doing in subsequent phases. In support of the qualitative dependency of phase quality identified by Cornick (1991), we shall build upon the model developed by Nero (1991).

We thus assume that a growth in quality loss will be experienced such that $ql_j \geq ql_i$ for all $j > i$. The nature of the growth in quality loss is debatable, however to enable quantification we shall assume similar to the defect growth model of Sears (1991) that the level of quality loss experienced in phase i, is a multiple, N_i of the level of quality loss present at the completion of phase $i - 1$.

Thus if no corrective measures are in place and assuming an inherent level of quality loss, ql_0, the total quality loss, QL of a project of n phases is given by,

$$QL_n = ql_0 \prod_{i=1}^{n} N_i$$

where $N_i \geq 1$.

Suppose that effort is employed in preventative and corrective measures for each phase and that from such efforts the proportion of quality loss remaining at the end of phase i is given by x_i, we may define the quality loss level at the completion of phase j as,

$$QL_j = ql_0 \prod_{i=1}^{j} N_i x_i \tag{2.46}$$

where $0 < x_i \leq 1$. It follows that for phase i the level of quality loss removed is $(1 - x_i)$ and the level of quality loss remaining is x_i.

2.8.2 The Cost of Corrective and Preventive Measures

The question of when and indeed where and how should quality control measures ideally be undertaken has been addressed in the reliability field as outlined by O'Connor (1991) where it is acknowledged that the discovery and correction of possible failure modes early on in a development programme is much more cost effective than bearing the cost of an unreliable product. We adopt a similar belief in the context of performing a project in that it is cheaper to rectify levels in quality loss during earlier project phases. We introduce a competing cost coefficient, c_o, to represent the cost to eliminate one unit of quality loss. As the project progresses the cost to eliminate a single unit of quality loss increases by a factor of M_i at each phase i.

A suitable function for the competing costs incurred in phase i, which incorporates the level of quality loss of (2.46), is given by,

$$CC_i = ql_0c_0(1 - x_i)\prod_{j=1}^{i} M_jN_jx_{j-1} \tag{2.47}$$

where $x_0 = 1$.

A refinement of the competing cost function is possible if we consider the reasoning behind the time-cost slope employed in the CPM approach. As documented in most standard Operational Research text books a reduction in the time to complete an activity requires additional resource and thus an increase in costs. For each activity a constraint, commonly referred to as the *crash time*, is in place representing the minimum amount of time an activity time can realistically be reduced to. In our context, instead of time we are dealing with quality loss, and shall thus assume similar to the crash constraint that a threshold exists for which reducing the quality loss level beyond this point will require extensive resources and hence will be reflected in increased costs. We may incorporate this feature by introducing a threshold parameter, t_i for each project phase i. A possible linear cost-quality loss relationship is,

$$cost = \frac{c_i(1 - x_i)}{1 - t_i}$$

where c_i is a cost estimate for activity i under maximum crashing and $x_i \geq t_i$. A possible non-linear cost-quality loss relationship is given by,

$$cost = \frac{1 - x_i}{x_i - t_i}$$

where a necessary condition is that $x_i > t_i$.

It follows that the total cost of competing efforts, TCC_n, with non-linear cost-quality loss relationships, of a project of n phases as,

$$TCC_n = ql_0c_0 \sum_{i=1}^{n} \frac{(1-x_i)^2}{x_i - t_i} \prod_{j=1}^{i} M_j N_j x_{j-1} \tag{2.48}$$

where as previously stated $x_0 = 1$.

If we suppose that the cost of unreliability due to failure, FC is a function of the level of quality loss experienced, the optimum strategy in competing against levels of quality loss in order to minimise overall costs may be identified. Thus if we define $FC = \alpha TQL$, then the total cost, TC, associated with competing efforts and unreliability costs is given by, $TC = TCC + FC$

2.8.3 Cost Optimisation

The structure of the formulation is in a format where we may identify the optimum competing cost efforts for each of the project phases. This may be achieved by identifying values of x_i such that an incremental change in competing efforts can be made at equal costs at any phase within the project. We thus consider the derivative of TC with respect to the contribution of costs from each phase i. Thus the optimum cost is achieved when,

$$\frac{dTC}{dx_1} = \frac{dTC}{dx_2} = \cdots = \frac{dTC}{dx_n} \tag{2.49}$$

We may perform all the necessary differentiation's and identify values of x_i which satisfy the optimum conditions. The differential equations may be easily solved using a software package such as *Mathematica*.

2.8.4 Example of Minimising Project Unreliability Costs

We consider the scenario indicated by Cornick (1991) in which the quality achieved in design affects the quality obtainable in construction. Assuming that the design phase and construction phase are sequential, we suppose that the inherent level of quality loss, ql_0 is one. This value may be a weighted value representing typical problems attributed to the project. Suppose that for each phase the level of quality loss doubles, thus $N_i = 2$ for $i = 1, 2$. Suppose that the cost to eliminate one unit of quality loss is ten units, and to represent the increase in costs experienced to rectify problems at later stages we shall assume that $M_i = 2$ for $i = 1, 2$. In addition suppose that in both phases all quality loss may be removed however to reflect the possibility that fewer problems are harder to find and correct, cost of rectification increases as quality loss decreases. Performing the necessary differentiation's and solving reveals that the optimum strategy to minimise the costs associated with unreliability is to spend an amount of 195 units in the design phase and an amount of 90 units in the construction phase. With competing measures in place the total cost with the entire program is 385 units. If no competing efforts were performed during the design phase, the program costs would be 1312 units, and if no competing measures were in place for either phase the program costs would be 4000 units. This example, even though oversimplified and requiring estimates for escalation parameters, illustrates the importance of addressing areas of quality loss earlier on in a project in order to minimise the costs associated with unreliability.

Further Measures

The above model highlights the savings of investing more effort into reducing quality loss at earlier stages in the project. It offers flexibility in setting project stage growth values for quality loss and cost depending on the project type. Other measures are readily available if we identify additional concepts. If for example, the failure rate associated with the *PO* performance is determined at project handover and denoted by λ, we may suppose that level of ql affects the wear out and degradation of the PO and thus that λ is a monotonically increasing function of ql. In support of this argument, in the context of building failure is the work by Dregner (1991). In this sense we are able to demonstrate reliability growth by reducing the level of wrong doing, thus implying that the interarrival times of failures of the PO become larger for decreasing levels of ql.

The fundamental difference between our scenario in analysing the effect quality loss has on reliability throughout the project and typical reliability studies in reliability growth is the non-existence of data. However our formulation is analogous to the considerations of the 'Initial Defects Model' of Cozzolino (1968) where the total level of wrong doing is referred to as the presence of errors committed during the production process and the possibility of unintended structural weaknesses. Unlike our formulation the wrong doings are quantified as defects where it was assumed that each defect has an exponential density time to failure with identical parameter, λ.

In conclusion, examination of the effects of quality loss on reliability is possible with modification of the parametric model defined above, however the limitations include the inability to represent the uncertainty associated with quality loss in addition to representing the characteristic measures time and cost.

2.9 Comments

After reviewing all the methods which may be adopted to perform a Project Risk Analysis we feel that progress in this area has been minimal, as both accuracy and meaningful results are restricted by algebraic and computational difficulties. In light of this we shall first indicate possible areas of improvement using Laplace transforms. The algebraic and computational difficulties may be overcome by utilising mathematical computer software, such as *Mathematica*, Wolfram (1988). However in order to develop a risk measure as defined in Chapter 1, rather than extending an existing methodology we shall utilise our sequential project formulation of Chapter 1 where possible. Interpretation of a fuzzy graph is unclear and therefore as the concept of 'probability' is generally recognised as an acceptable measure of risk we shall not pursue the development of a fuzzy risk measure.

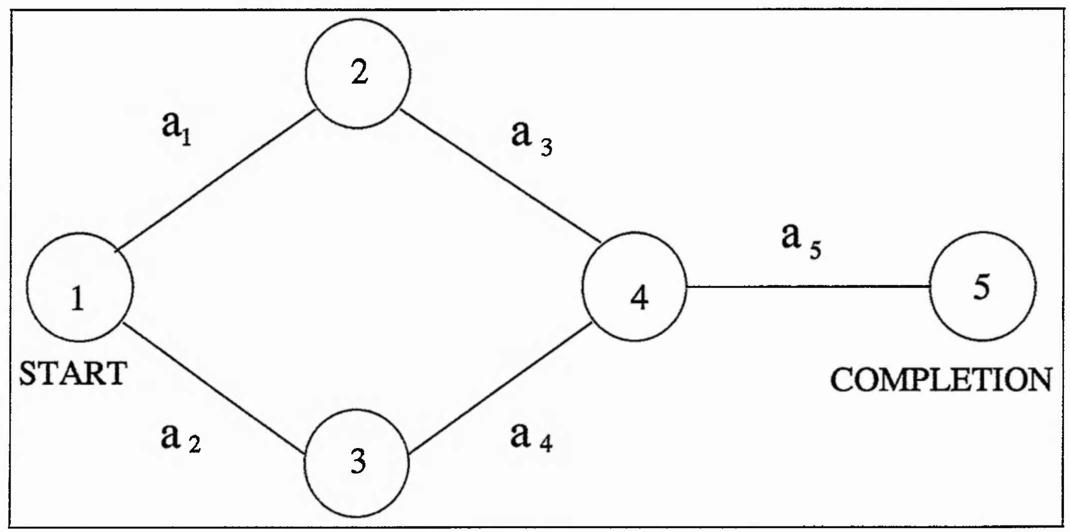


Figure 2.1: Two Path Network

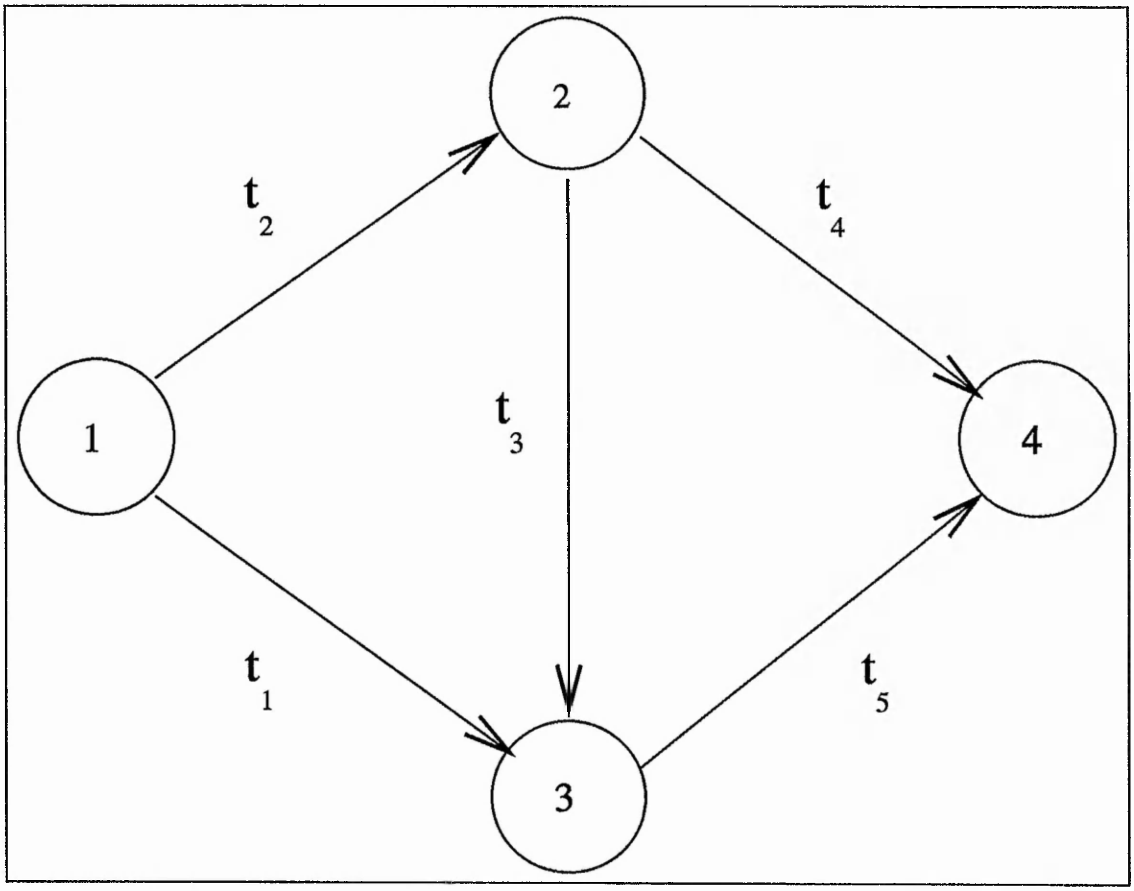


Figure 2.2: The Forbidden Network

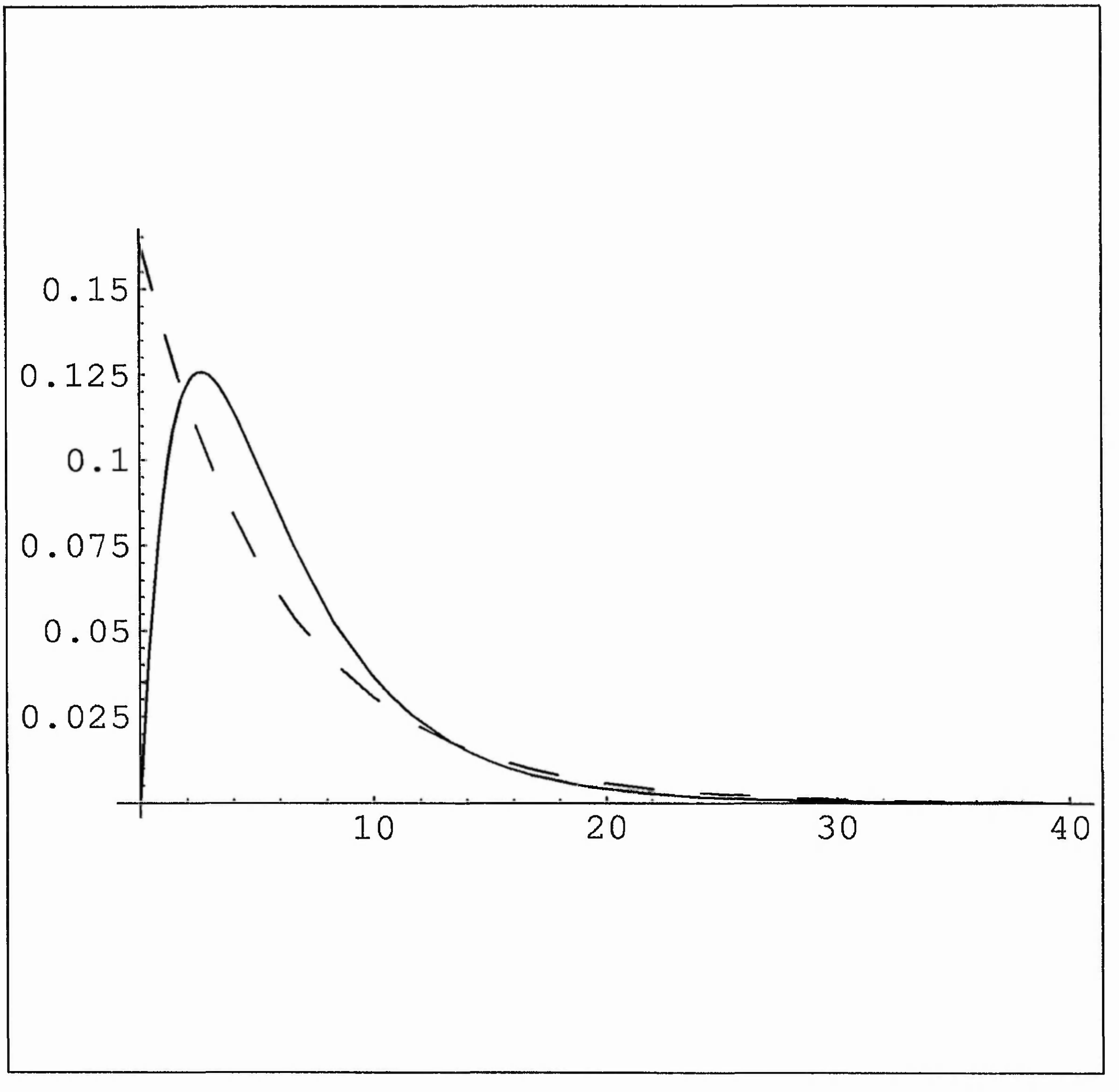


Figure 2.3: Approximation of the Maximum of two Exponential Variables with another Exponential Variable, ——$=f_e$, $- - - = f_a$

t	F_a	F_x
0.25	0.0001	0.0014
0.50	0.0056	0.0219
0.75	0.0374	0.0858
1.00	0.1141	0.1926
1.25	0.2317	0.3239
1.50	0.3700	0.4585
2.00	0.6287	0.6863
2.50	0.6287	0.6863
2.50	0.8060	0.8347
3.00	0.9058	0.9182
3.50	0.9562	0.9612
4.00	0.9803	0.9822
4.50	0.9913	0.9920
5.00	0.9962	0.9965
5.50	0.9984	0.9985
6.00	0.9993	0.9994
6.50	0.9997	0.9997
7.00	0.9999	0.9999

Table 2.1: Forbidden Network Results. (Probability Values for the Exact cdf, F_x, and Approximated cdf, F_a)

Chapter 3

A General Univariate Model using Laplace Transforms

3.1 Introduction

In the previous chapter we reported methods and techniques for assessing a univariate risk measure where traditionally a project is represented as a single characteristic activity network. In presenting a general univariate risk model, we recognise that sequential phases exist and thus a project network may be modified to incorporate individual sequential phases. Advantages of such a representation include the identification of phase measures, useful in monitoring the progress of a project and to aid decision making. Also calculations of project characteristics may be simplified if each sequential phase characteristic can be modelled by a suitable probability distribution that can be convoluted. In the unlikely event of each phase being identical, methods analogous to an ordinary renewal process where the distribution of component failure time is Erlangian may be used. In this chapter we consider the use of the special Erlangian distribution in modelling the times to complete network activities and project phases.

Currently the majority of risk analysis approaches employ simulation techniques to simulate the uncertainty associated with a project characteristic which is typically the project duration. We found that the software available for performing project type risk analysis typically employ Monte Carlo techniques to simulate the activities of the project network (see Appendix 1). In addition investigations into risk software distributed within the UK have revealed no unified approach in representation of the precedence of the project activities, however measures such as activity and path criticality may be easily generated.

Other than implementing the PERT approach as a computer program there have been few methods developed for commercial use. It is believed that the reasons for this include the computational difficulties associated with path dependency and criticality indices as outlined in Chapter 2. Also the convolution of continuous random variables as illustrated in Chapter 2 are cumbersome and hence the possible reason why only examples of discrete random variable convolutions have been documented in the context of a project risk analysis, Chapman et al (1983), and Dodin and Elmaghraby (1985).

In the following section we develop an algorithm that incorporates the approximation techniques of Dodin (1984a) and will collapse any project network structure. We provide examples where the exact and approximate project completion time distributions are obtained. If the parallel operator on quality loss is also the *maximum* then the algorithm may be utilised for determining the exact quality loss distribution. In the case of the average quality loss operator discussed in section 2.1.4, the algorithm may be used after modification.

To avoid some of the computational difficulties of convolutions we shall adopt the standard approach of Laplace transforms. In particular, to simplify the Laplace transform approach we suppose that each activity distribution may be suitably represented by an Erlangian density function. Unlike previous studies by Golenko-Ginzburg (1989) and Bendell (1995) we shall relax the condition requiring the activity distributions to be stable with respect to maximisation and convolution and obtain exact results assuming activity independence. In addition we shall explore the error in assuming that the maximum of Erlang densities is of the same type.

We implement the algorithm within the *Mathematica* software package of Wolfram (1988) and demonstrate, with examples, that the approximation measure to the exact completion time or quality loss is good and the measure of cost is exact.

3.2 Preliminaries

The advantage of employing Laplace transforms is that the integration difficulties involved in determining the explicit formula for the p.d.f of $X_1 + \cdots + X_n$ are avoided. Other measures of interest, such as the moments can be derived as the moment generating function may be determined from the Laplace transform simply by multiplying the dummy variable by -1, therefore enabling central moments to be computed.

Let $X_1, \ldots, X_n$ be continuous non-negative random variables independently distributed with p.d.f's $f_1(x), \ldots, f_n(x)$ and associated Laplace transforms denoted by $f_1^*(s), \ldots, f_n^*(s)$. Suppose that the sum of $X_1 + \ldots + X_n$ is denoted by the continuous random variable Y with p.d.f, $y_n(x)$. The Laplace transform of the p.d.f, denoted by $y_n^*(s)$, is determined as follows,

$$\begin{aligned}
y^*(s) &= E[\exp\{-s(X_1 + \ldots + X_n)\}] \\
&= E(e^{-sX_1} \ldots e^{sX_n}) \\
&= E(e^{-sX_1}) \ldots E(e^{-sX_n}) \\
&= f_1^*(s) \ldots f_n^*(s) \\
&= \prod_{i=1}^{n} f_i^*(s) \qquad (3.1)
\end{aligned}$$

where for completeness,

$$\begin{aligned}
f_i^*(s) &= E(e^{-sX_i}) \\
&= \int_0^\infty e^{-sx} f_i(x)\, dx. \qquad (3.2)
\end{aligned}$$

Once the product of Laplace transforms is known, the required p.d.f., $y(x)$ is obtained from inverting $y^*(s)$. An important result is that for all continuous functions, the function $y(x)$ can be uniquely determined. Inverting a Laplace transform is commonly referred to as the 'Inversion problem' which often involves recognising a function $y(x)$ for which $y^*(s)$ is the Laplace transform. Extensive tables are available, for example see Gradshteyn and Ryzhik (1980), Abramowitz and Stegan (1972), giving the Laplace transforms of common functions. In addition the *Mathematica* package may be utilised which enables the generation and inversion of Laplace transforms.

3.2.1 Use of Moment Generating Functions

The moment generating function of a convoluted function is given by the product of each individual m.g.f. To avoid confusion with the time variable, T, we shall denote the auxiliary variable of the m.g.f by r instead of the typical notation of t. Thus the m.g.f of Y is given by,

$$M_Y(r) = \prod_{i=1}^{n} M_{X_i}(r) \qquad (3.3)$$

From observing the similarity between a m.g.f which is a function of auxiliary variable r, specified by,

$$M(r) = \int_{-\infty}^{+\infty} e^{rx} f(x)\, dx \tag{3.4}$$

and the Laplace transform of a probability function, there is no need to determine the product of the moment generating functions as the auxiliary variable of the Laplace transform, often referred to as a dummy variable, can be multiplied by -1, hence giving the m.g.f, $M_Y(r)$. To avoid confusion between the functions we shall use the dummy variable s for Laplace transforms and the dummy variable r for moment generating functions.

3.2.2 Laplace Transforms using *Mathematica*

The Laplace transform package 'LaplaceTransform.m' version 2.0, originally written by Yehudai (1990) and extensively modified by Wolfram Research is based on tables of Laplace transforms by Oberhettinger and Badii (1973). It is distributed with *Mathematica* and enables Laplace transforms and inverse Laplace transforms to be determined. The package contains a list of functions which are selected according to the type of Laplace transform required. The limitations of the package include the non-exhaustive list of functions and, as noted by the author, the treatment of functions that are defined differently on different regions is needed yet would require substantial development of the package.

The relevant commands are as follows,

`Laplace Transform[`*expr*`, t, s]`	finds the Laplace transform of *expr*
`InverseLaplaceTransform[`*expr*`, s, t]`	finds the inverse Laplace transform of *expr*

3.2.3 Laplace Transform of a Beta Distribution

It is unsuitable at this stage to convolute beta distributions using the transformation (3.1), as the Laplace transform of the beta distribution is a confluent hypergeometric function which cannot be inverted once the convolution operator has been performed. In the general case, for all possible beta parameter values, computational difficulties are expected. However restricting one parameter to be of integer may enable convolutions to be performed. Further work, not performed by the author, is required in this area. If the approach is feasible it will enable the 'sum' operator of the PERT approach to be assessed without the need of simulation or approximation.

3.2.4 Special Erlangian Distribution

The special Erlangian distribution, sometimes referred to as the Erlang distribution is of the gamma distribution family where the shape parameter is restricted to integer values. It has been applied in the area of renewal theory, Cox (1962), where if it is assumed that failure takes place in α stages and the times spent in each stage are independent and exponentially distributed with identical densities, the time spent in all α stages is given by special Erlangian distribution defined by,

$$f(x) = \frac{\rho(\rho x)^{\alpha-1}e^{-\rho x}}{\Gamma(\alpha)} \tag{3.5}$$

where α is commonly referred to as the shape parameter, restricted to integer values and ρ is the scale parameter with positive value. For convenience in future discussion we shall refer to a special Erlangian distribution as an Erlang distribution. In addition a distribution is referred to as a general Erlangian distribution when the Laplace transform of the p.d.f. is a rational function of the dummy variable.

When $\alpha = 1$, the distribution is exponential. When $\alpha > 1$, the p.d.f is zero at the origin and has a maximum at $x = (\alpha - 1)/\rho$.

From integrating by parts it can be shown that,

$$F(x) = 1 - \sum_{r=0}^{\alpha-1} \frac{e^{-\rho x}(\rho x)^r}{r!} \tag{3.6}$$

Four cases of the special Erlangian distribution are shown in Figure 3.1. Each has a fixed mean of 1.

Similar to the PERT approach involving the beta distribution, the values a, m, b may be used to estimate the parameters of the distribution, see Bendell et al (1995). Thus in the case of the special Erlangian,

$$\hat{\mu} = \frac{a + 4m + b}{6} \tag{3.7}$$

$$\hat{\alpha} = \left(\frac{a + 4m + b}{b - a}\right)^2 \tag{3.8}$$

where,

$$\rho = \frac{\alpha}{\mu} = \frac{\text{shape parameter}}{\text{mean}} \tag{3.9}$$

From the variety of possible uncertainty representations indicated in Figure 3.1 and the parameter estimation for shape and scale, together with the fact that activity uncertainty is unlikely to be known exactly, we conclude that the special Erlangian provides the essential modelling capabilities to represent unimodal, skewed activity distributions. We are now in a position to consider both the distribution of a series of sequential activities and the distribution of parallel activities where each activity has a special Erlangian distribution. This is of particular interest as we found no reference to the application of special Erlangian distributions in the context of project network analysis.

3.3 Operators on a Special Erlangian

Desirable properties of the operations include (i) the ability to convolute Erlangian densities, (ii) determine the maximum of a series of convoluted Erlangian densities arranged in parallel, and importantly, (iii) be able to convolute the result of a maximum operation with a further Erlangian density. We shall address each area in turn and provide examples where appropriate.

3.3.1 Convolution of Erlangian Densities

This operation is relevant in the escalation of the time and quality loss of activities in series, the cost of activities either in series or parallel, and possibly the quality loss of activities in parallel. We present a general technique for the convolution of n activity densities which may be adopted for any of the characteristics. Since the Erlangian is a special case of a gamma distribution we could, as mentioned in section 2.3 of Chapter 2, use either of the results by Sim (1992) or Moschopoulos (1985) to convolute the special Erlangian distributions. However our interest at present is in the use of Laplace transforms. For convenience we shall sometimes use alternative notation, that is $\mathcal{L}[f_x]$ denoting the Laplace transform of p.d.f, f_x and $\mathcal{L}^{-1}[f_s]$ denoting the inverse Laplace transform of function, f_s.

Let Y represent the sum of the n random variables, $X_1, X_2 \ldots X_n$. The Laplace transform of a special Erlangian of activity i is given by,

$$f^*_{x_i}(s) = \frac{\rho_i^{\alpha_i}}{(s+\rho_i)^{\alpha_i}} \tag{3.10}$$

Thus the Laplace transform of the p.d.f. of $X_1, X_2 \ldots X_n$ is,

$$f^*_Y(s) = \prod_{i=1}^{n} \frac{\rho_i^{\alpha_i}}{(s+\rho_i)^{\alpha_i}} \tag{3.11}$$

As already mentioned a special case of the Special Erlangian is when the shape parameter, α_i, is equal to 1, which gives an exponential distribution. The

Laplace transform of the product of n Laplace transforms of the exponential type with scale parameters $\rho_1, \ldots, \rho_n$ is

$$\frac{\rho_1 \cdots \rho_n}{(\rho_1 + s) \ldots (\rho_n + s)}. \tag{3.12}$$

Note however, no matter the type of special Erlangian used for any of the n activities, the Laplace transform will always be of the form,

$$\frac{A(s)}{B(s)} \tag{3.13}$$

where,

$$B(s) = (s + \rho_1)^{\alpha_1} (s + \rho_2)^{\alpha_2} \cdots (s + \rho_n)^{\alpha_n}$$

3.3.2 Laplace Transform Inversion

A result of interest is provided by Eshbach and Souders (1975) where the inverse Laplace transform is given for $\mathcal{L}^{-1}\left[\frac{A(s)}{B(s)}\right]$. In this case $B(s)$ is of the form $(s - s_1)^{m_1}(s - s_2)^{m_2} \ldots (s - s_n)^{m_n}$. Recognising that $A(s)$ of the Laplace transform considered for n special Erlangian is always of the form $\prod_{i=1}^{n} \rho_i^{\alpha_i}$ a refined inversion for our case is given by,

$$f_Y(x) = \prod_{i=1}^{n} \rho_i^{\alpha_i} \sum_{k=1}^{n} \sum_{j=1}^{\alpha_k} \frac{\theta_{kj} x^{\alpha_k - j}}{(\alpha_k - j)!} e^{-\rho_k x} \tag{3.14}$$

where,

$$\theta_{kj} = \frac{1}{(j-1)!} \left[\frac{d^{j-1}}{ds^{j-1}} \frac{(s + \rho_k)^{\alpha_k}}{B(s)} \right]_{s=-\rho_k} \tag{3.15}$$

When $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 1$ the inversion may be performed in one step applying Heaviside's expansion theorem as defined in Abramowitz and Stegun

(1972). In brief, since the Laplace transform associated with the convolution of n exponential distributions is of the form $\frac{p(s)}{q(s)}$, where

$$q(s) = (s - a_1)(s - a_2)\ldots(s - a_n)$$

and $p(s)$ is a polynomial of degree less than n, the inverted Laplace transform, hence the required p.d.f is given by,

$$\sum_{i=1}^{n} \frac{p(a_i)}{q'(a_i)} e^{a_i x}$$

When two or more ρ_i are equal, modifications are necessary before inverting the Laplace transform.

3.4 Maximum of n-Special Erlangian Densities

In section 2.6.3 we presented the approach of Bendell et al (1995) for determining the central moments for the maximum of two Erlang densities. Here we present an approach for determining the exact distribution of n Erlang densities in parallel which is of a suitable form to enable convolutions with further Erlang densities. As a comparison of the accuracy of results obtained from performing a maximisation followed by a convolution, we demonstrate in Example 3 the approaches of simulation, method of moments, PERT and our functional form approach. In each case we assess the ability of the approach to determine suitable risk measures.

Difficulties that are experienced with other density functions, such as beta or normal, include the inability to define the distribution function in functional form. In such cases the integral can not always be evaluated in closed form. However since the special Erlang distribution exists in closed form we may proceed as follows. The p.d.f. of n distributions in parallel is given by,

$$g(x) = \sum_{i=1}^{n} \prod_{\substack{j=1 \\ j \neq i}}^{n} f_i(x) F_j(x)$$

Thus, the density of n-special Erlangians in parallel is given by,

$$f(x) = \sum_{i=1}^{n} \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{\rho_i(\rho_i x)^{\alpha_i - 1} e^{-\rho_i x}}{\Gamma(\alpha_i)} \left(1 - e^{-\rho_j x} \left[\sum_{k=0}^{\alpha_j - 1} \frac{(\rho_j x)^k}{k!}\right]\right) \tag{3.16}$$

For example the maximum of two paths in parallel, where each path is either Erlang distributed or the convolution of Erlang densities denoted by f_1 and f_2 with c.d.f.'s F_1 and F_2 respectively, is given by,

$$f_1 F_2 + f_2 F_1$$

Similarly the maximum of three paths in parallel is given by,

$$f_1 F_2 F_3 + f_2 F_1 F_3 + f_3 F_1 F_2$$

Since the Erlangian c.d.f. as given in (3.16) requires the parameter α_i to be known we may only be able to specify the density function of the maximum conditional on the shape parameters. We demonstrate the maximum operator with examples in the next section.

3.4.1 Use of Univariate Special Erlangian Activity Distributions

Example 3.1

Suppose that a project consists of n activities, where the cost distribution for each activity i is exponential with parameter ρ_i. Assuming all activity costs are independent and contribute to the total cost, the density function associated with project costs is given by,

$$f_C(c) = \sum_i^n A_i \rho_i e^{-\rho_i c}$$

where,

$$A_i = \prod_{j \neq i} \frac{\rho_j}{(\rho_j - \rho_i)}$$

Example 3.2

Suppose that a project comprises of three sequential phases, P_1, P_2, P_3, and that the quality loss experienced in each phase is Erlang distributed with parameters given in Table 3.1. The Laplace transforms of the p.d.f. of the total quality loss from project start to the completion of phase P_i, denoted by $\mathcal{L}[f_{CP_i}(ql)]$, are presented in Table 3.2. The explicit density functions of the total quality loss experienced to phase i may be determined with care using Eshbach's inversion formula or can be performed more efficiently by utilising *Mathematica*. From experience, the results generated by *Mathematica* even with using the 'Simplify[*expr*]' command, are not always in the simplest form. The inversions performed with *Mathematica* and hence the p.d.f.'s, $f_{CP_i}(ql)$, are shown below.

$$\begin{aligned}
f_{CP_1}(ql) &= \rho_1 e^{-\rho_1 ql} \\
f_{CP_2}(ql) &= \frac{\rho_1 \rho_2^2 (e^{-\rho_1 ql} - e^{-\rho_2 ql})}{(\rho_1 - \rho_2)^2} + \frac{\rho_1 \rho_2^2 ql e^{-\rho_2 ql}}{(\rho_1 - \rho_2)} \\
f_{CP_3}(ql) &= \frac{e^{-\rho_1 ql}}{(\rho_1 + \rho_2)^2 (\rho_1 - \rho_3)^2} + \frac{e^{-\rho_2 ql}(2\rho_1 - 3\rho_2 + \rho_3)}{(\rho_1 - \rho_2)^2 (\rho_2 - \rho_3)^3} \\
&+ \frac{e^{-\rho_3 ql}(2\rho_1 + \rho_2 - 3\rho_3)}{(\rho_1 - \rho_3)^2 (\rho_3 - \rho_2)^3} + \frac{ql e^{-\rho_3 ql}}{(\rho_1 - \rho_3)(\rho_3 - \rho_2)^2} \\
&+ \frac{ql e^{-\rho_2 ql}}{(\rho_2 - \rho_1)^2 (\rho_2 - \rho_3)^2} + \frac{ql e^{-\rho_3 ql}}{(\rho_1 - \rho_3)(\rho_2 - \rho_3)^2}
\end{aligned}$$

To avoid presenting all terms associated with $f_{CP_3}(ql)$ the first two central moments of the total quality loss experienced in all three activities may be obtained either by the moment generating function of f_{CP_3} or by simply working directly with the moments of each quality loss distribution as covered in section 2.6. Thus the mean, variance of f_{CP_3} are, $\frac{1}{\rho_1} + \frac{2}{\rho_2} + \frac{3}{\rho_3}$ and $(\frac{1}{\rho_1})^2 + 2(\frac{1}{\rho_2})^2 + 3(\frac{1}{\rho_3})^2$ respectively.

Example 3.3

Suppose there are two activities to be performed in parallel, where the time to complete each activity has a special Erlangian density. Consider the case where the density parameters are $f_1(t_1; \alpha_1 = 2, \rho_1)$ and $f_2(t_2; \alpha_2 = 3, \rho_2)$.

The time to complete both activities, $g(t)$, is determined from performing a maximum operation, thus,

$$g(t) = \rho_1^2 t e^{-\rho_1 t}(1 - e^{-\rho_2 t}(1 + \rho_2 t + \frac{(\rho_2 t)^2}{2})) + \frac{(\rho_2^3 t^2 e^{-\rho_2 t})}{2}(1 - e^{-\rho_1 t}(1 + \rho_1 t))$$

If a further activity is to be performed following the completion of both activities in parallel, which has density parameters $f_3(t_3; \alpha = 3, \rho_3)$ the time to complete all three activities, assuming path independence, may be determined from performing a convolution of the maximum result with f_3. The Laplace transform associated with $g(t)$ has the desired form, in that the denominator, $B(s) = (\rho_1 + s)^2(\rho_2 + s)^3(\rho_1 + \rho_2 + s)^4$ and thus the Laplace transform of the convoluted function is of the required form for inversion using either Eshbachs formulation or with *Mathematica*. Note that the explicit function for the completion density is long and therefore not presented here. Instead we show the moment generating function obtained directly from the Laplace transform,

$$M(r) = \frac{p_1^2 p_2^3 p_3^3}{(p_1 - r)^2 (p_2 - r)^3 (p_1 + p_2 - r)^4 (p_3 - r)^3}$$
$$\Big[p_1^4 + 4p_1^3 p_2 + 6p_1^2 p_2^2 + 4p_1 p_2^3 + p_2^4 - 6p_1^3 r - 20p_1^2 p_2 r - 24 p_1 p_2^2 r$$
$$-7p_2^3 r + 15p_1^2 r^2 + 40 p_1 p_2 r^2 + 21 p_2^2 r^2 - 20 p_1 r^3 - 25 p_2 r^3 + 10 r^4\Big]$$

where for given parameter values, for example where $\rho_1 = 0.3$, $\rho_2 = 0.2$ and $\rho_3 = 0.1$, we establish numerical values for the central moments. Thus in this case the moments to two decimal places are $\mu_1 = 45.81$, $\mu_2 = 368.11$, $\mu_3 = 6691.64$ and $\mu_4 = 597463.00$. With skewness of 0.95 and kurtosis of 4.41 we conclude the distribution is asymmetric and skewed to the right. If we apply the method of moments , with $\sqrt{\beta_1} = 0.95$ and $\beta_2 = 4.41$, interpolation is necessary. Using Bessel's interpolation formula where the second differences are formed, we obtained identical numerical values for the central moments.

To evaluate the accuracy of the approaches we identify risk levels for the distribution in functional form and from the use of Pearson curves. For convenience we define the risk level t_α such that $P[T \geq t_\alpha] = \alpha$. In addition for comparison we provide values for the PERT approach. Results are presented in Table 3.3, and as a check for accuracy the c.d.f obtained from a simulation of 10000 iterations is compared to the functional form in Figure 3.2. If we assume that the simulation results provide a basis for assessment of the accuracy it is clear that in this example the method for determining the approximate distribution generates accurate results. Also in agreement are the risk levels obtained with the method of moments. As expected the PERT approach underestimates the risk level for values in the extreme tail. An advantage of the functional form approach over the moments approach is that either the time value for a given risk level or the risk level associated with a given time can be obtained with ease. In addition for simple network configurations the moments approach is time consuming as interpolation of tables is often necessary, however as noted by Jaafar (1992) the approximation algorithm of Davis and Stephens (1983) may be implemented. Also noted by Jaafar (1992), the errors obtained when applying the method of moments are due to the assumption of activity independence.

We shall therefore consider complex project networks later in this chapter to further explore the use of the functional form approach.

Example 3.4

Suppose that a project phase consists of n activities each of which are performed in parallel where both the completion time density and quality loss density for each activity, i are Erlangian represented by $f_i(t; \alpha_i, \rho_i)$ and $g_i(ql; \beta_i, \varrho_i)$ respectively. Assuming that the cost to complete each activity is time dependent where the cost per unit time is x_i for given activity i and the unit cost of quality loss reduction is a constant value, Y, we may determine the expected cost of the phase. For realism we shall also assume that a penalty clause is in place, often specified in practice as a 'liquidated and ascertained damage' clause, Smith and Keenan (1979), where a cost of Z is charged for each time unit above the specified phase completion time of k. Thus the expected phase cost, C, is given by,

$$C = \sum_{i=1}^{n} x_i \frac{\alpha_i}{\rho_i} + Y \sum_{i=1}^{n} \frac{\beta_i}{\varrho_i} + Z \int_k^{\infty} t \sum_{i=1}^{n} \prod_{j \neq i} f_i(t) F_j(t) dt$$

Where an acceptable level of quality loss is set as θ the phase cost is adjusted by replacing the term $Y \sum_{i=1}^{n} \frac{\beta_i}{\varrho_i}$ with $Y \int_{\theta}^{\infty} ql \; g_p(ql) \; dql$ where $g_p(ql)$ is the density function of the total quality loss obtained from inverting the Laplace transform $\prod_{i=1}^{n} \frac{\varrho_i^{\beta}}{(\varrho_i + s)_i^{\beta}}$.

3.4.2 Remarks

In addition to the results presented in previous sections the Laplace transform of the density resulting from a maximum operation of Erlang densities is always a rational proper fraction and thus of the required form to enable further convolutions and maximum operations. We have demonstrated the effectiveness of the special Erlangian in representing uncertainty for the project characteristics and the ability to obtain exact solutions for the convolution and maximum operations.

The scope of *Mathematica* to perform the operations has been indicated with examples. Since a project may consist of numerous parallel and series connected activities it is desirable to have an approach which determines a measure for the overall project. Rather than considering the 'critical path' approach associated with PERT, we concentrate on possible methods of network reduction and implement a modified version of the method proposed by Dodin (1984a).

3.5 Procedures for Collapsing a Network

The techniques for performing convolutions and determining the maximum, although well defined, performing many of these operations is involved and thus not feasible in a large project network consisting of many parallel paths where each may share common activities. Difficulties are also apparent where independent paths cannot be assumed. Such cases have been discussed in section 2.5.1. Our objective is to define an algorithm which is efficient, will collapse any network of Erlang activities, and although will not always determine the exact project measure distribution, will generate the best lower bound approximation similar to the method proposed by Dodin (1984a). As commented by Sculli (1983), and Jaafar (1992) determining the c.d.f may not always be possible analytically for combinations of various distributions and hence bound to cause computational difficulties.

With both the objective and modelling difficulties in mind we shall consider two methods for collapsing a project network. Method A has been used by Jaafar (1992) as a means to evaluate the moments approach and Method B as used by Dodin (1984a) with exponential activities shall be developed as an algorithm and implemented with *Mathematica.*

3.5.1 Method A - Independent paths

Consider all j paths in the network as a series of activities For each path i of length n, perform the convolution of the n densities ,

$$p_i = f_1 * f_2 * \cdots * f_n \tag{3.17}$$

To evaluate the overall project time, perform a maximisation of the j paths,

$$f(t) = \sum_{i=1}^{j} \prod_{k \neq i} p_i(t) P_k(t) \tag{3.18}$$

3.5.2 Method B - Dynamic Update

Although a similar type of analysis has been used effectively in practice by the author for simulating a research and development programme where each activity duration is represented by a triangular probability distribution neither the implied algorithm has been documented in relevant project risk literature, nor the use of Erlang activity measures reported.

The essence of this method is as follows. Scan the network from left to right, performing a summation if a single activity leads to a node or a maximisation if more than one activity leads to a node. This in effect collapses a network from start to completion activity reducing paths where maximisation occurs to a single density function. This method is simple and only requires at maximum a convolution of two densities for any operation. It is assumed that each activity is independent.

We begin by stating notation necessary for the development of an algorithm to perform the dynamic update network reduction method.

Notation

Let $G(A, N)$ denote a directed, connected, acyclic graph where A is defined as the ordered set of activities, $\{a_1, a_2, \ldots a_n\}$. For a network with k events, N is a $k-1 \times k$ matrix, see Figure 3.3. Let $n_{i,j}$ denote the activity that is performed between event i and event j. Where $n_{i,j} = 0$ there is no activity connecting event i to event j.

Where each activity time is special Erlangian, let $a_i = f_i(t; \rho_i, c_i)$.

Let q denote the number of activities leading into node j. The value of q is important, in that it determines the operation to be performed. If $q = 1$ a convolution is performed. If $q > 1$ then the maximum of q parallel paths is found.

We also require a few additional variables specific to the algorithm defined. Let $\text{store}_q = [i, [i, j]]$. This collects information regarding the activities and preceeding events leading into a given event j, essential for computing the distribution of parallel activities. V is a $k-1$ array of density functions where v_j stores the time distribution to complete all activities up to event j.

Algorithm

1. Start at position $i = 1, j = 1$ of N.
2. Set activity count to zero, $q = 0$.
3. Scan N from top to bottom, (from $i = 1$ to $k-1$)

   If $N_{i,j} \geq 1$

   - increment activity count, q.
   - store previous event i together with the activity number, labelled a in $\text{store}(q, i, a)$.

4. After scanned column,

   If count$= 1$ Perform convolution, using store$(1, i, a)$, of $a * v(i)$.

   If count> 1 Perform parallel operation of the q paths. Firstly compute the convolution of each path leading to event j, (store$(1, i, 1)$ to store(q, i, a)), then compute maximum.

   Record result Store the result of either the convolution or parallel operation in $v(j)$.

5. Move to the $j + 1^{\text{th}}$ column and repeat from Step 2.

6. End.

3.5.3 Implementation of Algorithm

Performing the above algorithm on a k event network in effect simplifies the network to a single activity with time distribution $v_k(t)$. In addition the time distribution in reaching event j, where $j < k$ is given by $v_j(t)$. *Mathematica* procedures, see Appendix 2, were used to implement the algorithm, primarily because of the availability of the Laplace transform package to facilitate the convolution of the activity densities and the ease of matrix type network representation using lists of lists. A disadvantage of such an implementation included the execution time, where even for small networks times at least one day were experienced, sometimes without a result. A possible improvement in execution time and allowing for complex network analysis may be obtained in implementing the algorithm in a language such as C++, however this approach would require the development of Laplace transform generation and inversion procedures.

3.5.4 Network Reduction Examples

Surprisingly, apart from the *forbidden network* configuration, there are few appropriate project network examples in which asymmetric probability distributions are considered for activity characteristic. To demonstrate the algorithm we shall consider four examples, in which paths through all networks are near critical and compare the results to those obtained from Method A, PERT and depending on the complexity of the path dependency, present either the exact or simulated results. For all example networks considered, the algorithm procedure was executed on a Sun Workstation with *Mathematica* version 2.2 for SPARC.

Example 3.5

Consider a project where five activities are arranged as a *forbidden network*. For comparison to the exact solution of (2.21) we shall assume that each activity time is exponential with parameter λ.

The input values to the computer procedures, corresponding to the forbidden network configuration of Figure 2.2 are,

$$\text{network} = \{\{0,2,1,0\},\{0,0,3,4\},\{0,0,0,5\}\}$$

and

$$\text{activity} = \{\{\lambda,1\},\{\lambda,1\},\{\lambda,1\},\{\lambda,1\}\}$$

The generated c.d.f for the project completion time is,

$$\begin{aligned} F(t) &= 1 - e^{-\lambda t}[3\lambda t + 0.5(\lambda t)^2] \\ &\quad - e^{-2\lambda t}[3 - 2.5(\lambda t)^2 - 0.5(\lambda t)^3] \\ &\quad + e^{-3\lambda t}[2 + 3\lambda t + (\lambda t)^2]. \end{aligned}$$

For comparison, as indicated in Figure 3.4, c.d.f's are graphed for the exact distribution and those generated when implementing Method A, Method B with $\lambda = 2$. In addition we present risk values generated with both the exact and approximated distribution, see Table 3.4. We considered using the method of moments, however to enable the use of the maximum operation we must assume that the Erlang distribution is stable under maximisation. This assumption, however, is not needed when using the functional form, and therefore is considered more accurate than employing the method of moments. Further consideration of the stability assumption is covered in Example 3.7.

Example 3.6

This example was taken from Bendell at al (1995), where the network of Figure 3.5 is analysed with Erlang activity densities. Using the parameters presented by Bendell where the density shape parameter take on values between two and ten revealed a limitation of the *Mathematica* package on our computer system. It appears, as warned in Wolfram (1991), the intermediate expressions used in generating and inverting Laplace transforms together with determining the cumulative distribution functions when performing maximisations is demanding on computer memory and may cause the computer to halt. This happened for medium sized networks and above, even when the shape parameter for each activity is set to one. However in order to demonstrate the effectiveness of the algorithm we implement the aforementioned network with activity density input parameters, $\rho_A = 4, \rho_B = 5, \rho_C = 2, \rho_D = 3, \rho_E = 4, \rho_F = 3, \rho_G = 1, \rho_H = 2, \rho_I = 1$ and assume the shape parameter for all densities is two.

To assess the accuracy of this approach we determine the c.d.f of the project duration with our algorithm and compare with results obtained from simulation, PERT, Method A and the method of moments approach where we match the first two moments, see Figure 3.6. In addition the risk levels are specified for each approach, see Table 3.5. As previously assumed, if once again we believe the simulation approach to give the most accurate results then in this example the use of our algorithm provides the best approximation compared to the method of moments and Method A.

Remarks

In performing the method of moments approach, a similar procedure to Method B was adhered to which as concluded by Jaafar (1992) does not violate the independence assumption as much as Method A. However, where a maximisation is performed, stability is assumed in that an Erlang density can suitably approximate the true function, from matching the first two moments. Our approach in using the algorithm, although in some instances may take longer to perform will, in general, generate more accurate results as estimation of parameters after maximisation is not required. A possible refinement to the method of moments approach of Jaafar would be to implement our algorithm and incorporate the facility to determine variate values of non-normal univariate frequency curves.

To overcome the limitations caused by the intermediate calculations within *Mathematica* we suggest, if feasible, to restructure the project yet still implement the algorithm. Suppose a project comprises of n independent project phases where for each phase i there are k_i independent parallel activities. Rather than considering the entire project as a single network we may isolate the phases, thus representing the project network as n independent networks and use the algorithm at each stage to determine the phase risk measure distribution function. Care must be taken in the correct representation of parallel networks to be implemented within the algorithm where for each phase $k_i + 1$ events are required.

In addition we suggest, instead of convoluting the exact probability densities, to match the first two moments in order to approximate a suitable Erlang distribution for each project phase. This in turn requires less time to compute, however a certain amount of accuracy will be lost. In practice we believe this not to be of great concern since in most cases the activity densities are based on subjective estimates and thus the errors caused by approximation are not significant. We demonstrate the approximation approach from matching moments in the following example.

Example 3.7

Here we consider the use of matching moments of the phase distribution in order to approximate a suitable Erlangian distribution as outlined in the concluding remarks of Example 3.6. Suppose a project phase consists of two activities in parallel, where we use Erlangian densities, f_1, f_2, to specify the activity completion times from subjective information given as optimistic, most likely, and pessimistic times. Given that the subjective information for activities one and two are $a_1 = 1, m_1 = 4, b_1 = 13$ and $a_2 = 2, m_2 = 3, b_2 = 22$ suitable parameter estimates for f_1 and f_2 are, $\alpha_1 = 6, \rho_1 = 1.2$ and $\alpha_2 = 3, \rho_2 = 0.5$ respectively. Results which may be easily obtained include the p.d.f, c.d.f, and the moments associated with the phase completion time. Given the first two central moments, μ and σ^2, we may use the parameter estimators $\rho = \frac{\mu}{\sigma^2}$ and $c = (\frac{\mu}{\sigma})^2$, as given in Hastings and Peacock (1974), to approximate the maximum of two Erlang variables with another Erlang variable. Thus with $\mu = 7.065$ and $\sigma^2 = 8.889$ to three decimal places, suitable estimates for the scale and integer valued shape parameter are $\rho = 0.795$ and $\alpha = 6$. The exact p.d.f, f_x and approximated p.d.f, f_a are shown in Figure 3.7 indicating a reasonable visual fit. However in the context of approximating the risk associated with phase we are more concerned with accurately evaluating tail probabilities of the type $\int_k^\infty f_a(t)\, dt$. In this case we indicate the risk levels for given times, which may be considered as target times, located in the upper tail of the distribution and the error from the approximation.

Remarks

The error involved in assuming that the maximum of two Erlang variables is also an Erlang variable will depend on the particular values of the parameters. As covered in Section 2.6 we may use a Kolmogorov-Smirnov goodness of fit test to assess the significance of the errors. For comparative purposes we tested the Erlang approximation for a collection of subjective time estimate examples where we have assumed activity time is suitably represented by a unimodal density with skew greater or equal to zero.

Our results are shown in Table 3.6 where we report the test statistic, D, defined as $D = \sup_t |F_a(t) - F_x(t)|$, the mean absolute error. In addition an analysis of the risk levels is presented in Table 3.7 where the error in the risk levels at five time points located at equal interval above the mean, given by $\mu + \beta\frac{b_1+b_2-\mu}{5}$ where $\beta = 1, 2, \ldots, 5$. Also given in Table 3.7, we indicate the time error for a given risk level of $\alpha = 0.05$. Possible further analysis includes examination of the risk levels for cases where different distributions are used to model the subjective information, such as the errors involved from use of beta and Erlangian densities.

3.6 Conclusions

The use of Laplace transforms is a distinct advantage in convoluting project activity time distributions. In Chapter 4 we shall simplify our project representation to a number of successive phases in series and consider modelling both time and cost as jointly distributed random variables.

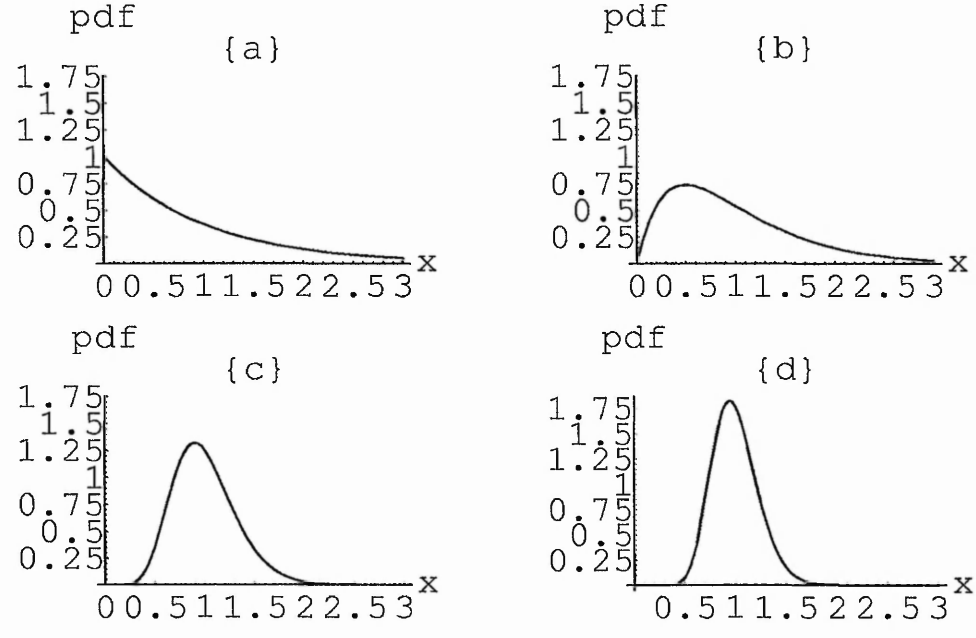


Figure 3.1: Cases of the Special Erlangian Distribution of Mean 1, (a) $\alpha = 1$, (b) $\alpha = 2$, (c) $\alpha = 10$, (d) $\alpha = 20$

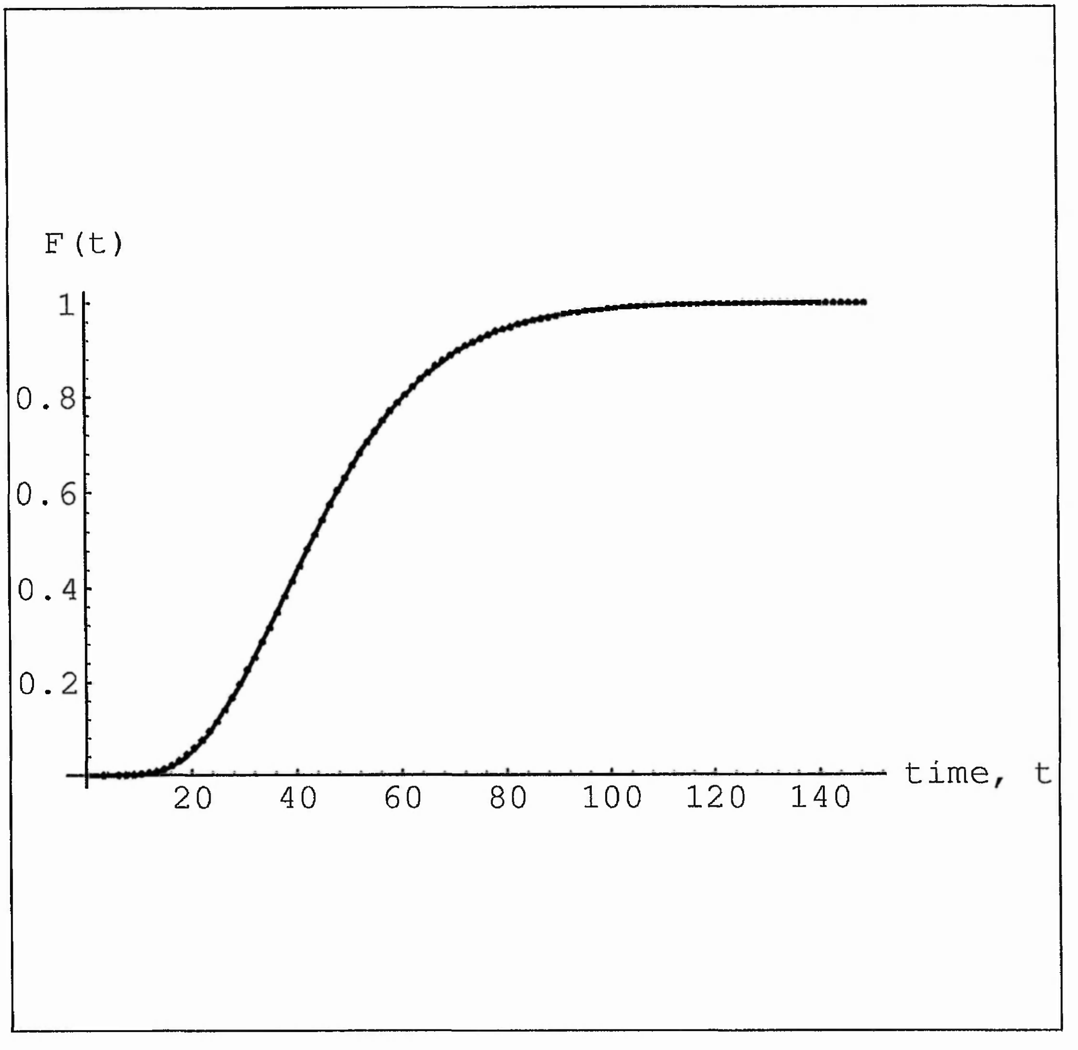


Figure 3.2: Functional Form Results Compared to Simulation Results of Example 3.3, ...simulation,—functional form

$$
N = \underbrace{\left(\begin{array}{cccc} n_{1,1} & n_{1,2} & \cdots & n_{1,k} \\ n_{2,1} & n_{2,2} & & \\ \vdots & & \ddots & \vdots \\ n_{k-1,1} & & \cdots & n_{k-1,k} \end{array} \right)}_{k \text{ columns}} \Bigg\} k-1 \text{ rows}
$$

Figure 3.3: Network Matrix

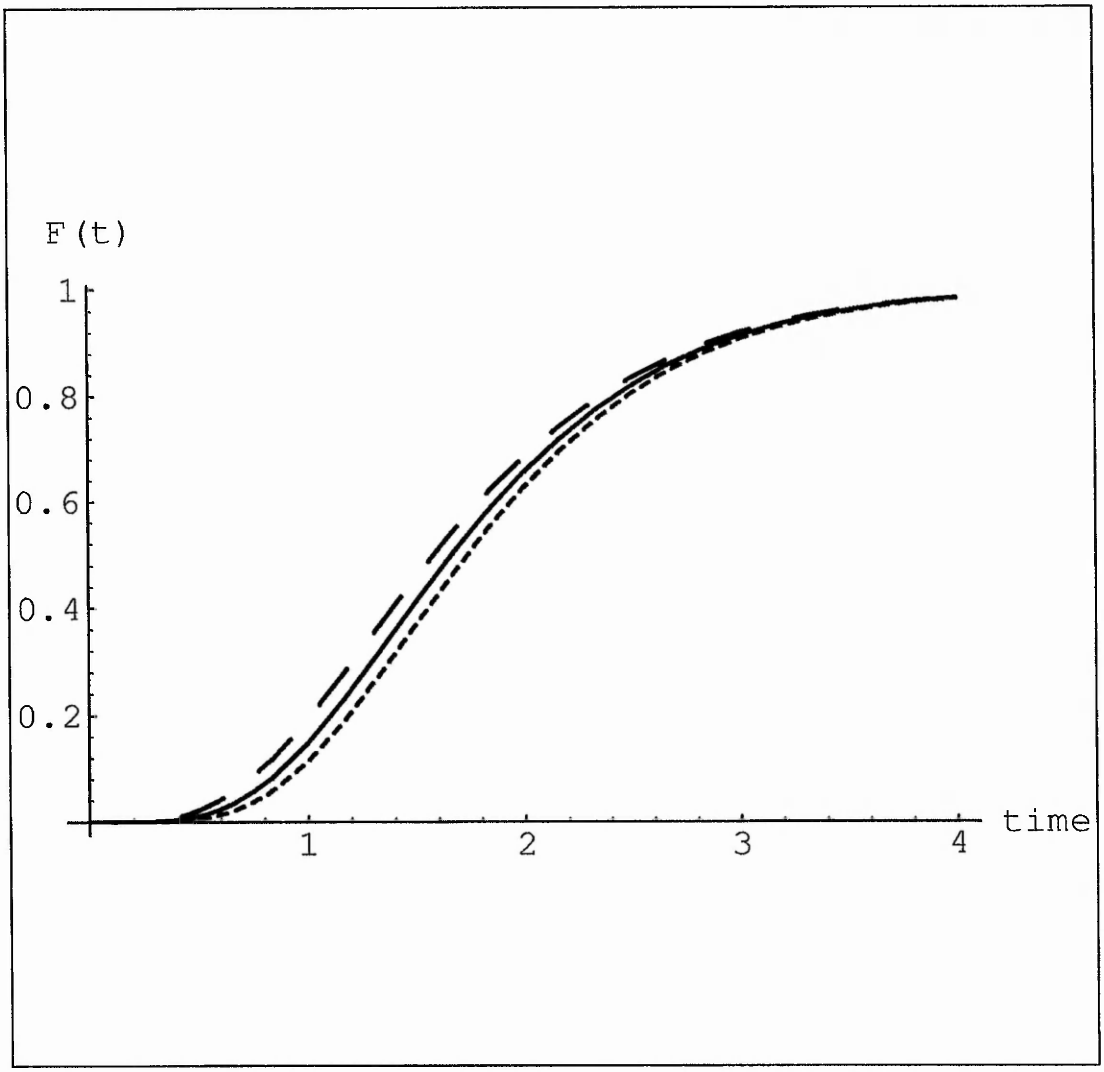


Figure 3.4: Distribution Functions for Forbidden Network, Example 3.5 (dash line - exact, line - method B, dotted line - method A

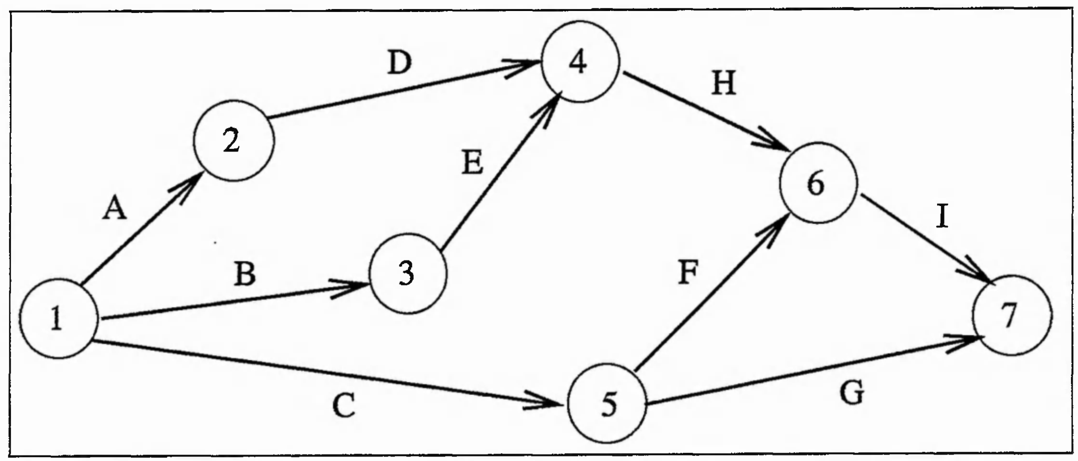


Figure 3.5: Network diagram for Example 3.6

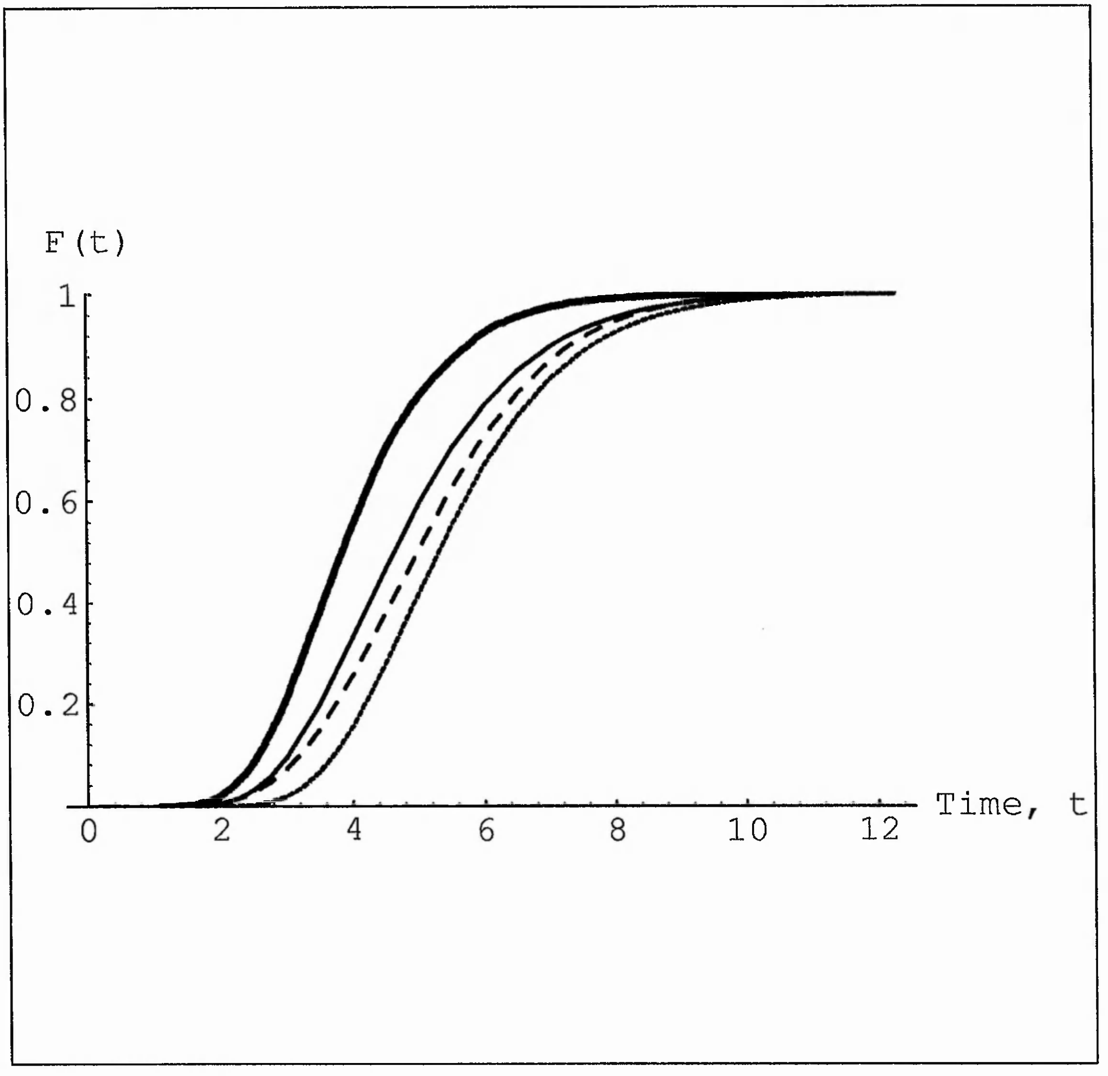


Figure 3.6: Distribution Functions for Example 3.6, (bold line - simulation, line - Method B, dotted line - Method A, dashed line - Method of Moments

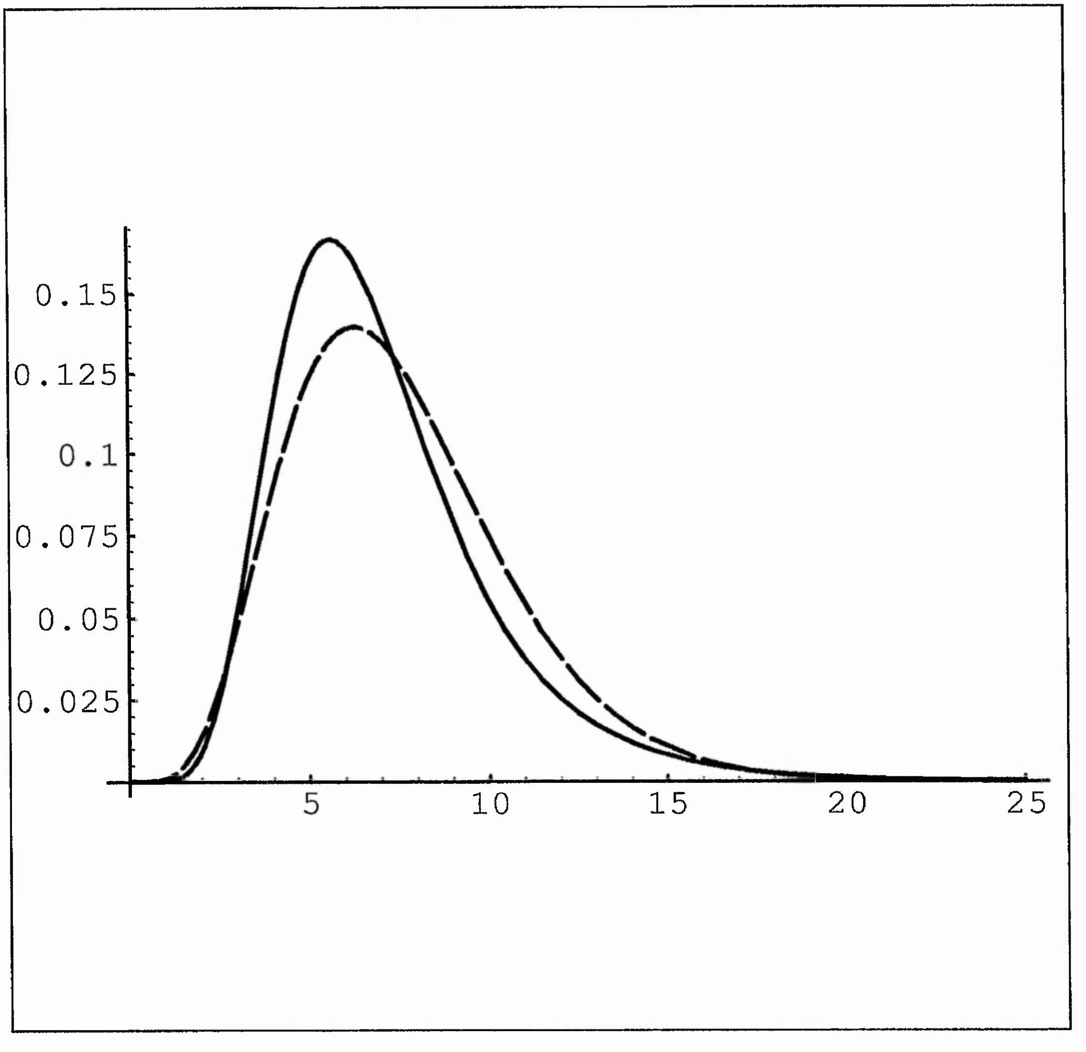


Figure 3.7: The Approximation of the Maximum of two Erlang Random Variables, ——$=f_x$, – – –$=f_a$

	Input parameters	
Phase	α	ρ
P_1	1	ρ_1
P_2	2	ρ_2
P_3	3	ρ_3

Table 3.1: Erlang Parameters for Example 3.2

Completion of Phase i	$\mathcal{L}[f_{CP_i}]$
1	$\frac{\rho_1}{(\rho_1+s)}$
1 and 2	$\frac{\rho_1\rho_2^2}{(\rho_1+s)(\rho_2+s)^2}$
1, 2 and 3	$\frac{\rho_1\rho_2^2\rho_3^3}{(\rho_1+s)(\rho_2+s)^2(\rho_3+s)^3}$

Table 3.2: Quality Loss Laplace Transforms for Example 3.2

Risk Level, α	Functional form	Approach Method of moments	PERT
0.010	103.25	103.22	90.04
0.025	91.14	91.15	80.99
0.050	81.55	81.58	76.85
0.100	71.41	71.45	69.81
0.250	56.55	56.56	58.06

Table 3.3: Risk levels for Example 3.3

		Approach	
Risk Level, α	Exact	Method A	Method B
0.010	4.35935	4.41462	4.38787
0.025	3.78345	3.85226	3.81921
0.050	3.33223	3.41449	3.37528
0.100	2.86094	2.96024	2.91331
0.250	2.18423	2.31277	2.25278

Table 3.4: Risk levels of Example 3.5

Risk Level, α	Simulation	Approach Algorithm (Method B)	Method of Moments	Method A
0.010	7.56	9.84	9.82	10.40
0.025	7.08	8.76	8.88	9.34
0.050	6.33	7.92	8.12	8.51
0.100	5.69	7.04	7.32	7.65
0.250	4.72	5.78	6.11	6.42

Table 3.5: Risk levels of Example 3.6

	Activity 1					Activity 2					Fitted Erlang			Mean abs.
Example	a_1	m_1	b_1	ρ_1	α_1	a_2	m_2	b_1	ρ_2	α_2	ρ	α	D	Error
1	1	2	6	3.60	9	1	2	7	2.63	7	3.97	1	0.0308	0.0058
2						1	2	10	1.26	4	1.94	7	0.0475	0.0081
3						1	3	7	3.30	11	4.18	15	0.0343	0.0045
4						1	3	10	1.83	7	2.29	9	0.0281	0.0044
5						1	4	7	4.00	16	4.49	18	0.0243	0.0056
6						1	4	10	2.00	9	2.25	10	0.0296	0.0080
7						1	5	10	2.32	12	2.26	12	0.0337	0.0074
8	1	3	6	4.42	14	1	2	7	2.63	7	4.83	17	0.0291	0.0037
9						1	2	10	1.26	4	2.56	10	0.0536	0.0088
10						1	3	7	3.30	11	4.79	18	0.0174	0.0023
11						1	3	10	1.83	7	2.80	12	0.0593	0.0083
12						1	4	7	4.00	16	4.88	20	0.0458	0.0082
13						1	4	10	2.00	9	2.57	12	0.0216	0.0046
14						1	5	10	2.32	12	2.63	14	0.0346	0.0054
15	2	4	6	9.00	36	1	2	7	2.63	7	5.81	25	0.1281	0.0173
16						1	2	10	1.26	4	4.46	20	0.1047	0.0126
17						1	3	7	3.30	11	8.45	36	0.0305	0.0035
18						1	3	10	1.83	7	4.38	20	0.0612	0.0087
19						1	4	7	4.00	16	7.65	34	0.0230	0.0032
20						1	4	10	2.00	9	3.50	17	0.0530	0.0087
21						1	5	10	2.32	12	3.17	17	0.0411	0.0072

Table 3.6: The Approximation of the Maximum of two Erlang Random Variables by another Erlang Random Variable

			β			5% risk level	
Example	1	2	3	4	5	abs. risk level Error	% risk level Error
1	0.0073	0.0006	0.0000	0.0000	0.0000	0.1190	2.53
2	0.0023	0.0031	0.0004	0.0000	0.0000	0.0612	1.00
3	0.0009	0.0004	0.0000	0.0000	0.0000	0.0315	0.61
4	0.0095	0.0019	0.0001	0.0000	0.0000	0.1768	2.72
5	0.0085	0.0006	0.0000	0.0000	0.0000	0.1144	1.98
6	0.0136	0.0023	0.0002	0.0000	0.0000	0.2268	3.14
7	0.0168	0.0026	0.0002	0.0000	0.0000	0.2080	2.65
8	0.0002	0.0003	0.0000	0.0000	0.0000	0.0353	0.71
9	0.0039	0.0036	0.0004	0.0000	0.0000	0.0399	0.64
10	0.0018	0.0004	0.0000	0.0000	0.0000	0.0059	0.10
11	0.0006	0.0019	0.0001	0.0000	0.0000	0.0007	0.01
12	0.0085	0.0007	0.0000	0.0000	0.0000	0.1147	1.97
13	0.0069	0.0024	0.0002	0.0000	0.0000	0.1332	1.85
14	0.0029	0.0009	0.0001	0.0000	0.0000	0.0103	0.13
15	0.0217	0.0002	0.0000	0.0000	0.0000	0.5262	9.95
16	0.0083	0.0035	0.0004	0.0000	0.0000	0.0556	0.90
17	0.0031	0.0004	0.0000	0.0000	0.0000	0.0202	0.37
18	0.0128	0.0024	0.0001	0.0000	0.0000	0.1352	2.08
19	0.0088	0.0008	0.0000	0.0000	0.0000	0.0830	1.42
20	0.0166	0.0037	0.0003	0.0000	0.0000	0.2850	3.91
21	0.0094	0.0033	0.0003	0.0000	0.0000	0.1725	2.20

Table 3.7: Risk Level Errors between the Maximum of two Erlang Random Variables and the Approximated Erlang Variable

Chapter 4

Models of Dependent Project Phase Characteristics

4.1 Introduction

We have assumed, based on our formulation stated in section 1.7, that project characteristics of concern which enable assessment of the project objectives are the time, the cost, and the level of quality loss incurred. Although the project objectives influence which measures are of key importance, it is believed that time-scales and budgets are typically of most concern and are relatively easy to quantify compared to a quality loss measure. In support of this much effort is expended in practice in the collection and updating of information to enable the forecasting of time and costs. Evidence of this is reflected in the methods commonly applied, Moder et al (1983), Nicholas (1990) and Turner (1994).

Even though time and cost are common measures the relationship between time and cost is often overlooked. To date only the time-cost slope of the classic CPM technique as covered in Moder et al (1983) has been used effectively to relate time and cost. This technique, however, is limited, in that uncertainty cannot be modelled and also is not the best model which reflects the optimal project compression from an increase in resources. A recent review by Elmaghraby (1995), indicates the lack of progress for a much demanded realistic relationship between time and cost and concludes that there is an important gap to fill.

In addition to assessing the ability in meeting targets for given project characteristics, other measures may be related to one or more of the overall project characteristic measures. It is often a project planner's concern to have an indication of the performance of the project outcome. Performance in this context may refer to a specific project output, PO, or a collection of i outputs of the project indicated by PO_i. This may include output measures, efficiency, reliability, or even income generated. In the area of electronic systems the performance measure, reliability is an essential part of the engineering programme. In an engineering context, reliability is considered an integral part of the product development and, as shown by many competitive, high technological companies, reliability is a high value property. The effects of high reliability help ensure low cost of repairs and maintenance together with less quantifiable effects such as customer goodwill and product reputation, O'Connor (1991).

In the context of project planning, where projects are often one-off ventures, different concepts and techniques require development to enable a new class of project management model to be established. As indicated in the introduction of Chapter 1, there is very little development of a model incorporating such features other than the qualitative relationships as discussed by Barnes (1988), Ward et al (1991), Kidd (1987), and Klein (1993).

To obtain realistic measures, suitable relationships and structured models require formulating. In this chapter we define a suitable bivariate time-cost distribution and examine the criteria for project success. In addition we present an application of the proportional hazards model where the reliability of the project outcome can be related to the characteristic value for all project phases.

4.2 Time-Cost Model

For modelling purposes we use a bivariate random variable, specifically of exponential type, which allows the uncertainty and dependency between time and cost to be modelled. Literature on the application of bivariate exponential distributions, apart from the distribution of Gumbel (1960) and Farlie-Gumbel-Morgenstern, see Johnson and Kotz (1972), is mainly in the area of reliability modelling. The bivariate exponential is used to model the lifetime of components in a two component system. One of the first multivariate distributions involving the exponential distribution was that of Freund (1961). This two component system model involved the lifetime of one component depending on the other. Since the marginal distributions were not exponential it was called a bivariate exponential extension. The most widely referenced distribution is that of Marshall and Olkin (1967) which gives rise to the following models; 'fatal shock' model, 'non-fatal shock' model, and a random sums model. Similarly Downton (1970) used a special case of the bivariate gamma distribution due to Kibble (1941) in modelling the times between shocks and the number of shocks to failure. For a complete review of bivariate exponential distributions, see Kotz et al (1983).

4.2.1 Construction of a Time-Cost Distribution

Intuitively the more time spent on a particular activity the greater chance that the activity will cost more. This relationship is illustrated in Figure 4.1 where the threshold/shift in the cost distribution is dependent on the realisation time. We use this idea to construct a bivariate distribution specific to this context. As a starting point, since the exponential distribution has been used to model time in the univariate case, Dodin (1984) and Kulkarni and Adlakha (1986), we construct a suitable bivariate exponential type distribution and investigate the use of a convolution operator. Initial investigations were directed in formulating a bivariate function incorporating dependency between the exponential parameters. Computational difficulties were encountered in generating and inverting the Laplace transform and therefore another model specification was searched for.

Work by Seshadri and Patil (1964) showed a bivariate distribution may be constructed from a conditional density assuming that the conditional is of univariate exponential form. The random variable C will be used for cost and the random variable T used for time. In practice the costs incurred will often depend on the time spent on the activity. To incorporate this feature, the threshold value of the cost density depends on the realisation of the random variable T. The joint density function and the conditional density function are given by,

$$f(c,t) = f(c|t)f(t) \tag{4.1}$$

$$f(c|t) = \frac{f(c,t)}{f(t)} \tag{4.2}$$

Suppose that activity time can be modelled using an exponential distribution with density function having a threshold parameter, b,

$$f_T(t) = \gamma exp[-\gamma(t-b)] \quad \gamma > 0,\ b \geq 0,\ t > b. \tag{4.3}$$

and since cost is dependent on time realised a plausible relationship is given by,

$$f(c \mid t) = \lambda exp[-\lambda(c - g(t;\alpha,k))] \tag{4.4}$$

where,

$$\lambda > 0$$

$$g(t;\alpha,k) \geq 0$$

$$c > g(t;\alpha,k)$$

The function $g(t;\alpha,k)$ enables the threshold value of the cost density to be dependent on the time realised. In the most simple case the threshold value is a linear function of the time realised,

$$g(t;\alpha,k) = \alpha + kt \quad \alpha \geq 0 \; k > 0 \tag{4.5}$$

The intuitive appeal of such a function includes that ability to model the fixed cost of performing an activity by assigning an appropriate value to the threshold parameter α. Applying equation (4.1), gives the joint p.d.f. for activity time and cost,

$$f(c,t) = \lambda\gamma exp\left[-\gamma(t-b) - \lambda(c-(\alpha+kt))\right] \tag{4.6}$$

where,

$$c > \alpha + kt$$

$$t > b$$

The joint c.d.f. for time and cost is thus given by evaluating,

$$F(c,t) = \int_b^{\min(t,\frac{c-\alpha}{k})} \int_{\alpha+kt}^{c} f(x,y)\, dx\, dy \tag{4.7}$$

In the simplest case of exponential time we may suppose that $b = 0$. Where there is no dependency or thresholds the joint density function reduces to the product of two independent exponential densities,

$$f(c,t) = \lambda\gamma e^{-\gamma t - \lambda c}$$

From the conditional distribution the expected value of C given $T = t$ can be identified,

$$\begin{aligned} E[C \mid T=t] &= \frac{1}{\lambda}\int_{g(t;\alpha,k)}^{\infty} c\, exp[-\lambda(c-(\alpha+kt))]dc \\ &= \frac{\lambda kt + \lambda\alpha + 1}{\lambda^2} \end{aligned} \tag{4.8}$$

Also the expected cost value may be determined,

$$\begin{aligned} E[C] &= \int_{-\infty}^{\infty} E[C|T=t] f_T(t)dt \\ &= \frac{\gamma + \alpha\gamma\lambda + \lambda k + b\gamma\lambda k}{\gamma\lambda^2} \end{aligned} \tag{4.9}$$

The marginal density of C is obtained by integrating the joint density with respect to t. Note the upper limit is determined from the threshold relationship $c > \alpha + kt$, hence $t < \frac{c-\alpha}{k}$.

$$\begin{aligned} f(c) &= \int_{b}^{\frac{c-\alpha}{k}} \gamma\lambda exp[-\lambda(c-(\alpha+kt)) - \gamma(t-b)]dt \\ &= \frac{\gamma\lambda}{\gamma - \lambda k}\left(e^{\lambda(\alpha-c+bk)} - e^{\frac{\gamma}{k}(\alpha-c+bk)}\right) \end{aligned} \tag{4.10}$$

where $c > \alpha + bk$.

Remarks

As for Freund's distribution, since the marginal distribution is not exponential we shall also call our distribution a bivariate exponential extension. Further work is possible investigating the case where $\frac{k}{\gamma} = \frac{1}{\lambda}$.

4.2.2 Examples of Time-Cost Distributions

We illustrate the modelling capabilities of the time-cost distribution by graphing the joint p.d.f. for a variety of distribution parameters. Figure 4.2 illustrates a distribution where no dependency between time and cost exists and thresholds are set to zero.

The fixed cost, FC, of performing an activity may be incorporated by setting $\alpha = FC$. Similarly the least time that an activity will take to perform is accounted for with a suitable value of b. A distribution where the fixed cost is four units and the least time is two units, yet time and cost remain independent, is shown in Figure 4.3(a). The dependency between time and cost is incorporated by setting $k > 0$. Increasing the value of k and keeping all other parameters constant, increases the range of possible cost values for any given time as shown in Figure 4.3(b) where $k = 1$, Figure 4.3(c) where $k = 2$ and Figure 4.3(d) where $k = 5$.

Based on the approach outlined in Chapter 2 for convoluting probability density functions, Laplace transforms and inversions are now considered for our bivariate exponential density function.

4.3 A Time-Cost Project Model

For each phase assume that a bivariate exponential extension distribution adequately describes the relationship between time and cost. See Figure 4.4 for project representation.

Although we naturally concentrate upon using an exponential type distribution to avoid computation and algebraic difficulties, the time-cost relationship is of great benefit in providing a guideline to the performance of the project. Similar to the use of Laplace transforms of special Erlangian distributed activity characteristics, we explore the use of two dimensional Laplace transforms in order to perform the convolution of bivariate densities.

The project time and cost distribution may thus be determined using the general equation,

$$Z = A_1 + A_2 + \ldots + A_n$$

where, $A_i = [C_i, T_i]$. Here A_i may be thought of as a random variable of two dimensions, representing the characteristics phase cost and phase time.

For simplicity we assume that there is no lower time value and all costs are time dependent. An appropriate joint density function is of the form,

$$f(c,t) = \lambda\gamma exp[-\gamma t - \lambda(c - kt)] \tag{4.11}$$

which has the Laplace transform,

$$\begin{aligned} \mathcal{L}[f(c,t)] &= \int_0^\infty \int_{kt}^\infty e^{-uc-vt} f(c,t) dc\, dt \\ &= \frac{\lambda\gamma}{(\lambda+u)(\gamma+ku+v)} \end{aligned} \tag{4.12}$$

Hence, for n random variables with densities of the above form, the density of the sum, Z, is given by the inversion of the following,

$$\prod_{i=1}^{n} \mathcal{L}[f_{C_iT_i}(c,t)] = \prod_{i=1}^{n} \frac{\lambda_i\gamma_i}{(\lambda_i+u)(\gamma_i+k_iu+v)} \tag{4.13}$$

A general formulation of the sum of n independent random variables with bivariate exponential distributions of the threshold dependency type can be obtained from inverting equation (4.13). We first tried invoking the Laplace transform inversion procedure of *Mathematica* with no success and thus studied alternative equation forms. A summary of the necessary steps to invert the product of n Laplace transforms is as follows: the inversion with respect to v is first determined,

$$\prod_{i=1}^{n} \frac{\lambda_i\gamma_i}{\lambda_i+u} \sum_{i=1}^{n} \prod_{j\neq i} \frac{1}{k_j-k_i} \frac{exp[-t\gamma_i - utk_i]}{\left(\frac{\gamma_j-\gamma_i}{k_j-k_i}+u\right)} \tag{4.14}$$

To invert with respect to u it is desirable that the transformed function is expressed in partial fractions. It can be seen that the number of u denominator terms is $2n-1$. Rearranging (4.14) gives,

$$Y\sum_{l=1}^{n}\prod_{\substack{m=1\\m\neq n+l}}^{2n}\frac{1}{(\lambda_{m,l}+u)(k_{m,l})}e^{-t\gamma_l-utk_l} \tag{4.15}$$

where, $Y=\prod_{i=1}^{n}\lambda_i\gamma_i$ and,

$$\begin{aligned}&\text{for } m=1,2,\ldots,n \text{ and } l=1,2,\ldots,n;\\&\quad k_{m,l}=1,\quad \lambda_{m,l}=\lambda_m\\&\text{for } m=n+1,n+2,\ldots,2n \text{ and } l=1,2,\ldots,n;\\&\quad k_{m,l}=k_{m-n}-k_l,\quad \lambda_{m,l}=\frac{\gamma_{m-n}-\gamma_l}{k_{m-n}-k_l}\end{aligned}$$

Inverting with respect to u is more involved and requires the use of a translation by first applying Heavisides expansion, see Abramowitz and Stegan (1972). Thus the p.d.f. of the convolution of n independent bivariate exponential densities, where the cost threshold is time dependent is given by,

$$f_Z=Y\sum_{l=1}^{n}\sum_{\substack{m=1\\m\neq n+l}}^{2n}\prod_{j\neq m}^{2n}\frac{e^{-t\gamma_l}e^{-\lambda_{m,l}(c-k_lt)}}{(\lambda_{j,l}-\lambda_{m,l})}\cdot\frac{1}{k_{j,l}}\cdot\frac{1}{k_{m,l}} \tag{4.16}$$

where,

$$\begin{aligned}c&>k_lt\\t&>0\end{aligned}$$

A necessary condition for the use of this inversion is that $\lambda_i\neq\lambda_j$, $\gamma_i\neq\gamma_j$, and $k_i\neq k_j$ for all $j\neq i$. Similar to the univariate case, modifications are needed when these conditions are not met.

From knowing the Laplace transform of a given density function the moment generating function can be easily determined simply from multiplying each dummy variable, (u and v), of the Laplace transform by -1, hence,

$$M_{CT}(u,v) = \frac{\lambda\gamma}{((\lambda-u)(\gamma-ku-v))} \tag{4.17}$$

4.3.1 Example of Risk Measures

To demonstrate the convolution of bivariate exponential densities and quantify possible risk measures of interest we consider a project comprising of two sequential phases. Suppose that the joint time-cost densities are denoted by $f_1(t_1, c_1; \lambda_1, \gamma_1, k_1)$ for phase 1 and $f_2(t_2, c_2; \lambda_2, \gamma_2, k_2)$ for phase 2. The joint p.d.f. of the project completion, f_Z is given by expanding equation(4.16) with $n = 2$. The result consists of six parts,

$$f_Z(t) = \prod_{i=1}^{2}\lambda_i\gamma_i \left[\frac{e^{-t\gamma_1}}{k_2-k_1}\cdot \left(\frac{e^{-\lambda_1(c-k_1t)}}{(\lambda_2-\lambda_1)(\frac{\gamma_2-\gamma_1}{k_2-k_1}-\lambda_1)} + \frac{e^{-\lambda_2(c-k_1t)}}{(\lambda_1-\lambda_2)(\frac{\gamma_2-\gamma_1}{k_2-k_1}-\lambda_2)} + \frac{e^{-\frac{\gamma_2-\gamma_1}{k_2-k_1}(c-k_1t)}}{(\lambda_1-\frac{\gamma_2-\gamma_1}{k_2-k_1})(\lambda_2-\frac{\gamma_2-\gamma_1}{k_2-k_1})}\right) + \frac{e^{-t\gamma_2}}{k_1-k_2}\cdot \left(\frac{e^{-\lambda_1(c-k_2t)}}{(\lambda_2-\lambda_1)(\frac{\gamma_1-\gamma_2}{k_1-k_2}-\lambda_1)} + \frac{e^{-\lambda_2(c-k_2t)}}{(\lambda_1-\lambda_2)(\frac{\gamma_1-\gamma_2}{k_1-k_2}-\lambda_2)} + \frac{e^{-\frac{\gamma_1-\gamma_2}{k_1-k_2}(c-k_2t)}}{(\lambda_1-\frac{\gamma_1-\gamma_2}{k_1-k_2})(\lambda_2-\frac{\gamma_1-\gamma_2}{k_1-k_2})}\right)\right] \tag{4.19}$$

where $c > k_1t$ for first 3 expressions and $c > k_2t$ for remaining expressions, $t > 0$.

To provide a graphical representation suppose the joint density for activity 1 has parameters, $\lambda_1 = 0.5$, $\gamma_1 = 0.5$, $k_1 = 2$, and the joint density for activity

2 has parameters, $\lambda_2 = 0.4$, $\gamma_2 = 0.4$, $k_2 = 3$. Thus the joint p.d.f of the time and cost for each activity are shown in Figure 4.5.

The convolution of the two joint densities is given by,

$$f(c,t) = \begin{cases} \frac{2}{15}e^{0.1c-0.7t} - \frac{8}{10}e^{-0.4c+0.3t} + \frac{2}{3}e^{-0.5c+0.5t} & \text{if } 2t < c \leq 3t \\ \frac{2}{3}(e^{-0.5c+0.5t} - e^{-0.5c+1.1t}) & \text{if } c > 3t \\ 0 & \text{elsewhere} \end{cases} \quad (4.20)$$

and is represented graphically in Figure 4.6.

4.4 Incorporating Quality

Based on our discussion of quality in section 1.7, we consider how quality may be related to the time spent in completing a given phase.

4.4.1 A Time-Quality Loss Relationship

We assume that the accumulation of quality loss throughout the project is phase dependent according to the nature of the phase activities and is predominately affected by the phase duration. To model possible types of phase quality loss accumulation we identify three intuitive time-quality loss relationships.

Relationship 1

Suppose that quality monotonically increases with time at a decreasing rate. In this case the level of quality loss is initially high and diminishes as time is expended. A feasible explanation is that more time ensures greater care is taken possibly in testing and achieving quality control measures. However as identified by Sears (1991) in recognition of the difficulty in finding and correcting small amounts of defects, we assume that the time taken to rectify deficiencies in quality increases as the level of quality loss reduces.

Relationship 2

On the other hand, suppose quality monotonically decreases at an increasing rate. Here the level of quality loss is initially low yet as more time is spent, quality loss increases. Such a relationship stresses the importance of completing the phases as soon as possible. Possible representations include, phase activities which are affected by external uncontrollable factors, such as the weather, which have severe effects on the outcome of phase which inevitably could affect the performance of the project outcome.

Relationship 3

Possibly the most realistic representation incorporates aspects of the above two relationships, in which quality is increased over a period to a given time point at which further efforts contribute to quality loss and thus reduce the level of quality.

Based on the intuitive relationships discussed above and adaptation of the mixed exponential distribution, see Christensen (1984), we propose a suitable time-quality loss function is given by,

$$ql = ae^{bt} + \alpha(1 - a)e^{-dt} \tag{4.21}$$

where $0 \leq a \leq 1$, and $b, d \geq 0$. The values b, d control the rate of growth and decrease in quality loss and α acts as a scaling factor in relationship one to allow a suitable range in quality loss levels. In the case of the third relationship, discussed above, we believe that it is unlikely that the rate of quality loss removal and the rate of quality loss increase, are identical, and therefore the variable a allows for various levels of asymmetry. When $a \neq 0.5$ asymmetry may be achieved. For $a = 0.5$ and assuming positive values for b and d where $b = d$, a symmetrical relationship is obtained. As indicated in Figure 4.7, from varying the values of a, b and d all three time-quality loss relationships may be modelled. The relationship may be simplified by replacing $\alpha(1 - a)$ with the parameter β.

4.5 Comments

The bivariate exponential extension distribution is a suitable function for modelling the relationship between both time and cost for independent phases in sequence. The distribution has intuitive appeal in that the chance of higher costs increases with the time realised. Also the threshold parameters enable minimum time and cost values to be incorporated. Further studies in using such a distribution include the examination of project networks where each activity has a bivariate exponential extension. To determine exact results, consideration to path dependency, and the development of suitable operators in the cumulation of both time and cost are needed.

Evaluating the performance of a project is commonly conducted with the aid of time management techniques which unfortunately do not address all areas of concern which may govern the success of a project. Similarly if cost of the project phases is solely modelled this also can limit the realistic decisions and actions taken during a project. For example, if it is likely that duration targets may be exceeded, more resource may be injected to offset, i.e. compete against, the event of completion delay. Where contracts are agreed to complete work in a predetermined period and agreed costs, the additional costs incurred to compete against the event of time delay may be passed on to the contractors. Even though both time and cost are important measures in monitoring the completion of the project they alone do not guarantee the success of the project outcome. Consider any project and assume that effort has been taken to ensure that it is completed on time to a required cost. It is a fallacy to assume that the project is a success merely because both time and cost constraints have been adhered to. Often, especially where there is high uncertainty in the project being performed the outcome may not perform to expectation. For example, projects which have ended in disaster include the Space Shuttle disaster of 1986, Bell and Esch (1989) and many bridge failures reported by Scott (1976). These are extreme cases in which shortfalls in reliability were experienced. It is interesting to note that investigations into the space shuttle disaster revealed that if a probabilistic risk analysis was performed instead of qualitative techniques, the mission would have likely been aborted. However in many cases probabilistic approaches are considered a hindrance as many projects are aborted unnecessarily due to pessimistic results. This gives rise to a 'catch 22' situation.

In light of the sensitivity of the data we found few examples where data is presented for such disasters. Other cases, less extreme yet still of concern include the deficiency of the project outcome. This deficiency may be in terms of the non-availability of the project outcome or possibly in terms of reduced output. Once again obtaining evidence of such measures is difficult in light of many companies policy to the confidentiality of data. However, there are many practical examples, for instance, it is well known that power plants are often less efficient than required resulting in potential financial losses. Problems lie not only in how the plant is operated or even the design of the plant, but rather the problems in the construction and testing stages of the plant. Possible reasons include the inability to quantify how well a project outcome is being achieved with the current concept of project performance and the inadequacy of project management techniques including computer software.

We now consider the use of proportional hazards modelling as a means of relating the reliability of the project outcome to the characteristics of the project phases.

4.6 Proportional Hazards Modelling

This technique has been widely applied in reliability studies and largely owes its origin to Cox (1972). Applied examples of interest are provided by Wightman and Bendell (1986) and Argent et al (1986). The essence of the basic proportional hazards model is that it attempts to identify the independent effects of various associated variables thought to influence the performance criteria. In reliability studies the performance criteria is typically the life length of a component. In the context of our formulation the associated variables may be the time, cost and quality loss experienced during the project phases, denoted by the vectors $\underline{t}$, $\underline{c}$ and $\underline{ql}$.

The proportional hazards model is structured upon the familiar hazard function, equation (1.9). A typical use of this technique involves the analysis of system components, where it is assumed for each component the associated hazard function can be decomposed into a baseline hazard function and a function dependent on the covariates. The covariate function suggested by Cox is an exponential term incorporating the effect of the values of the covariates. Thus the decomposition is typically written as,

$$h(t; z_1, z_2, \dots, z_k) = h_0(t) exp(\beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k z_k) \tag{4.22}$$

where the β_i's are the unknown parameters of the model defining the effects of each of the covariates, z_i. The covariates are assumed to act multiplicatively on the base-line hazard function, that is for different significant covariate values the hazard functions are proportional to each other over all time t. The β_i parameters are typically required to be estimated and tested to see whether each covariate has an effect on the hazard function. The baseline hazard function $h_0(t)$ represents the hazard function associated when the covariates take the baseline value zero. Either a parametric form may be taken for $h_0(t)$, such as the Weibull hazard or a distribution free approach may be used in which no particular form is assumed.

The effect on the reliability function from modelling the effects of the covariates is a power one given by,

$$R(t; z_1, z_2, \ldots, z_k) = [R_0(t)]^{exp(\beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_k z_k)} \tag{4.23}$$

where,

$$R_0(t) = exp\left[-\int_0^t h_0(x)dx\right]$$

4.6.1 Use of the Proportional Hazards Model

Other than stating the covariates relevant to our formulation at this stage we recognise further work is possible in collecting data for estimating model parameters. However to demonstrate the effects of the covariates we assume values for the β parameters and a two parameter Weibull baseline hazard function given by,

$$h_0(t) = \frac{\alpha}{\eta^\alpha} t^{\alpha - 1}$$

where α is the shape parameter and η is the scale parameter. We suppose the case where the $\alpha > 1$ indicating an increasing hazard rate with time. This may apply to the failures associated with buildings. For example as investigated by Scott (1976) many buildings constructed with modern techniques since the war are prone to failure after a period of ten to fifteen years. Common causes of failures include the use of unskilled labour during the project phases, design faults and lack of sufficient maintenance. If we suppose that maintenance is not performed then design faults and the consequences of unskilled labour in terms of the quality loss may be modelled. In such cases the cause of design faults, may be attributed to spending insufficient effort, in terms of time and money in the project phases and model the covariates $\underline{t}$ and $\underline{c}$. Alternatively where quality loss may easily be quantified, the hazard function may be defined in terms of the quality loss experienced over project phases. Where quality loss is additive over project phases, the total quality loss experienced may be modelled as a covariate.

Consider a project of three phases, design, construction and operation of a building. Suppose that the baseline hazard function is known with parameters $\alpha = 3$ and $\eta = 10$ and the significant covariates are ql_1 and ql_2 denoting the quality loss experienced in design and construction respectively and the covariate parameters are $\beta_1 = 0.2$ and $\beta_2 = 0.1$. From a change in the levels of quality loss experienced in the project phases we may analyse the effects on the building reliability. Using the uniform quality loss dependency model, to be covered in section 5.6 and modelling the covariate data with expected quality loss values a hazard function may be determined together with an associated reliability function. If $\mu_{ql_1} = 5$ and $\mu_{ql_2} = 20$, we may use the related function, (4.23), to determine the associated reliability function. Suppose this scenario is the worst case considered with regards to quality loss and the best case scenario is when $\mu_{ql_1} = 4$ and $\mu_{ql_2} = 12$. If we also consider the baseline reliability function, assuming no loss in quality over the project phases we may illustrate the reliability profile of the building in terms of the optimistic, pessimistic and most desirable levels of quality loss experienced during the project phases, see Figure 4.8.

The purpose of the book by Scott (1976) was to provoke the necessary actions to address construction failures cased by common causes. The problem that he encountered in practice was providing evidence to support his claims. We believe, where suitable data is available the reliability profile of given building types may be determined from modelling attributes of the project phases. In addition other factors such as weather conditions, and the morale of employees may be incorporated into the model to test for significance. As revealed in Chapter 1, there is a great concern within the construction industry in relating the outcome of a project, such as a building or chemical plant, to how the project was performed. We have illustrated a possible technique which is suited to the requirements of the analyst yet field data is essential to model efficiently. Further work is thus possible if a suitable database is available regarding performance criteria during the project phases and operation data.

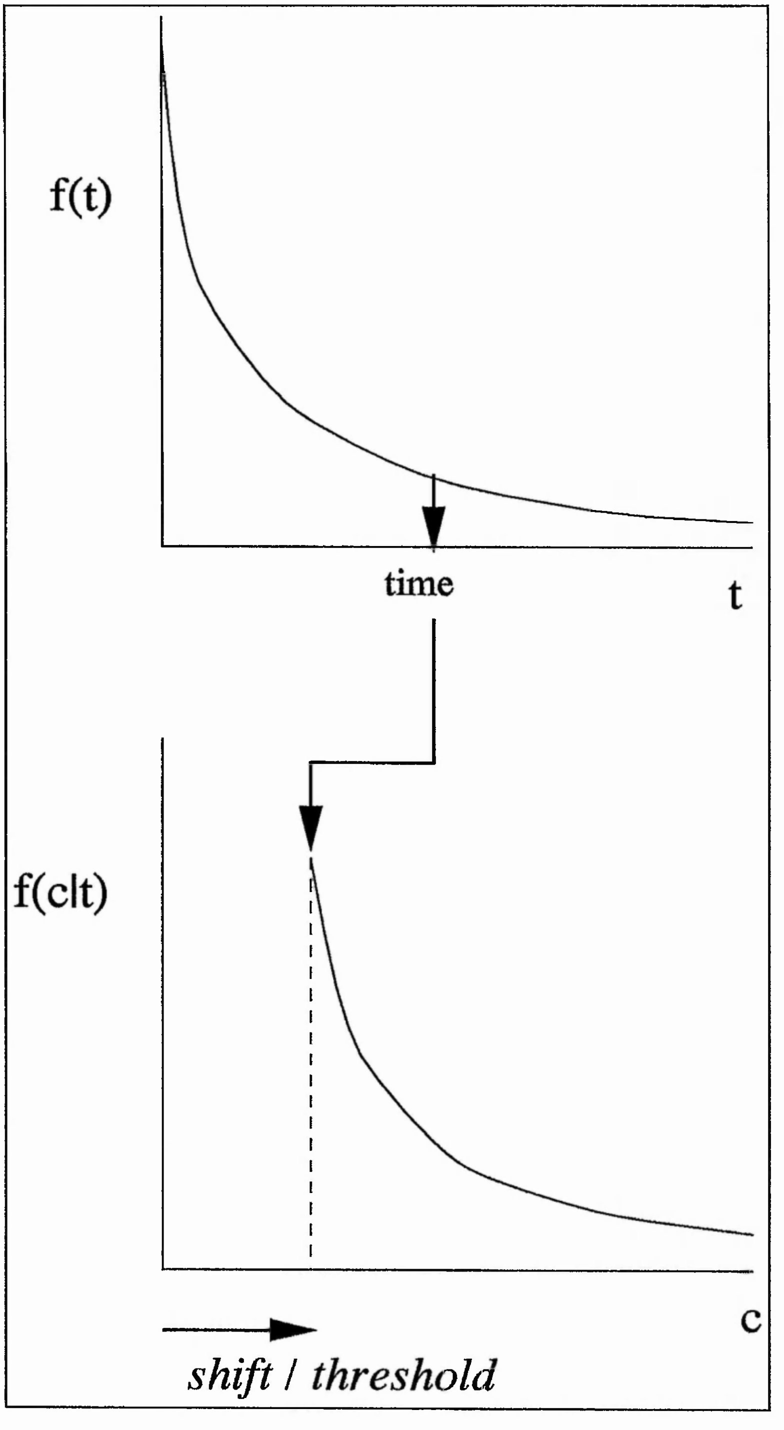


Figure 4.1: Time-Cost Relationship

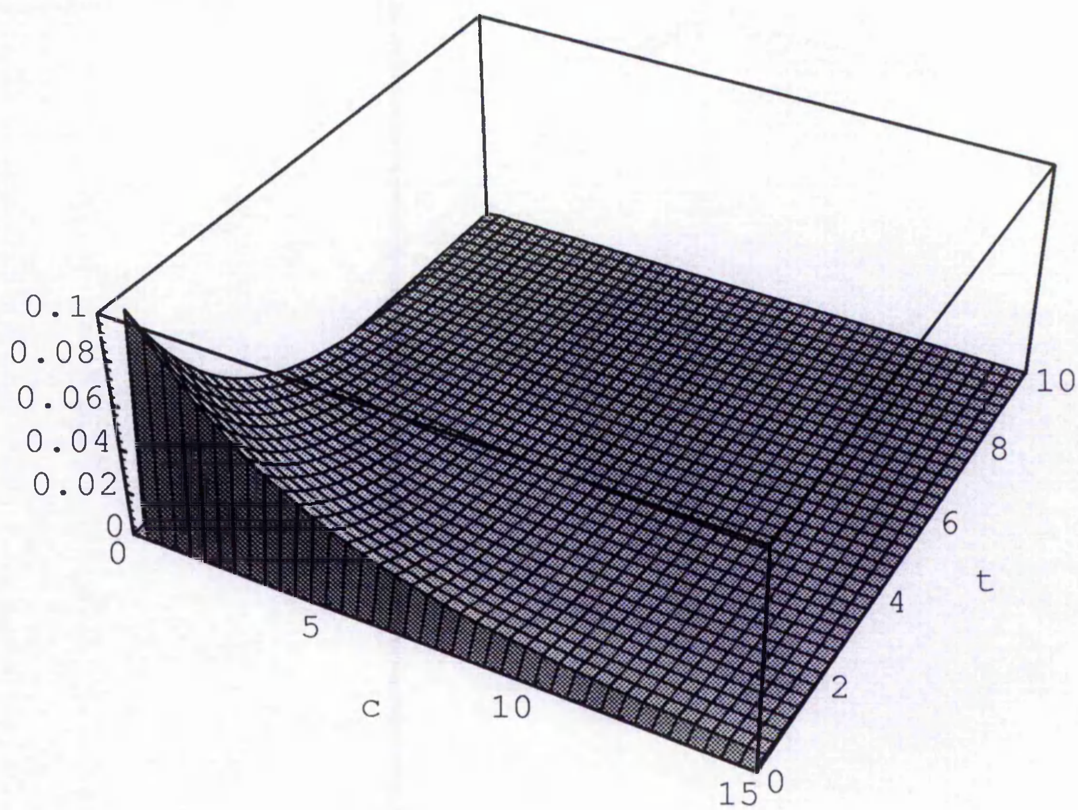
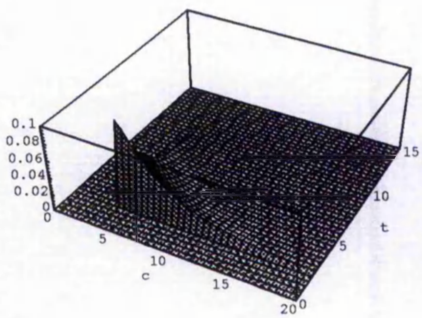
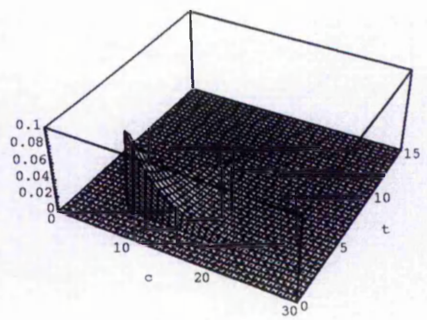


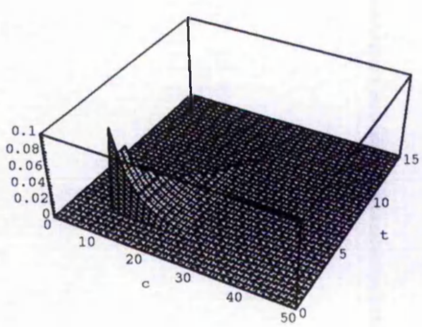
Figure 4.2: Time-Cost Dependency Plot; no Dependency, no Thresholds, $f(c,t;\lambda=0.25,\gamma=0.5,\alpha=0,k=0,b=0)$



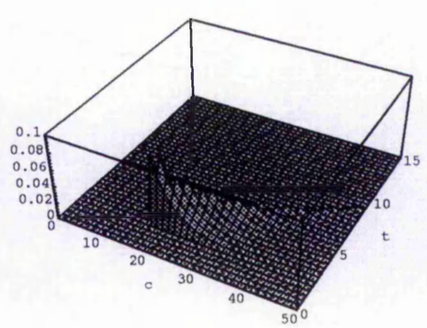
(a) $\alpha = 4, b = 2, k = 0$



(b) $\alpha = 4, b = 2, k = 1$



(c) $\alpha = 4, b = 2, k = 2$



(d) $\alpha = 4, b = 2, k = 5$

Figure 4.3: Time-Cost Probability Density Plots

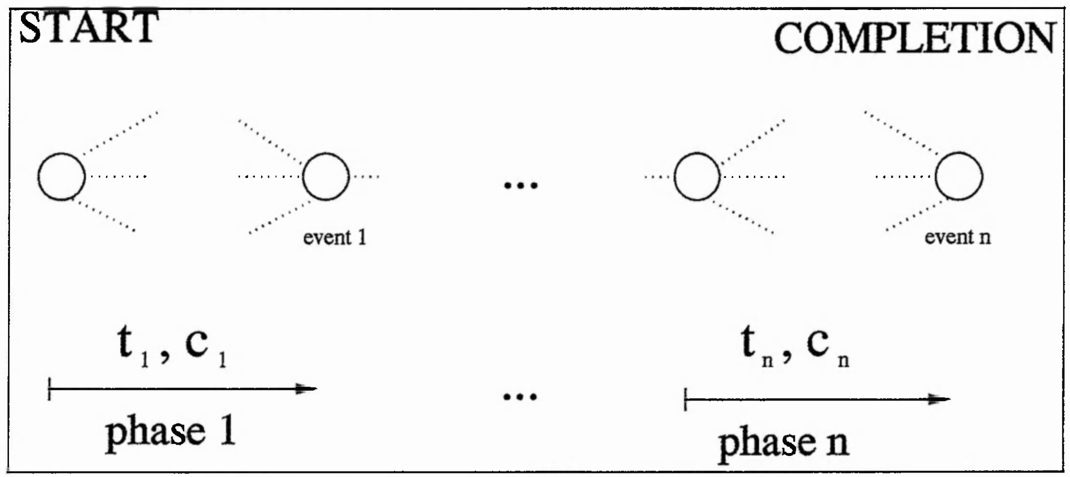
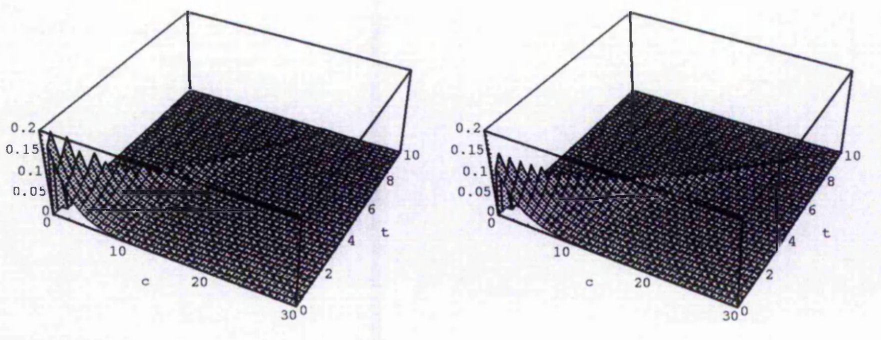


Figure 4.4: Project phases



(a) $f(c_1, t_1; \lambda_1 = 0.5, \gamma_1 = 0.5, k_1 = 2)$

(b) $f(c_2, t_2; \lambda_2 = 0.4, \gamma_2 = 0.4, k_2 = 3)$

Figure 4.5: Bivariate Activity Densities

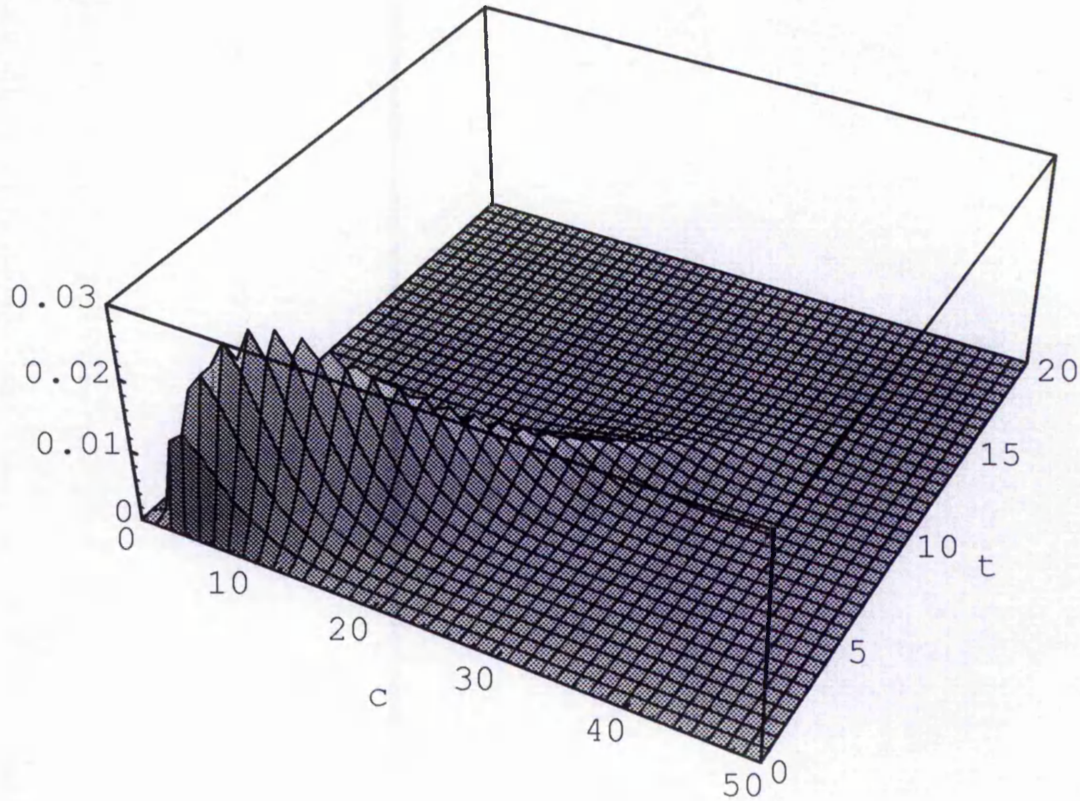
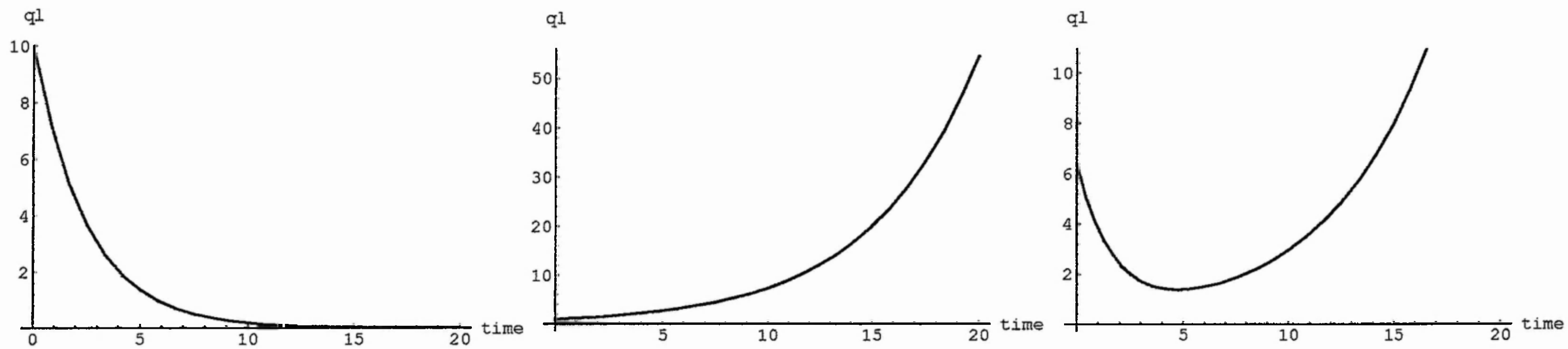


Figure 4.6: Joint Density Function of the Convolution of two Bivariate Densities



(a) $a = 0, b = 0, d = 0.4, \alpha = 10$

(b) $a = 1, b = 0.2, d = 0, \alpha = 0$

(c) $a = 0.4, b = 0.2, d = 0.6, \alpha = 10$

Figure 4.7: Time-Quality Loss Dependency Plots

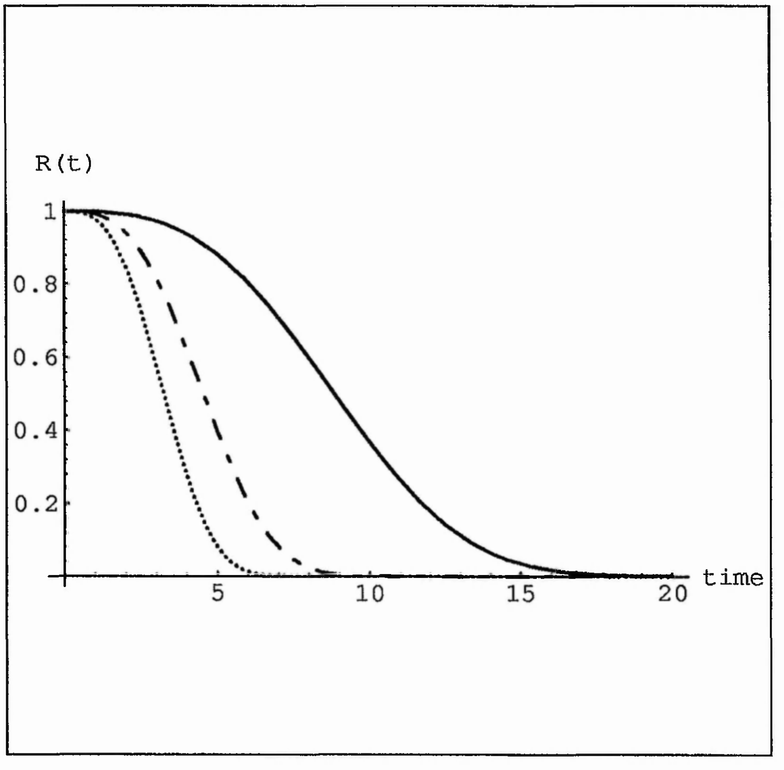


Figure 4.8: Building Reliability for Optimistic, Pessimistic and Desirable levels of Quality Loss levels during Project Phases; ....Worse Case,.-.-.-.-.- Optimistic, ________Desirable

Chapter 5

Generalised Stochastic Model

5.1 Introduction

In previous chapters we treated the time to complete project phases as random variables with defined p.d.f.'s. In addition we developed a time-cost relationship where time is considered the driving characteristic which influences the cost. We now consider the dependency between project phases as outlined in our formulation of Chapter 1 and provide a further model where operation measures are dependent on phase characteristics.

Based on our formulation we recognise two classes of model of particular interest. The first captures the stochastic nature of the project phases, where transition rates between project phases may be explored. With this type we discuss variations of the use of the Markov property. The Markov property has been explored previously by Kulkarni and Adlakha (1986) where the state space of a project network is discrete and it is assumed each activity may be either active, dormant or idle. Calculations are simplified by assuming exponential activity densities, however a limitation of the approach is that the state space can grow rapidly with network size and thus generally requires the use of a computer to store state values.

The second model class involves relating performance characteristics during the project phases to operation phase measures, such as the time to operational failure. To formulate a suitable model for the second type we draw upon studies and methods available to the reliability analyst and use the partial operation reliability model of Bendell and Humble (1985) to model the depedency between operation reliability and project quality.

In order to provide a suitable model structure for examination of the project phases we first examine properties common to stochastic processes.

5.2 Preliminaries

5.2.1 Stochastic Process

Stochastic processes are commonly used to investigate phenomena that are essentially concerned with the flow of events in time. For example birth and death, growth and decay, change and transformation can all be modelled as a stochastic process. Many of the stochastic processes are model dependent and require careful formulation. The type of stochastic process selected is based on the characteristics of the model. As an introduction to developing a stochastic process specific to the modelling of the performance of a project we give a broad review of the more general definitions. For a fuller account of the theory of stochastic processes see Cox and Miller (1965), and Feller (1966). Applied examples of stochastic processes in the context of renewal theory are available in Cox (1962) and applied biological processes covering many of the theoretical models are presented in Bailey (1963).

Definitions and Ideas

Since a process evolves over time a set of 'time' points is defined. These may be either a set of integers, $(n = 0, 1, 2 \ldots)$, called discrete time, or an interval $(-\infty < t < \infty)$, called continuous time. It is convention to have a collection of random variables $\{X_n\}$ or $\{X(t)\}$ defined for time points n (discrete time) or t (continuous time). The state space is defined as the set of possible values of an individual X_n or $X(t)$. It may be either discrete or continuous and be multidimensional.

A special class of process called a 'Markov process' enables the transition between states to be dependent only on the current state. Therefore given arbitrary times $\ldots < t_{n-2} < t_{n-1} < t_n$,

$$P(X_{t_n} = x | X_{t_{n-1}} = y, X_{t_{n-2}} = z, \ldots) = P(X_{t_n} = x | X_{t_{n-1}} = y).$$

In order to determine the conditional p.d.f.'s over the entire time interval the Chapman-Kolmogorov equation is used,

$$P(X_{t_n} = x | X_{t_{n-2}} = z) = \int_{-\infty}^{\infty} P(X_{t_{n-1}} = y | X_{t_{n-2}} = z)$$
$$P(X_{t_n} = x | X_{t_{n-1}} = y)\, dy.$$

A corresponding definition in continuous time, where $t_{n-2} < t_{n-1} < t_n$ is,

$$P(X(t) = x | X(t_{n-2}) = z) = \int_{-\infty}^{\infty} P(X(t_{n-1}) = y | X(t_{n-2}) = z)$$
$$P(X(t_n) = x | X(t_{n-1}) = y))\, dy$$

These enable the conditional p.d.f's to be constructed over the 'long' time interval (t_{n-2}, t_n), from those over the short time intervals (t_{n-2}, t_{n-1}) and (t_{n-1}, t_n). Due to the complex nature of possible stochastic processes many of the models available make assumptions with regards to the transition probability mechanism. In most cases in order to obtain results of interest the transition probabilities are assumed to be independent of time. To incorporate a change in the transition probability structure possibilities include setting the transition probability as a function of the current state, which is typical of birth and death processes, Bailey (1964). Non-homogenous processes that are mathematically tractable include the Non-Homogeneous Poisson Process (NHPP) and the Polya Process, however both such models are not suited to our formulation.

Discussion

There are numerous possibilities to how our project phase formulation may be adapted as a finite stochastic process which are dependent on how we define the state space and the nature of time. One such possibility is if we assume X_i is a three dimensional continuous state space for time, cost, and quality loss and the process time is indexed on the completion of phase i. For a continuous time process the state space may consist of two continuous dimensions for cost and quality loss and a discrete dimension for project phase. Alternatively we may consider the project evolving as time, cost, and quality loss accumulate thus modelling the state space as the project phase. For simplicity we consider two cases. The first case we assume time is continuous and formulate a stochastic process where project progress is dependent on time, however similar to the time-cost model of Chapter 4, the characteristics cost and quality loss may be determined from the state of the project at a given time. The second case we assume time is discrete representing the number of phases performed and discuss intuitive dependencies affecting the phase characteristics time, cost and quality loss.

5.3 Formulation with Continuous Time

We now turn to a general formulation where methods based on transitions in a small time interval may be applied. We define $S(t)$ as the stochastic process indicating the number, s of project phases completed at time $t \geq 0$. We assume that the random variable $S(t)$ is discrete, so that the state space is defined as the integer values in the range $[0, n]$ where n is the number of phases to be performed in series. We define $S(t) = 0$ to be the achievement of no phases at time t, and $S(t) = n$ to be the completion of the project at time t. Once the project is complete no further transitions are possible and therefore n is defined as an absorbing state. The process thus has a probability distribution such that,

$$p_s(t) \geq 0 \text{ for s in the range } [0, n]$$

and

$$\sum_{s=0}^{n} p_s(t) = 1, \quad \text{for all t}$$

Using this formulation we may structure a model where in general the discrete parts of the distribution may be identified, so that,

$$\sum_{s=1}^{n-1} p_s(t) + p_0(t) + p_n(t) = 1, \text{ for all t}$$

To incorporate a Markov property, such that the instantaneous transition rates at t depend on progress made in phase s at t and not on states previously passed through, we may augment the definition of the states. We introduce the continuous variable τ_s to represent the time spent in state s, as indicated in Figure 5.2 which corresponds to time spent to complete phase $s + 1$. Now the state space is partly discrete and partly continuous.

We may specify the system as being in state (s, τ_s), $\forall 0 \leq s < n$, or in state n. State (s, τ_s) may be interpreted as having accomplished the first s phases and have spent time τ_s performing the successive phase without completion. For convenience let $p_s(\tau_s; t)$ denote the probability associated with the states that include both discrete and continuous parts. Our general model is modified to,

$$\sum_{s=0}^{n-1} \int_0^t p_s(\tau_s, t) d\tau_i + p_n(t) = 1, \text{ for all t}$$

We define $\lambda_s(t)$ as the transition rate probability function from state $(s, t - \sum_{j=0}^{s-1} \tau_j)$ to state $(s+1, 0)$ at time t. Alternatively without loss of meaning we may define the transition rate probability function as $\lambda_s(\tau_s)$ from state (s, τ_s) to state $(s+1, 0)$ at time τ_s in state s. Hence $\lambda_s(\tau_s)\Delta t + o(\Delta t)$ is the instantaneous probability of a transition to state $(s+1, 0)$ at the end of the infinitesimal interval $\tau_s + \Delta t$ given that phase $s+1$ is being performed at time t. Thus in summary our transition rate probability functions are,

$$\lambda_0(t) = \lim_{\Delta t \to 0} \frac{P[S(t+\Delta t) = (1,0) | S(t) = 0]}{\Delta t}$$
$$\lambda_s(t) = \lim_{\Delta t \to 0} \frac{P[S(t+\Delta t) = (s+1,0) | S(t) = (s, \tau_s)]}{\Delta t}, \quad 1 \leq s < n$$

A reasonable initial condition for the system of equations is,

$$\begin{aligned} p_0(0) &= 1 \\ p_n(0) &= 0 \\ p_s(\tau_s, 0) &= 0 \quad 0 < s < n \end{aligned} \tag{5.1}$$

Since state $s = n$ is an absorbing state this implies that an absorptive sink with probability collected into a discrete atom will develop at $s = n$. Consequently one would hope that by time infinity, all phases will be complete and the distribution will be concentrated onto the state of project completion, so that

$$\lim_{t\to\infty} p(s,t) = 0, \quad 0 < s < n$$
$$\lim_{t\to\infty} P_0(t) = 0, \tag{5.2}$$
$$\lim_{t\to\infty} P_n(t) = 1.$$

The Chapman Kolmogorov differential equations for the system are

$$\begin{aligned} \frac{dP_0}{dt} &= -P_0(t)\lambda_1(t) \\ \frac{dP_n}{dt} &= \int_0^t p_{n-1}(\tau_{n-1};t)\lambda_{n-1}(\tau_{n-1})d\tau_{n-1} \\ \frac{\partial p_s(\tau_s;t)}{\partial \tau_s} + \frac{\partial p_s(\tau_s;t)}{\partial t} &= \lambda(\tau_s)p_s(\tau_s;t) \end{aligned} \tag{5.3}$$

where a necessary condition is $p_s(0;t) = p_{s-1}\lambda_{s-1}(\tau_{s-1})$.

Further work using this formulation includes the development of project measures. Such measures include the following. Assessment of which phase is being performed at time, t, is possible from computing $p(s,t)$. With specified duration target values, assessment of meeting target completion given time τ_s has been spent in the current phase s and time t has elapsed from project start is given by $p(meet\,target|t,\tau_s,s)$. The expected time to phase i and the expected time from phase i to phase j are further measures. By incorporating the time-cost and time-quality loss relationships of Chapter 4, measures regarding all project characteristics may be obtained.

5.4 Formulation with Discrete Time

Whether a particular system leads to a Markov process depends on how the random variables specifying the stochastic process are defined. Other than using the Markov property to determine the probability of leaving a given phase, it is possible to define conditional type dependencies between phases. We shall refer to this type of dependency as inter-dependency which we shall use throughout this chapter and discuss further in Chapter 6.

The evolution of a project requires the consumption of resources. These resources have constraints and targets which can determine the success of a project. For convenience we assume that the use of resource has a direct influence of the characteristics time, cost, and quality loss of the project, $\underline{t}$, $\underline{c}$, and $\underline{ql}$.

As in section 1.7.5, and modelled in Chapter 4, in general it is unlikely that all three characteristics are independent within a phase. In addition in general the characteristics of a given phase may affect any or all of the characteristics of future phases. Such a dependency may not be modelled with a Markovian model as a 'memory' property is needed. We explore cases where the 'memory' of previous states is needed with the aid of simulation in Chapter 6.

In order to employ the Markov property, we assume a Markov chain with a continuous state space and suppose that dependency may only exist between successive phases. The discrete time points in this instance represent the completion of a given phase. Thus for adjacent phases indexed by i and $i-1$, we may incorporate the discrete time Markov property for each characteristic, where the characteristic state transition probability is defined by,

$$P(T_i = t, C_i = c, QL_i = ql | T_{i-1} = t_{i-1}, C_{i-1} = c_{i-1}, QL_{i-1} = c_{i-1}) \tag{5.4}$$

In this case we assume that the phase characteristics are independent within a phase. Thus the amount of time spent completing phase $i-1$ may only affect the time spent in phase i and have no influence on any other characteristic. Similar relationships also exist with the characteristics cost and quality loss. To retain realism associated with a project we would expect the transitions to be non-homogeneous, thus the transition probability between different adjacent phases will differ.

Incorporating dependency within phase is more involved as numerous dependencies may exist between characteristics of a given phase in addition to dependencies between characteristics of adjacent phases. A possible formulation, which is based on the assumption applied in Chapter 4 where the characteristic time is considered the most influential characteristic is to suppose that,

$$P(T_i = t | T_{i-1} = t_{i-1}, C_{i-1} = c_{i-1}, QL_{i-1} = ql_{i-1}) \tag{5.5}$$

and,

$$P(C_i = c | T_i = t)$$
$$P(QL_i = ql | T_i = t, C_i = c)$$

Once again we explore this type of dependency further in section 6.3 of Chapter 6. Where specific measures of concern are of importance we may simplify the dependency. For example as developed in section 2.8 of Chapter 2 we considered a parametric model for the escalation in quality loss where dependency existed between phases. If a particular phase is performed badly it may affect performance of a successive phase. To capture this type of dependency we may model the inter-phase dependency for one characteristic, say quality loss.

Another consideration is the introduction of targets. Targets are of great importance. From comparing the actual progress made within a project to targets enables assessment of how the project is evolving. From an understanding of the dependencies throughout the project, improved decision making is possible at any given phase as future consequences may be assessed.

Based on our formulation of Chapter 1 and the above considerations we state a mechanism for incorporating phase dependency.

5.5 Transition Mechanism for Phase Dependency

We define S_j as random variable of three dimensions indicating the levels of cost, time, and quality loss at phase j and Z_k, also as a random variable of three dimensions, associated to S_j by $Z_k = \sum_{i=1}^{k} S_i$. For ease of reference we introduce the following vector notation. For a given phase i we shall use the vector $\underline{s}_i$ to refer to the realisations (t_i, c_i, ql_i). Similarly we shall use the vector $\underline{z}_k$ to denote, $(\sum_{i=1}^{k} t_i, \sum_{i=1}^{k} c_i, \sum_{i=1}^{k} ql_i)$.

The dependency between phases may be incorporated by the transition to project phase j being dependent on one or a combination of appropriate measures of time, cost and quality loss to phase $j-1$. Features of interest include examination of the transition dependency.

Cumulative Dependency

Suppose the performance of the current phase depends upon the total time and money that has already been spent together with the total amount of quality loss experienced. Thus,

$$P(S_j = \underline{s}j | \underline{z}_{j-1}) \tag{5.6}$$

Inter-Phase Dependency

For particular phases the level of performance may only depend on how the previous phase was conducted and not on the escalation of all components. Thus the Markov dependency may be of the form,

$$f(S_j = \underline{s}_j | S_{j-1} = \underline{s}_{j-1}) \tag{5.7}$$

5.5.1 Implementation of Transition Mechanism

The mechanism for transition dependency is intuitive and appealing as it reflects the type of dependency experienced between project phases. By incorporating target values a useful project management tool is obtained, however to provide a means of efficiently modelling such dependency we turn to simulation to generate results of interest, see Chapter 6. To illustrate the difficulty in modelling phase dependency we develop a probabilistic univariate quality loss model.

5.6 A Quality Loss Phase Dependency Model

We now develop a model which illustrates the stochastic nature of quality loss over the project phases. In what follows, j subscripts on QL and $f(ql)$, denote the random variable quality loss for phase j and the p.d.f of the quality loss of phase j respectively. For convenience we let $Z_j = \sum_{i=1}^{j} QL_i$ and $z_j=\sum_{i=1}^{j} ql_i$.

In order to demonstrate a possible model, based on the ease of generating and inverting Laplace transforms we first assumed that for a given phase i, the quality loss random variable follows an exponential distribution given by,

$$h(ql_i) = \gamma_i \, e^{-\gamma_i \, ql_i}$$

Thus, the p.d.f of the total quality loss experienced in n phases, f_{Z_n} is given by,

$$f_{Z_n} = \mathcal{L}^{-1}\{\prod_{i=1}^{n} \mathcal{L}\{h(ql_i)\}\}$$

We assume that the quality loss experienced in phase i depends on all the previous 'wrong doings' up to and including phase $i-1$. Thus a suitable function determining the performance of the current phase is using a conditional density function on the realisation of all quality loss previously experienced,

$$g(ql_i|z_{i-1}) = \gamma_i(z_{i-1})\, e^{-\gamma_i(z_{i-1})\, ql_i} \tag{5.8}$$

In addition, we build upon the covariate structure and introduce other contributing factors such as planned quality, funds available, and time allowed denoted by $\overline{q}$, $\overline{c}$ and $\overline{t}$ respectively. We assume that these contributing factors are known and therefore are constant. The time available to perform a given task may contribute differently to the quality loss experienced. When quality loss is proportional to the amount of time available, as discussed in section 4.4, a suitable function for the γ_i parameter may be defined as,

$$\gamma_i(z_{i-1}, \overline{q}_i, \overline{c}_i, \overline{t}_i) = \frac{\overline{q}_i \overline{c}_i \eta_i}{z_{i-1} ql_i \overline{t}_i} \tag{5.9}$$

where η_i is a suitable scaling factor.

Here the mean quality loss, μ_{ql_i} tends to zero as either $\overline{c}_i$ or $\overline{q}_i$ tend to infinity assuming $\overline{t}_i$ remains constant. Similarly as z_{i-1} tends to infinity, μ_{ql_i} also tends to infinity. Alternatively where quality loss is inversely proportional to the amount of time available, also discussed in section 4.4, the γ_i parameter may be defined as,

$$\gamma_i(z_{i-1} ql_j, \overline{q}_i, \overline{c}_i, \overline{t}_i) = \frac{\overline{q}_i \overline{c}_i \overline{t}_i \eta_i}{z_{i-1} ql_j} \tag{5.10}$$

It follows that the quality loss density function during phase i is given by integrating over the total quality loss values of previous phases denoted by z_{i-1}. Thus,

$$h(ql_i) = \int_0^{\infty} g(ql_i|z_{i-1})\, dz_{i-1} \tag{5.11}$$

5.6.1 Example

Suppose that the quality loss density function for phase one is defined as,

$$h(ql_1) = \gamma_1 \, e^{-\gamma_1 ql_1} \tag{5.12}$$

Since this is the initial phase, $h(ql_1)$ is also the p.d.f for QL_1, therefore the quality loss density function for phase two is derived as follows,

$$h(ql_2) = \int_0^\infty g(ql_2|ql_1) \, dql_1 \tag{5.13}$$

Letting η_2 incorporate all constant contributing values and performing the integration gives,

$$h(ql_2) = 2\gamma_1\eta_2 k_0(2\sqrt{\gamma_1}\sqrt{\eta_2 ql_2}) \tag{5.14}$$

where $k_0(z)$ is a modified Bessel function of the second kind. To obtain $h(ql_3)$ we must perform the convolution of $h(ql_1)$ and $h(ql_2)$. When convoluted the result includes a Laguerre Polynomial which cannot be simplified to a useable form. In conclusion, a closed form equation for the quality loss at phase $i \geq 2$ is involved and therefore considered too complex to develop as a realistic management aid. Even though the exponential distribution was effective when determining the dependency within a project phase, as shown in section 4.2 of Chapter 4, the approach is not tractable with phase dependency. Even using an exponential quality density function on the range $[0, 1]$ causes computational problems. The probability density for phase two quality contains a Kummer confluent hypergeometric function which also cannot be simplified.

Resolution of computational difficulties may be overcome by adopting a different density function where phase dependency may be incorporated. A suggested simplification is to model the relationships with a uniform p.d.f.

5.6.2 Univariate Dependency using the Uniform Distribution

Once again, to demonstrate the computation steps we shall determine the quality loss density function at any phase i. The function will always be of shape illustrated in Figure 5.1 and the assumptions include,

- for any phase there is always a chance of zero quality loss
- the range of the quality loss distribution is dependent on both the previous phase quality loss value, ql_{i-1} and the factors $\overline{t}$, $\overline{c}$, and $\overline{ql}$
- each quality loss value in the range is equally likely of occurring.

A suitable dependency formulation for the uniform parameter b_i is,

$$b_i = \frac{\alpha_i(ql_{i-1} + 1)}{\overline{c_i}\overline{t_i}} \tag{5.15}$$

where α_i is a suitable scaling factor and the addition of one is incorporated to overcome the possibility of the numerator being zero. If no dependency exists the uncertainty in the level of quality loss is dependent on both the time and money available.

In all cases a necessary condition is that the uniform probability value, k_i, for the quality loss p.d.f is given by,

$$k_i = \frac{1}{b_i} \tag{5.16}$$

Thus letting η_i represent all constant factors at phase i, implies the quality loss at phase i is dependent on the previous phase, $i - 1$, by the conditional density function,

$$f(ql_i|ql_{i-1}) = \frac{\eta_i}{ql_{i-1}+1} \tag{5.17}$$

The quality density function at phase i is given by,

$$h(ql_i) = k_1 \prod_{j=2}^{i} \eta_i \ln[1 + \frac{1}{k_{i-1}}] \tag{5.18}$$

The uncertainty associated with the total quality loss experienced to the completion of phase n, denoted as $h(z)$, where $Z = QL_1 + QL_2 + \cdots + QL_n$ can be determined by convoluting the quality loss density functions from start to phase n. To facilitate the determination of this result we may invert the product of the Laplace transform associated with each quality loss density function. The Laplace transform associated with $h(ql_i)$ is,

$$\mathcal{L}(h(ql_i)) = \int_0^{\frac{1}{k_i}} e^{-s\,ql_i} k_1 \prod_{j=2}^{i} \ln[1 + \frac{1}{k_{j-1}}]\, dql_i \tag{5.19}$$

Thus the Laplace transform of n phases is given by,

$$\mathcal{L}\{h(z)\} = \frac{1}{s^n} \prod_{i=1}^{n} \frac{1 - e^{-b_i s}}{b_i} \tag{5.20}$$

Inverting gives,

$$h(ql) = \frac{1}{(n-1)!\prod_{i=1}^{n} b_i}(ql^{n-1} + \sum_{(i/n)} (ql + (-1)^i\, b_{(i/n)})^{n-1}|ql - b_{(i/n)}|) \tag{5.21}$$

where $\sum_{(i/n)}$ is the sum over all subsets of size i from n_j, $i = 1, 2, \ldots, n$ and $b_{(i/n)}$ the corresponding product of the b_i's of each subset.

Further work includes evaluation of the quality loss dependency model with real project data. A possible application is in modelling software reliability where dependency exists between phases of the software development.

5.7 Formulation of a Project Outcome Performance Measure

Suppose that the performance of the project outcome is not dichotomic in the sense of either performing or not performing, but various levels of performance are achievable. In this case we may use the partial operation reliability model of Bendell and Humble (1985), and incorporating a dependency between how well the project was conducted. For completeness we shall provide formulation of the transition probabilities and a relevant Kolmogorov-Chapman forward equation in order to specify differential equations for solution.

We define $S(t)$ as the stochastic process indicating the level of performance of the PO at time t in use of the operation phase. For convenience we assume $S(t)$ takes values in the interval $[0, 1]$ and define 0 to represent the case when the PO has failed in some sense resulting in a zero level of operation, 1 to be the full operation of the PO and values in-between representing the proportion of operation achieved. Assuming $S(t)$ is continuous with probability density say $p(s, t)$ on $0 < s < 1$ we may represent the distribution as a composite of discrete and continuous parts where atoms of probability, $P_0(t)$ and $P_1(t)$ are used for the two extreme levels in operation. Thus,

$$\begin{aligned}
&\int_{0+}^{1-} p(s,t)ds + P_0(t) + P_1(t) = 1, \quad \text{for all } t \\
&p(s,t) \geq 0 \\
&P_0(t) \geq 0 \\
&P_1(t) \geq 0
\end{aligned} \tag{5.22}$$

If the transitions between states of PO performance satisfy the first order Markov property that the instantaneous rates at t depend only on the state at t and to some extent the total quality loss experienced during the project phases, but not on previous states of performance, we may define $\phi(s, r, t, ql)$ as the transition rate probability function from state s to state r at time t $(r \neq s)$. For

$0 < r < 1$ and $r \neq s$, $\phi(s, r, t, ql)$ is the instantaneous probability of a transition to states $r + \delta r$ at the end of the infinitesimal interval t to $t + \delta t$ given that the project outcome performance is in state s at t and the quality loss level prior to operation is ql. The non-negative random variable QL with probability density f_{QL} provides a measure of the quality loss experienced during the project phases. This may be obtained by various approaches previously discussed. Possibilities include the use of the univariate approach of Chapter 3 where the quality loss of each phases is independent and follows an Erlangian distribution. Alternatively where time and cost for each phase are modelled by a bivariate exponential extension we may determine a distribution of quality loss from using relationship (4.21) of Chapter 4. Where phase dependency exists the uniform distribution may be used as indicated in section 5.6.

As applied in the proportional hazards model, it is the quality loss measure that enables the dependency between the operation phase and how well the project is performed during the project phases.

We assume that QL at $t = 0$ is the total quality loss experienced immediately prior to PO use. In most cases this will be at the time point of project completion however this may not always be true. If delays are incurred in putting the PO to use possible changes in the quality loss may be experienced such as degradation. On the other hand initial testing of the PO prior to use may identify shortfalls and allow for improvements, thus reducing the level of quality loss of the PO. Such issues will not be explored in our discussion. In summary the limiting probabilities of the transition rate function are,

$$\begin{aligned} \phi(s, r, t, ql)\delta r &= \lim_{\delta t \to 0} \frac{P[r \leq S(t + \delta t) \leq r + \delta r | S(t) = s, QL = ql]}{\delta t}, \quad 0 < r < 1, r \neq s \\ \phi(s, 1, t, ql) &= \lim_{\delta t \to 0} \frac{P[S(t + \delta t = 1 | S(t) = s, QL = ql]}{\delta t}, \quad s \neq 1 \\ \phi(s, 0, t, ql) &= \lim_{\delta t \to 0} \frac{P[S(t + \delta t) = 0 | S(t) = s, QL = ql]}{\delta t}, \quad s \neq 0 \end{aligned} \tag{5.23}$$

To provide an indication of the PO operation level, s at time t we derive fundamental differential equations for the process by considering the forward Chapman-Kolmogorov equation as given in Bailey (1964) under section titled *Diffusion Processes* and also outlined in Cox and Miller (1965). Thus considering the states $0 < s < 1$,

$$\begin{aligned} p(s,t+\delta t) =& p(s,t)\left[1 - \delta t \int_{\substack{0+\\ r\neq s}}^{1-} \phi(s,r,t,ql)dr - \delta t\phi(s,1,t,ql) - \delta t\phi(s,0,t,ql)\right] \\ &+ \delta t \int_{\substack{0+\\ r\neq s}}^{1-} p(r,t)\phi(r,s,t,ql)dr \\ &+ \delta t P_1(t)\phi(1,s,t,ql) + \delta t P_0(t)\phi(0,s,t,ql) \\ &+ \o(\delta t) \end{aligned} \tag{5.24}$$

This leads us to the following forward differential equation,

$$\begin{aligned} \frac{d\,p(s,t)}{dt} =& -p(s,t)\left[\int_{\substack{0+\\ r\neq s}}^{1-} \phi(s,r,t,ql)dr + \phi(s,1,t,ql) + \phi(s,0,t,ql)\right] \\ &+ \int_{\substack{0+\\ r\neq s}}^{1-} p(r,t)\phi(r,s,t,ql)dr \\ &+ P_1(t)\phi(1,s,t,ql) + P_0(t)\phi(0,s,t,ql) \end{aligned} \tag{5.25}$$

If it is assumed that full performance of the PO is available at $t = 0$ then suitable initial conditions for the system of equations are,

$$\begin{aligned} p(s,0) &= 0, \quad \text{for } 0 < s < 1 \\ P_0(0) &= 0 \\ P_1(0) &= 1 \end{aligned} \tag{5.26}$$

In this case it is assumed that the PO is achieved and full performance is initially possible. If we assume that repairs, replacement or any other form of maintenance of the PO is not undertaken the evaluation of $p(s,t)$, $P_0(t)$ and $P_1(t)$ is simplified. The control of the decline in performance is possible by specifying suitable functions for $\phi(s, r, t, ql)$ such that the absorbing state $s = 0$ is reached much faster for higher levels of ql.

If we assume that maintenance is not performed then,

$$\phi(s, r, t, ql) = 0 \quad \text{for all } t \text{ and all } r > s \tag{5.27}$$

Further work includes the development of project outcome performance measures such as the expected time from start of operation, $t = 0$, until PO performance declines to a level β. Scenarios of project conduct may be considered enabling assessment of PO performance given the expected quality loss experienced. Similar model formulations may be investigated for the measures operation availability and operation output previously discussed in Chapter 1.

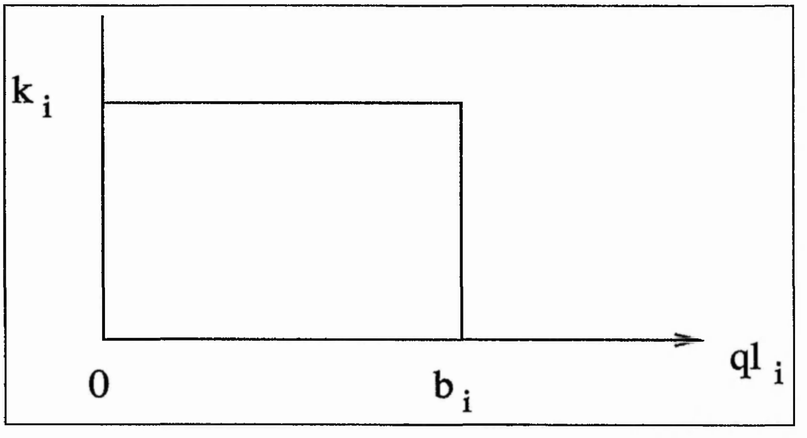


Figure 5.1: Quality Loss p.d.f with Phase Dependency

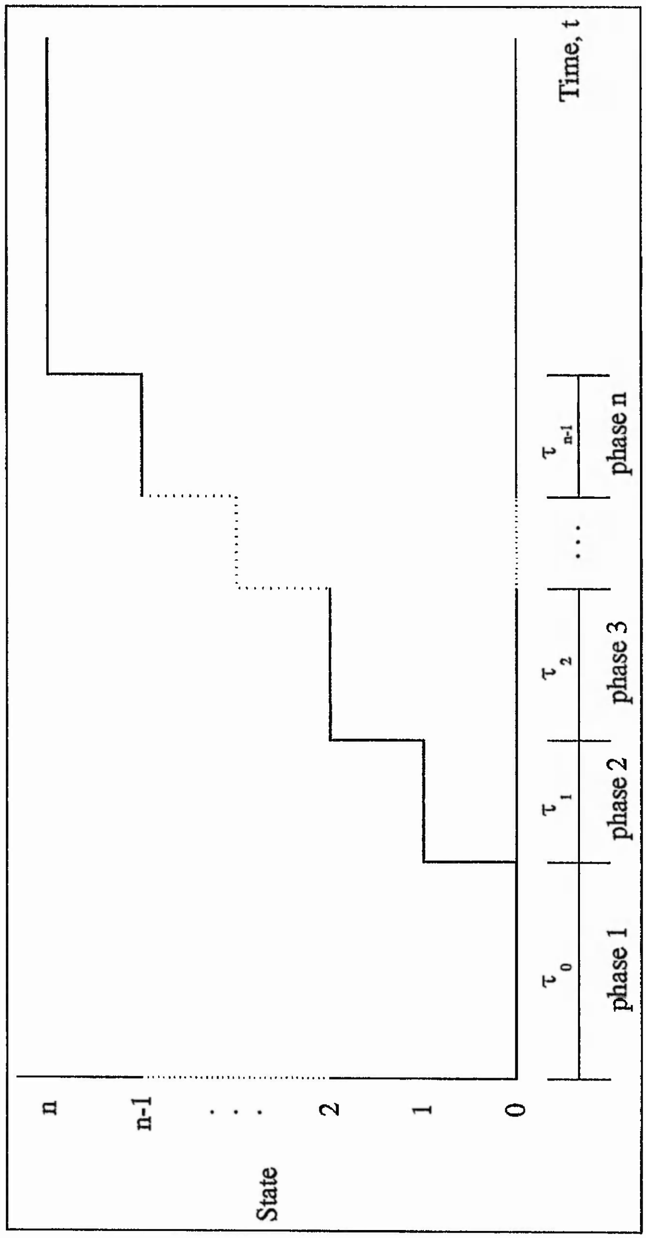


Figure 5.2: A Markov Project Process

Chapter 6

Simulation of Dependent Project Characteristics

6.1 Introduction

We now turn to the use of simulation to overcome the modelling limitations encountered when representing the evolution of a project as a stochastic process. By relaxing the Markov assumption, dependencies may be modelled between non-adjacent phases in addition to modelling the dependency between like and unlike project characteristics. Also we may generalise the formulation of Chapter 1 to include parallel project phases prior to operation, which is beneficial in the analysis of many project networks. As modelled in Chapter 5 we assume that the proportional hazard model is a suitable technique to model the reliability of the project outcome during the operation phase.

6.1.1 Review of Project Risk Simulation Approaches

The use of simulation is a common tool in project management, typically used to assess the duration and cost characteristics of a project. Currently there are numerous risk analysis software packages available, of which the majority interface or require the use of additional packages such as spreadsheet or project management software.

Stand alone packages including Risknet have tended to be expensive and for specific project types. For a summary of the details of a selection of common risk analysis packages available in the UK see Appendix 1. In general we found many of the packages to be similar in terms of the input probability distributions available, the type and presentation of the results and in most cases, as employed in the PERT approach, project activities and characteristics are assumed independent. However, the non-commercial VERT program not covered in Appendix 1, was initially developed by Moeller (1972), and provides the facility to aggregate up to three uncertainties for each activity of a project network. As identified the three uncertainties typical of most projects are time, cost and performance which, if assumed stochastic within VERT, may be represented as three independent probability distributions. The dependency facility that is available is in the use of an empirical relationship between two characteristics such as time and cost where, for example, a sampled value from the time distribution determines a cost value.

This dependency is also adopted in the software, RISNET II, as reported by Fennell (1989) in addressing the concept of Total Risk Assessing Cost Estimating (TRACE), implemented in the US Army and Navy. In this case, where $Cost = A + B \times Time$, not only is the time characteristic stochastic, but the linear parameters, A and B are assumed stochastic and suitable probability distributions are allocated to represent the consequences of technical problems, schedule slips and other unplanned occurrences.

A distinction between RISNET II and the majority of relevant simulation programs is unlike most simulation packages, the initial representation of activity time is deterministic and referred to as the 'baseline' values. Variability is modelled for each 'baseline' value and thus enables the uncertainty associated with project completion to be simulated. Unlike typical risk measures quantified as a probability RISNET II quantifies the risk as the difference between the baseline schedule and the value at a specific probability level. Thus risk may be quantified in terms of either time or cost.

Another useful modelling feature, as indicated by Kidd (1987) in advocating the use of VERT, also available within RISNET II, is in the handling of node logic where the aggregation in the project measures may be controlled and contingencies enforced according to predetermined conditions. Possible node logic available include 'AND' analogous to the maximum operator where all paths to a node must be completed, 'OR' which requires at least a single path to be completed, and 'PAND' standing for Partial And which requires at least one input path to be completed before continuing however all incoming paths must be completed before the project is considered complete. Also FILTERS may be set which ensure conditions are met, such as adhering to target boundaries for time, cost, and performance before project continuation is possible.

Another key point is that, in most examples discussed within the risk software documentation, projects are represented as a network of activities. However as indicated in BS6046 (1984), nearly all projects may be divided into phases to indicate the type of tasks or activities to be conducted in some logical sequence. In the simplest case we may assume that a project comprises of sequential phases as formulated in Chapter 1. For more advanced studies we may relax this assumption and suppose that phases may be performed in parallel, which is typical of a project activity network analysis.

6.2 Simulation Objectives

From work of previous chapters, where emphasis has been on project characteristics, dependencies and risk measures, we believe that building upon the ideas employed in VERT and RISNET II, further dependencies may be modelled with the aid of simulation. Our objective through the use of simulation is to identify further modelling capabilities which overcome the restrictions associated with analytical studies. In brief, as revealed in Chapter 5, it can be seen that our problem of modelling the uncertainties and dependencies within a project is not amenable to a totally mathematical solution. Restrictions are apparent in the use of the Markov property in that dependency between past events cannot be considered in determining the performance of future events. In addition, conditional dependency based on target measures was not possible and parallel operations were not considered due to the computational complexities associated with dependent paths. Therefore in achieving our objective, further measures and relationships of interest may be quantified, which in essence may enhance the knowledge associated with a project and provide justification for management decisions. For example, the decision to continue with a project where characteristic target values have been exceeded may be analysed effectively from simulating the consequence dependency.

In summary we shall make as few restricting assumptions as necessary in order to provide flexibility in the modelling possibilities. Similar to RISNET we shall provide the option of setting a baseline type measure for each activity characteristic, however unlike the deterministic value used our initial characteristic values are stochastic when used. It is however assumed that a baseline stochastic distribution is specified for all influencing characteristics of a given activity which are not dependent on any other characteristics of the given activity.

To enable the inclusion of such features we shall now develop a simulation tool to model the complex dynamic stochastic characteristics of a project. This in turn enables the generation of statistics associated with the performance of the project and assessment of the risks associated with the project characteristics. Where relationships are defined between the PO measures and project performance characteristics we may assess the availability, reliability and output of the PO as discussed in our general formulation outlined in Chapter 1.

As is common in performing simulations, as indicated by Van Slyke (1963), a criteria for a successful simulation includes the speed of generating the events within the simulation and the suitability of the statistics to the measures of performance of the process.

6.3 Characteristic Dependency

We shall assume that for a given activity i, the characteristics, which may be time, cost, and quality loss, are not necessarily independent, are stochastic, and for convenience represented by random variables T_i, C_i, QL_i. We refer to within activity dependency as *intra-dependency.* In addition we recognise the possibility of dependency between non-adjacent activities and shall refer to this as *inter-dependency.* To provide a structure for *inter-dependency* we suggest that where a dependency exists the performance of subsequent activities is dependent on the realisation of the previous activity. In addition we shall suppose that the dependency, or knock-on effects, are only enforced when a certain condition such as exceeding a characteristic target is experienced. Thus assuming all activities are performed within predetermined target values each activity is *inter-independent.* Where dependency is set and targets are exceeded the knock on effects will alter the stochastic nature of subsequent activity characteristics.

Since no general model exists in the representation of project events and dependencies we shall represent our project as sequential phases, where it is recognised that each phase may comprise of an activity network.

6.3.1 Intra-Dependency

To explore possible configurations of *intra-dependencies* allowing for various intuitive dependencies we shall consider the general case where k characteristics are modelled for a single activity. As investigated in Chapter 4 in developing a suitable bivariate time-cost exponential distribution, we supposed that time has the greatest influence over the characteristic cost, and thus assumed time is independent of cost, yet cost may be dependent or independent of time. If we denote the ith characteristic by X_i, we may represent the possible characteristic dependencies with the aid of directed graphs. To avoid paradoxes where there is the possibility for a two way dependency, we shall assume that if a dependency exists between any two given characteristics the dependency must be one way. Thus for $k = 2$ the possible dependency configurations are indicated in Figure 6.1.

In the case where $k = 3$, to avoid the repetition of similar configurations, as shown in Figure 6.1(b), we only present the unique dependency configurations, see Figure 6.2.

The recognition of the qualitative dependency between the characteristics has been reported by many authors, however as indicated by Klein (1993) the quantitative relationship between the characteristics could be complex and advises to employ non-quantitative relationships where possible. Such an approach enhances the understanding of activity characteristic relationships, however useful analysis is limited in obtaining quantitative measures. We begin by a further discussion of possible dependencies between three activity characteristics. If we suppose that the characteristics X_1, X_2 and X_3 are time, cost and quality loss respectively, we may interpret Figure 6.2 as follows. Configuration (a) where no dependency exists is the case assumed in the VERT program, however the characteristic performance is used instead of our quality loss measure. In Chapter 4, we developed a time-cost dependency, where cost is dependent on the time realised, and discussed the case where quality loss has a mixed exponential relationship with time. Such a configuration is represented in (c). If we supposed that in addition to a time-quality loss dependency, that the amount of money spent on an activity effects the level of quality loss, as indicated in (f), we may model further realistic dependencies.

For example, if a proportion of the cost of performing an activity contributes towards quality control measures, often included under the heading of 'quality costs', such as inspection costs or tasks to ensure conformance to specification, one would expect quality loss to decrease with the amount of money spent. The remaining configurations (b), (d), and (e) are not as intuitively appealing with the characteristics as defined, however may be of interest with different activity characteristics.

Another plausible three level dependency for configuration (f) is where X_2 denotes quality loss and X_3 denotes cost. Here, the time spent on an activity once again affects both the cost and quality loss, however instead of the amount of effort determining the level of quality loss, it is level of quality loss that determines level amount of money that is spent. If we assume that quality loss is static in that, once determined for a given activity it cannot be altered, the cost comprises of the efforts to achieve the level of quality loss. For low levels of quality loss one may assume that greater efforts have been enforced and thus costs are high, however for high levels of quality loss, one may assume that efforts have been minimal and thus costs are low.

Since there is little material with regards to the quantification of the exact dependency between activity characteristics, we shall suppose that a suitable dependency may be achieved by adjustment of the expected value associated with the dependent characteristic. Prior to developing such a relationship we shall define a suitable p.d.f. to model the uncertainty associated with each characteristic. As we are not restricted to a distribution that is mathematically tractable and limiting the modelling capabilities we suggest the use of the gamma distribution to represent the uncertainty associated with each characteristic.

The gamma distribution is given by,

$$f(x;\alpha,\rho) = \frac{\rho(x)^{\alpha-1}e^{-\rho x}}{\Gamma(\alpha)} \tag{6.1}$$

for $0 \leq x \leq \infty$ and $\alpha, \rho > 0$.

The gamma distribution allows modelling variations as it reduces to an Erlangian distribution when α is integer and simplifies to an exponential when $\alpha = 1$.We shall now proceed with our intra-dependency formulation. Suppose that characteristic X, with p.d.f, $f(x;\alpha_x,\rho_x)$ and mean μ_x, is dependent on characteristic Y with p.d.f, $g(y;\alpha_y,\rho_y)$. To incorporate a suitable dependency, we suggest that the mean value μ_x is adjusted to a value μ'_x based upon a suitable function of μ_x and a realisation of Y. For convenience, similar to the dependency of RISNET II we adopt a linear dependency where for a realised value, y, the adjusted expected value of X is given by,

$$\mu'_x = \mu_x + b_y y \tag{6.2}$$

Thus from specification of a suitable parameter, b_y, and assuming the shape parameter α_x remains constant, we may determine the scale parameter ρ_x. In general where a given characteristic X is dependent on k characteristics denoted by $Y_1, \ldots Y_k$,

$$\mu'_x = \mu_x + \sum_{i=1}^{k} b_i y_i \tag{6.3}$$

A possible simplification is if a characteristic X is totally dependent on a characteristic Y, initial distributions need only be specified for independent characteristics Y_i. In this case, μ_x is initially set as zero and the p.d.f of characteristic X is totally dependent on the realisation of characteristic Y.

6.3.2 Inter-Dependency

Apart from the characteristic dependencies within an activity, dependencies may exist between phases in terms of characteristic relationships. As indicated by Nero (1991) in formulating a parametric quantitative model for the evaluation of construction quality, phase dependency may exist in that the quality of a given phase may be influenced by the quality achieved in a previous phase. He also recognised that the quality of a given phase is very much dependent on the contractors used which provides additional support in the need of quantifying a quality loss measure. We consider a more general approach in that for a given activity any of the characteristics defined may influence the characteristics associated with any successive activity.

Similar to the intra-dependency discussed above we incorporate dependency within the expected value of the dependent characteristic, however an additional target based feature is also included. We appreciate that this is not the only type of dependency possible, but at present provides the basis for further investigation. Consider the case where characteristic X is dependent on q characteristics of previous activities denoted by $Z_1, \ldots, Z_q$. Suppose that for each Z_i, a target value, θ_i is specified, which depending on the characteristic may be a time, cost, or quality loss target. From inclusion of such a feature we may model the consequences of not meeting targets at earlier stages. There are numerous possibilities, for example the consequence of not attaining the required level of quality in an earlier activity may result in more time and effort in a later stage to rectify areas of quality loss. Here Z_i is quality loss, θ_i possibly an upper limit to an acceptable level of quality loss, and X could either be the time or cost characteristic of a successive activity. If no contingency measures are in place to remove quality loss, growth in quality loss may be experienced throughout successive activities as demonstrated in the sequential parametric quality loss model of Chapter 2. In this case a dependency exists between X and a given Z_i which both measure levels of quality loss experienced. The dependency also provides a means of adjusting the initial characteristic distributions if they are considered either optimistic or pessimistic. In general if in addition to characteristics X and Z we include the k intra-dependent characteristics, denoted by Y_j, the modified expected value of X is given by,

$$\mu'_x = \mu_x + \sum_{i=1}^{q} b_{Z_i}(z_i - \theta_i)U(z_i - \theta_i) + \sum_{j=1}^{k} b_j y_j \tag{6.4}$$

where $U(t)$ is defined as the unit step function. Such a relationship enables the performance of a given activity to affect future activities. However if we supposed that it is the cumulative level for a characteristics measure that is of concern possibly the total time, total cost, and total quality loss experienced in a given phase, where for each characteristic measure a target value is specified, the dependency requires further information. We need to know how each characteristic measure escalates from both sequential and parallel activities.

6.4 Simulation Methods

Initial investigations were focused in the use of *Mathematica* to simulate a project with dependencies described above. Advantages included the use of a powerful graphics environment to illustrate the results generated from the add on statistics package 'Descriptive Statistics'. For simplicity and the availability of a random number generator within *Mathematica* to demonstrate the benefit of dependency modelling, we first adopted the use of the inversion method to generate a random variate from an Erlang distribution. The method applied is as follows. For a given p.d.f., $f_X(\cdot)$ with c.d.f. $F_X(\cdot)$ we can obtain the variate,

$$X = F_X^{-1}(R)$$

where,

$$R \sim U(0,1)$$

The accuracy of the variate was sufficient, however the speed to successfully simulate 10,000 iterations of a network, comprising of two phases, each of three characteristics was unacceptable, taking approximately twenty six hours on an unloaded Sun workstation. Based upon our simulation criteria of speed we did not pursue the further implementation of *Mathematica* to perform simulations.

We thus explored other methods for simulating a gamma type distribution. There are various algorithms that may be applied depending on the value of the gamma parameter α. If $\alpha < 1$, possible algorithms include the power method, or the switching algorithm. Studies carried out by Dagpunar (1988) where the algorithms were implemented in Fortran, and times observed for generating a random variate, showed the switching method to be the most time efficient. When $\alpha = 1$ the gamma distribution reduces to the exponential distribution. Common variate generation methods include the inversion method, the method of Forsythe (1972), and the method due to Maclaren et al (1964). When $\alpha > 1$ various methods are available including the ratio method, the Cauchy method, Ahrens and Dieter (1974), the Log-logistic method, Cheng (1977) and the t-distribution method proposed by Best (1978). In our literature search we came across references to NAG routines to generate pseudo-random variates from a gamma distribution. The NAG Fortran Library is a comprehensive collection of Fortran 77 routines for the solution of numerical and statistical solutions, NAG (1990).

The NAG routine G05FFF employs three algorithms which are implemented depending upon the value of the parameter α. For $\alpha < 1$ a switching algorithm is called as reported by Dagpunar (1988) as the most efficient method, for $\alpha = 1$ the logarithmic transformation method of a uniform random variate is used, and lastly for $\alpha > 1$ the t-distribution of Best (1978) is applied.

6.4.1 Overview of Simulation Program

In light of the efficient NAG routines available to generate a gamma variate we decided to code a program in the language Fortran 77, that would allow the modelling of both *inter* and *intra* dependencies, see Appendix 3 for program listing. The high quality graphical output easily generated with *Mathematica* is not easily achieved in Fortran even with implementing a NAG routine and therefore in order not to deviate from the purpose of performing a simulation we export all simulated values into *Mathematica* to generate graphical output. However similar to the output achieved with the 'Descriptive Statistics' package of *Mathematica* we may utilise the NAG routine G01AAF to generate summary statistics for each characteristic at project completion.

Based on our adopted approach there are three procedures involved in performing a full analysis, (i) Datafile setup, (ii) Simulation and (iii) Analysis. The simulation procedure is complex and contained within the program and is considered of limited interest here other than commenting on the input requirements set as a datafile and results generated as shown in the examples that follow.

6.4.2 Dependency Datafile Setup and Default Parameters

In this section we describe the layout of the project characteristic datafile that is assumed by the simulation procedure. The datafile has the same basic format irrespective of the dependency analysis being undertaken. The datafile used by our program is written in ASCII and has the following standard format, (note the colon and the following comments are not part of the file, but are explanatory notes. Also provided are the corresponding variable names as used in the simulation procedure).

nevents : Number of project events

nact : Number of project activities

nchar : Number of activity characteristics

ninterd : Number of inter-dependencies

iters : Number of simulation iterations

nanaly : Specifies the number of analyses to perform where each analysis provides results between two given project events indicated by the variable **fromto**

network : Represents the project network structure as a **nevents**×**nevents** array where values indicate the activity between event row_i and event $column_j$. A zero indicates no activity between events

intra : Specifies the intra-dependencies between each characteristic of each activity. In total **nact**×**nchar** rows and **nchar** columns of the realisation coefficient, b, as used in equation(6.2), are collected. Where a zero is entered it is assumed that no intra-dependency exists

shape : Specifies the Gamma shape parameter for the p.d.f of each activity characteristic. In total an array of **nact**×**nchar** values are required

mean : Specifies the mean value for the p.d.f of each activity characteristic. An array of **nact**×**nchar** values are required

parop : Specifies the operation to perform for each characteristic measure when parallel activities are performed. (1=maximum, 2=summation, 3=minimum, 4=average). **nchar** values are required

target : Specifies the target value for each activity characteristic. There are a total **nact**×**nchar** values that may be set

inter : The variable **ninterd** specified above determines the number of inter-dependencies. If **ninterd**$=0$ then no values are required. Otherwise for each inter-dependency, the prevailing event, the prevailing characteristic, the dependent activity, the dependent characteristic and the realisation coefficient are specified

fromto : Where a network analysis is required, determined by **nanaly**>0, two events are specified indicating the part of the network to report results for

6.5 Demonstration of Simulation Model

6.5.1 Example - Characteristic Dependency within a Phase

In this example we illustrate intra-dependency within a single phase. For convenience we implemented configuration (f) of Figure 6.2 where cost is dependent on time and quality loss is dependent on both the time and cost. Initial mean values for the characteristics time, cost, and quality loss are set as $\mu_t = 40$, $\mu_c = 100$, and $\mu_{ql} = 10$. For a duration of t, we have assumed that the μ_c increases by $0.2t$, and μ_{ql} increases by $0.5t$. To incorporate the effects of quality control we have assumed that given cost, c, μ_{ql} decreases by a value $0.2c$. In addition to simulating the dependency with 10,000 iterations, we also considered the case of independent characteristics for comparison. The results of both simulations are presented as descriptive statistics which were generated with the Fortran program, see Table 6.1, and c.d.f plots generated within *Mathematica*, see Figure 6.3 and Figure 6.4. Each simulation took approximately five seconds to run and generate the descriptive statistics which is a significant improvement over solely using *Mathematica*, where execution time often exceeded one day.

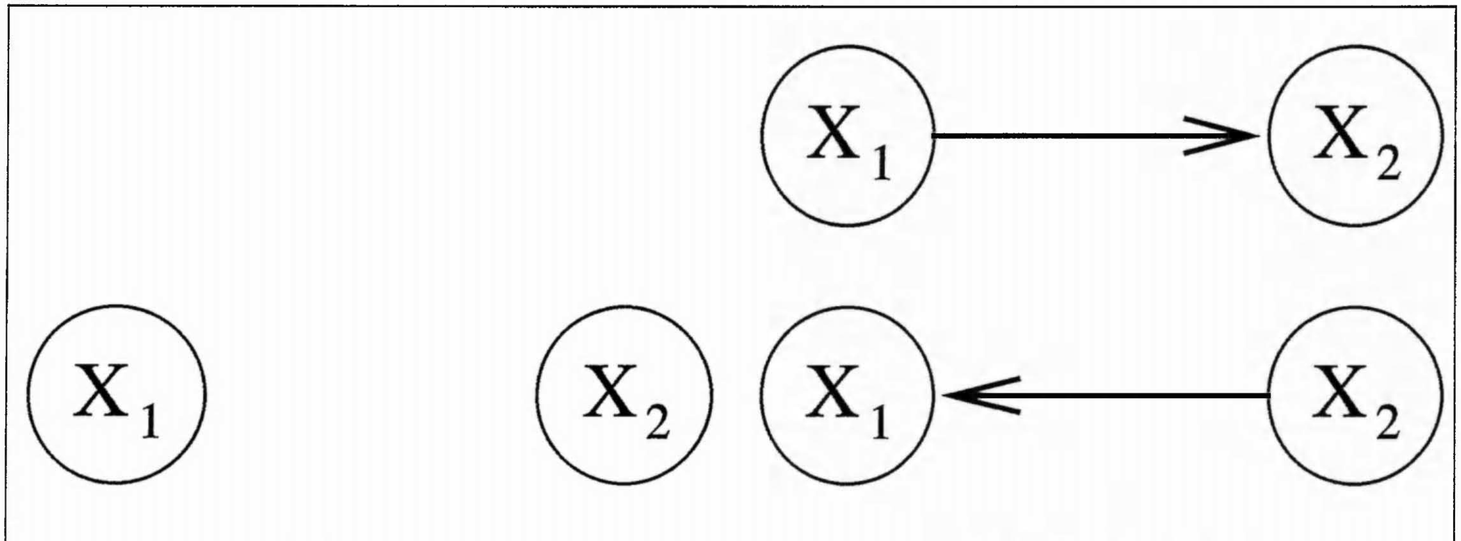
Since in both cases time is assumed independent, the c.d.f. for each plot is identical and therefore not reported. As shown in Figure 6.3 the cost distributions are similar implying risk values will be alike for a given cost. However as evident in Figure 6.4 the quality loss distributions differ greatly implying a significant difference in risk for a given level of quality loss.

Obviously all scenarios will differ according to the dependency parameter values modelled, however in this case, we have illustrated that different risk values may be obtained if the nature of the dependency which exists in practice can be captured with our dependency model.

6.5.2 Example - Characteristic Dependency between Phases

We now consider the case where not only dependency exists between characteristics of project phases, but if certain characteristic target values are exceeded, then a future phase will be affected. To illustrate this feature we consider two sequential phases where intra characteristic dependency exists as in the above example, however inter-dependency exists between the quality loss characteristic and the mean time to complete the successive phase. We now explore the differences between modelling intra-dependency with inter-independency and intra-dependency with inter-dependency.

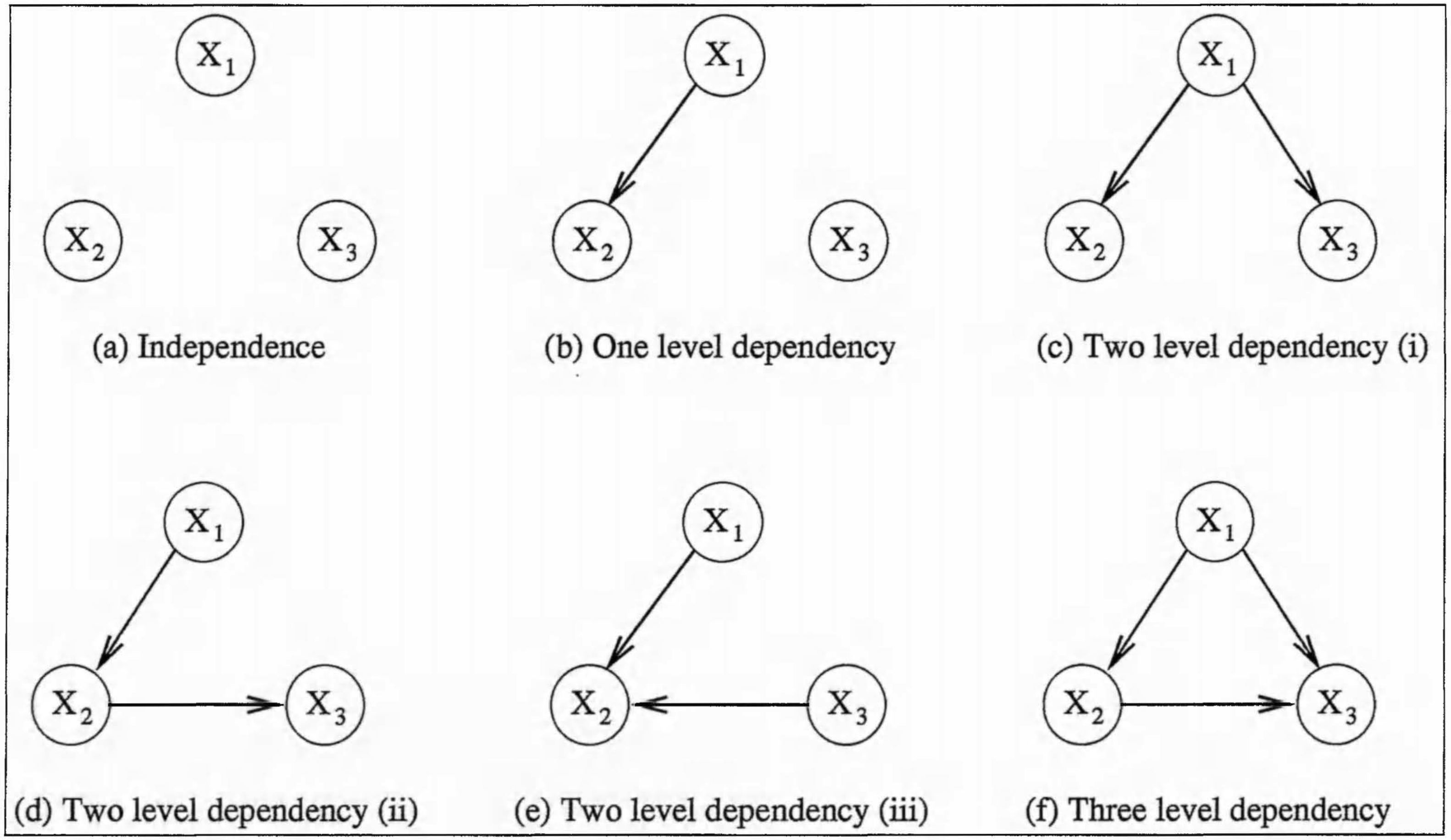
For phase one we set the upper limit (target) for quality loss as 30. If this value is exceeded then the consequences will be encountered, possibly in rectification, in phase two. We assume that the rectification of the excess in quality loss affects the duration of phase two, which in turn will affect the cost, and quality loss. Once again the time to perform each simulation was minimal, taking approximately seven seconds. The difference in results from both simulations in the total time, total cost and total quality loss, are indicated in Figure 6.5, Figure 6.6, and Figure 6.7 respectively. As shown the differences are mainly in the extreme tail of the distribution where the risk is typically calculated. This indicates that if dependency is overlooked when modelling the project the risk may be significantly underestimated in meeting given targets for the project characteristics, which in turn may affect the operation phase.



(a) Independence

(b) One way dependency

Figure 6.1: Dependency Configuration of two Characteristics



(a) Independence (b) One level dependency (c) Two level dependency (i)

(d) Two level dependency (ii) (e) Two level dependency (iii) (f) Three level dependency

Figure 6.2: Dependency Configuration of three Characteristics

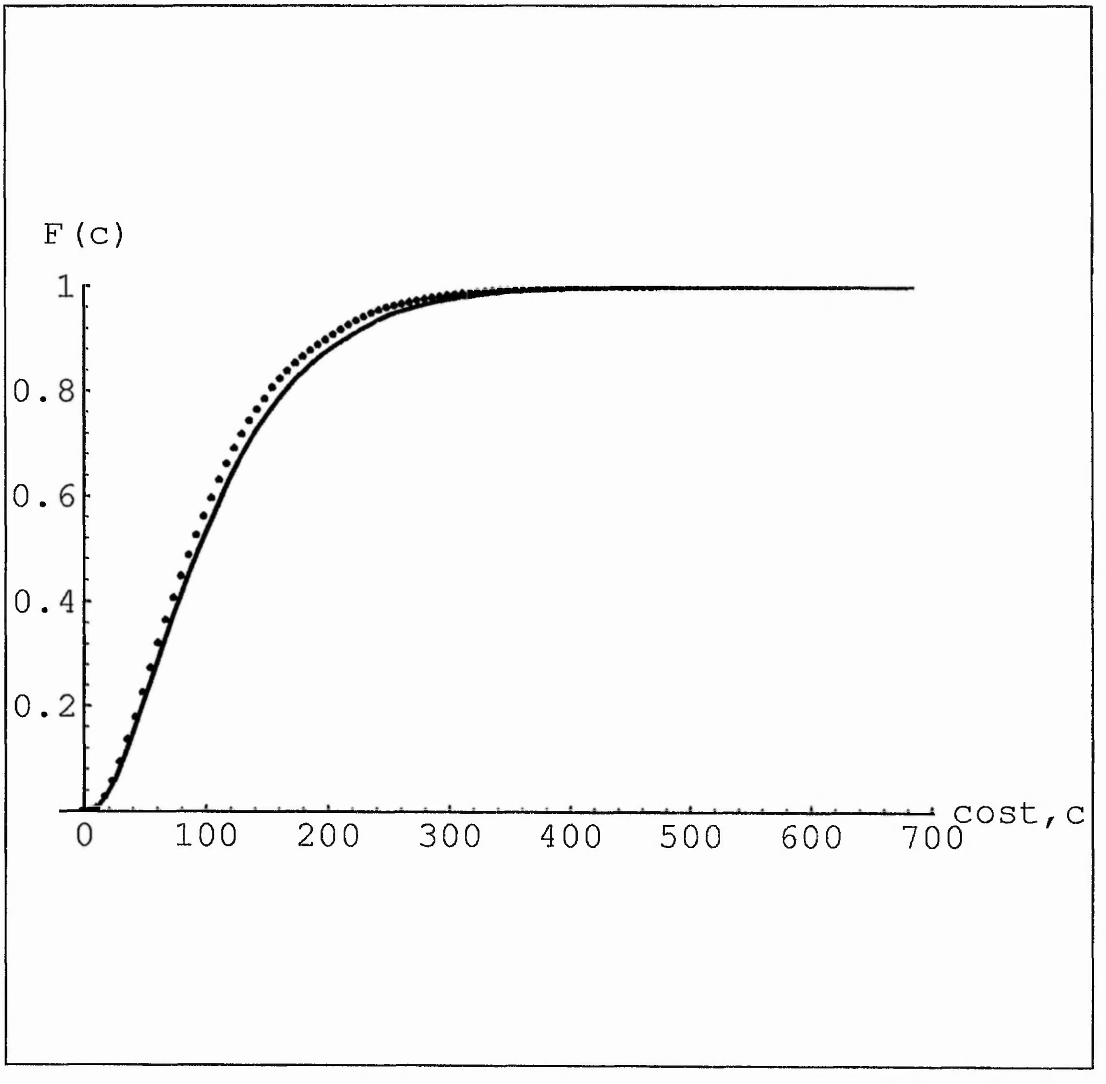


Figure 6.3: Simulation Results of the Phase Characteristic Cost; ....Independent Characteristics, ________Dependent Characteristics

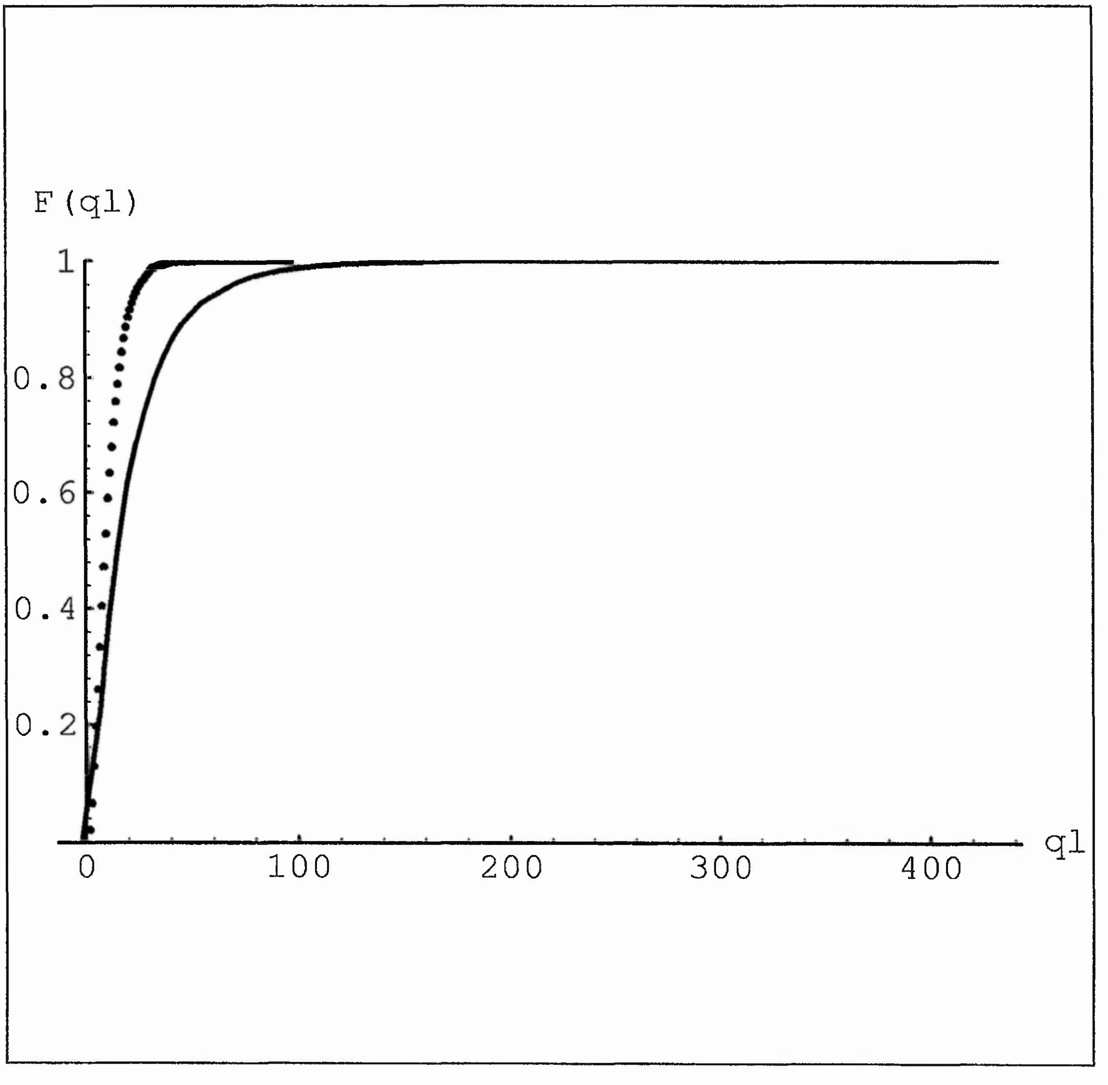


Figure 6.4: Simulation Results of the Phase Characteristic Quality Loss; ....Independent Characteristics, ________Dependent Characteristics

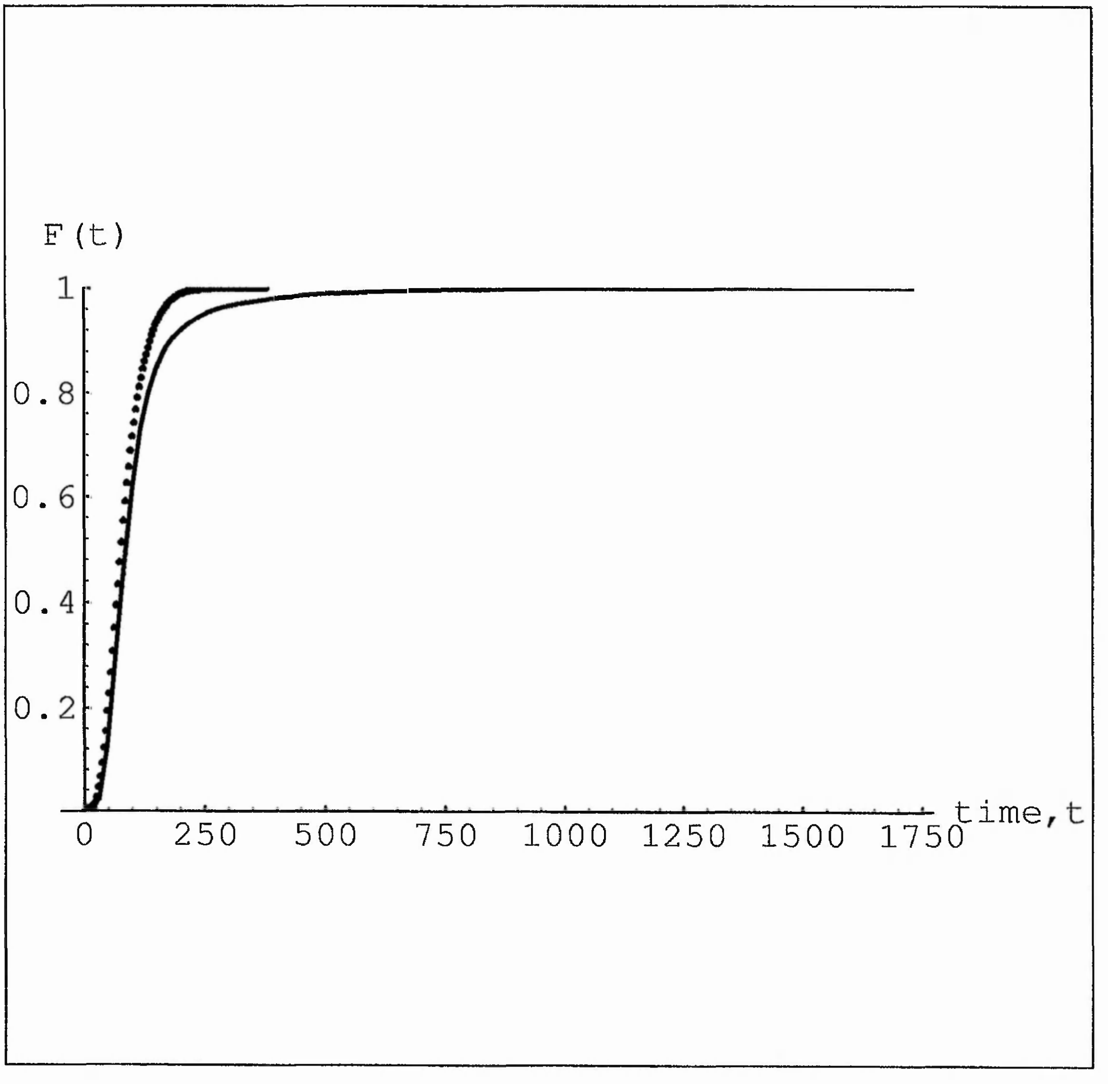


Figure 6.5: Phase Dependency Simulation Results with for the Characteristic Time; ....Independent Phases, ________Dependent Phases

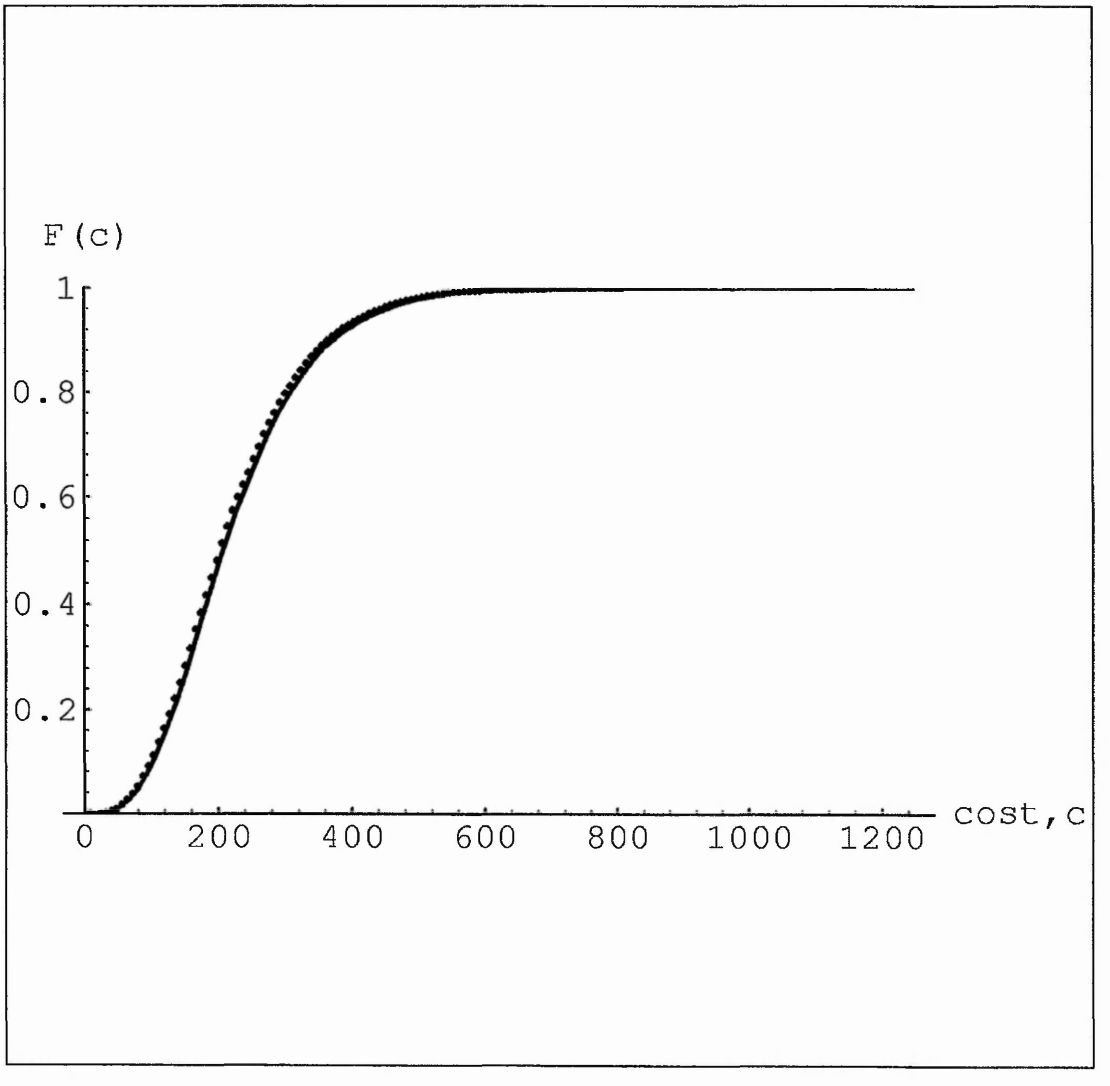


Figure 6.6: Phase Dependency Simulation Results for the Characteristic Cost; ....Independent Phases, ________Dependent Phases

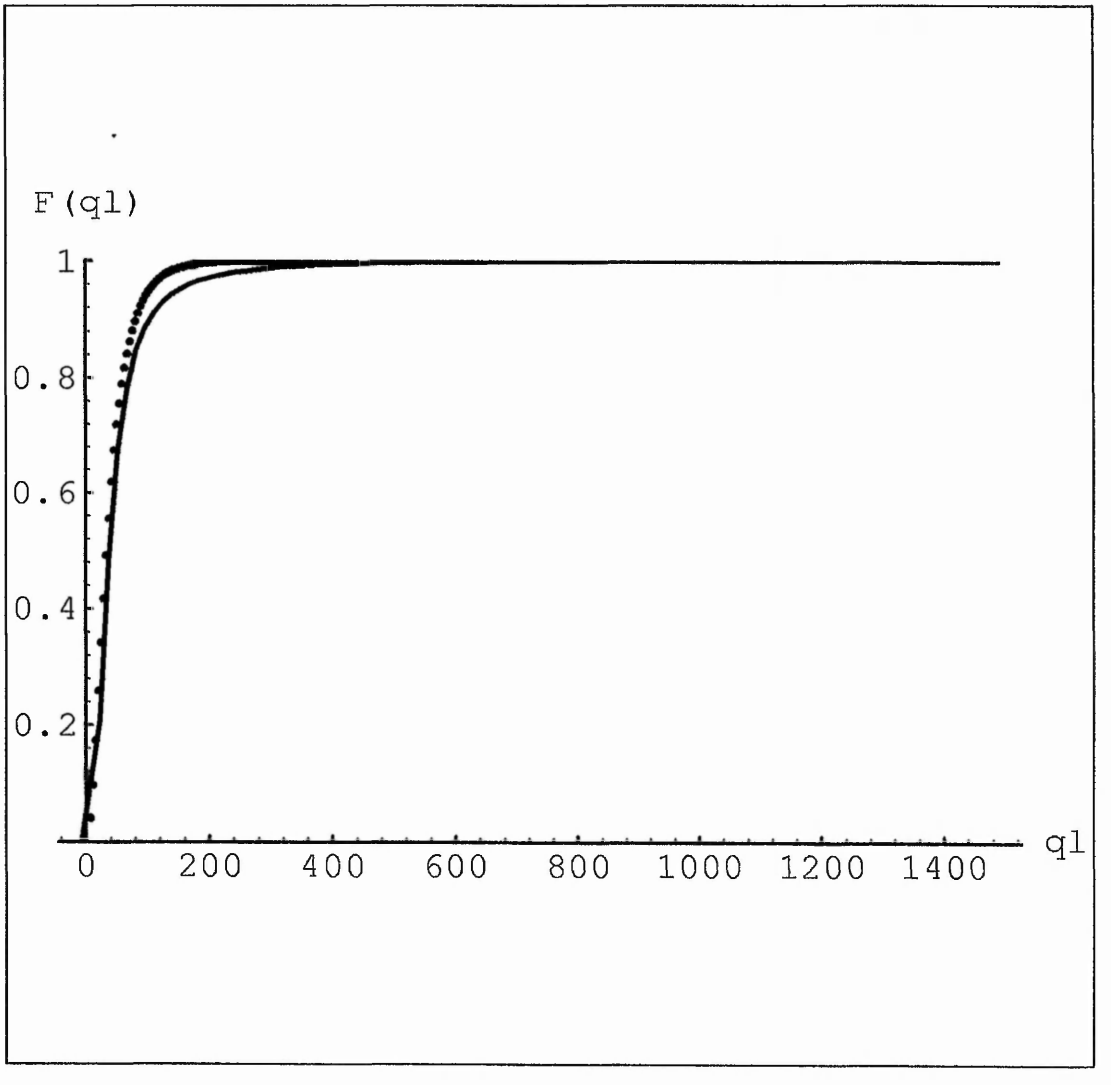


Figure 6.7: Phase Dependency Simulation Results for the Characteristic Quality Loss; ....Independent Phases, ________Dependent Phases

	(i) Dependent Characteristics			(ii) Independent Characteristics		
Statistic	Time	Cost	Quality Loss	Time	Cost	Quality Loss
μ	40.36	107.56	19.35	40.36	99.52	9.84
σ	28.13	75.83	22.36	28.13	69.85	7.02
skew	1.36	1.39	3.26	1.36	1.37	1.52
kurtosis	2.70	2.96	24.09	2.70	2.89	4.68
min.	0.15	0.88	0.00	0.15	0.75	0.10
max.	238.83	680.35	429.63	238.83	627.48	96

Table 6.1: Descriptive Statistics assuming (i) Dependent Characteristics, (ii) Independent Characteristics

Summary of Advances in Knowledge Achieved, Conclusions, and Further Work

This thesis is concerned with methods to model Project Risk where time, cost and quality are incorporated. Chapter 1 reviews the literature and outlines the basic concepts of a risk analysis. Limitations of current methodologies are highlighted with examples from the construction industry. A conceptual model is formulated which incorporates the operation phase of a project and provides a framework to measure the project characteristics time, cost and quality. We introduce a quality loss measure as a criteria for assessing the performance of a project and project reliability to assess operation performance.

In Chapter 2 we investigate the use of PERT, analytical approaches, method of moments and fuzzy numbers in performing a project risk analysis of a single project characteristic. Computational difficulties are discussed and improvements suggested. We develop a parametric quality loss model to investigate the costs associated with quality and unreliability.

Chapter 3 introduces the use of Laplace transforms to investigate project networks where activity probability distributions are special Erlangian. An algorithm is defined and implemented within *Mathematica* and illustrated with examples. It is shown to provide a better approximation of the c.d.f. of project completion time than the method of moments, PERT and the case where activities are assumed independent.

In Chapter 4 a time-cost dependency is constructed as a bivariate exponential extension such that cost is dependent on time. Convolution of the time-cost distribution is explored using Laplace transforms and project risk is measured assuming sequential project phases. The moments for time and cost are obtained. Time-quality loss relationships are discussed and a model defined. The use of the proportional hazards model is examined where performance measures of the project phases are modelled as covariates.

The dependency issue is continued in Chapter 5 where a stochastic process is formulated for the performance of a project. A complete model is given where the state space is partly discrete and partly continuous representing the project phase and time in phase. We use the partial operation model of Bendell and Humble (1985) to model the dependency between project quality and operation reliability.

In Chapter 6 the concepts of inter and intra dependencies are introduced. A simulation program is developed to extend the dependencies in Chapter 5 to allow the relationships between the characteristics within and between phases to be modelled. Examples illustrating the benefits of such analysis are provided.

The overall conclusion of this thesis is that it is feasible and desirable to construct project risk models where dependency exists between project characteristics and operation measures. Limitations of exact approaches can be overcome with a sequential phase formulation and the scope of modelling features is extended with the use of simulation procedures.

Recommendations for Further Work

Collection of data for project types with common attributes is necessary to evaluate the application of the author's models in practice. Further investigation of methods to quantify the quality loss characteristic for specific project types is possible.

For the time-cost model of Chapter 4, the special case where $\frac{k}{\gamma} = \frac{1}{\lambda}$ requires investigation. An improvement in the software execution time of the implemented network reduction algorithm of Chapter 3 may be possible using alternative software such as C++ or Fortran.

Further work regarding the continuous time project formulation of Chapter 5, includes the development of expected time measures and incorporating the characteristics cost and quality loss. The operation performance model, also of Chapter 5, may be developed to include operation availability and operation output.

To evaluate the *inter* and *intra* dependencies specified in the simulation model of Chapter 6, the collection and analysis of project characteristic measures is needed.

References

ABIKO, A. K. (1991). Quality housing in the context of a developing country. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

ABRAMOWITZ, M. and STEGAN, A. (1972). *Handbook of Mathematical Functiobs*, Dover Publications, New York.

AHRENS, J. H. and DIETER, U. (1974). Computer methods for sampling from gamma, beta, Poisson and binomial distributions. *Computing*, **27**, 927-937.

ANDERSON, D. R., SWEENEY, D. J., and WILLIAMS, T. A. (1991). *An Introduction to Management Science*, 6th ed., West Publishing Company.

ANG, A. H. S. , ASCE, M., ABDELNOUR, J. and CHAKER, A. A. (1975). The Analysis of Activity Networks under Uncertainty. *J. Engr. Mech. Div.*, **101:8**, 373-387.

ANKLESORIA, K. P. and DREZNER, Z. (1986). A multivariate approach to estimating the completion time for PERT networks. *J. Opl Res. Soc*, **37:8**, 811-815.

ANSELL, J. I. and PHILLIPS, M. J. (1994). *Practical Methods for Reliability Data Analysis*, Oxford Statistical Science Series 14, Clarendon Press, Oxford.

ANSELL, J. I. and WHARTON, F. (1992). *RISK: Analysis, Assessment and Management.* Wiley.

ANSELL, J.I. (1992). *Reliability: Industrial Risk Assessment*, Ansell and Wharton (1992), Chapter 7, 105-122.

ARGENT S. J., BENDELL, A., MANNING, P. T., MARSHALL, J., RYAN, S. G. and WIGHTMAN, D. W. (1986). Proportional hazard modelling in the analysis of transmission failure statistics, *Proc. 5th Euredata Conference*, Heidelberg.

ASCHER, H. and FEINGOLD, H. (1984). *Repairable Systems Reliability*, Lecture Notes in Statistics, Volume 7, Marcel Dekker, inc., New York.

BAILEY, N. T. J. (1963). The simple stochastic epidemic: a complete solution in terms of known functions, *Biometrika*, **50**, 11.

BAILEY, N. T. J. (1964). *The Elements of Stochastic Processes; With Applications to the Natural Sciences*, Wiley.

BARNES, M. (1988). Construction Project Management. *Project Mgmt, 6:2*, 69-79.

BAXENDALE, A. T. (1991). Management information systems: the use of work breakdown structure in the integration of time and cost. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

BELL, T. E. and ESCH, K. (1989). The Space Shuttle: a case of subjective engineering, *IEEE Spectrum*, **26:6**, 42-46.

BENDELL, A. and HUMBLE, S. (1985). Reliability model with states of partial operation. *Naval Logistics Quarterly*, **32**, 509-535.

BENDELL, A., SOLOMAN, D. and CARTER, M. (1995). Evaluating Project Completion Times When Activity Times are Erlang Distributed, *J. Opl. Res. Soc.*, **46**, 867-882.

BENDELL, A. and PAYNE, B. (1996). The dependence of operating reliability on project quality. *Conference presentation*, Lulea University of Technology, Sweden.

BENDELL, A. (1998). *The use of FMECA to perform a risk analysis*, Personal communication.

BERNY, J. (1989). A new distribution function for risk analysis, *J.Opl. Res. Soc.*, **40:12**, 1121-1127.

BEST, D. J. (1978). Letter to the Editor. *Appl. Statist.*, **29**, 181.

BEZELGA, A and BRANDON, P. (1991). *Management, Quality and Economics in Building.* Transactions of the European Symposium on Management, Quality and Economics in Housing and other Building sectors, Lisbon, Portugal.

BEZELGA, A and De SOUSA, E. T. (1991). Quality assurance and management in building construction research and practical contributions. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

BIRREL, G. S. (1991). Factors and variables to induce effective construction schedules and efficient construction. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

BLANCHARD, B. S. (1978). *Design and Manage to Life Cycle Cost.* M/A Press, Oregan.

BRAZ, A. J. OLIVEIRA (1991). The use of aluminium systems in buildings. A model for quality assurance on installation. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

BS5760: *Reliability of Systems, Equipments and Components*, British Standards Institution, London.

BS6046: Part 1 (1984). *Guide to the use of management, planning, review and reporting procedures*, British Standards Institution, London.

BUCKLEY, J. J. (1989). Fuzzy PERT, *Applications of Fuzzy Set Methodology in Industrial Engineering*, 103-114.

BUREAU OF NAVAL WEAPONS (1958). PERT. Programs evaluation and review task, phase I, *Technical report*, Department of Navy, Washington D.C.

BURT, J. M. and GARMAN, M. B. (1971). Conditional Monte Carlo: A simulation Technique for Stochastic Network Analysis, *Management Science*, **18:3**, 207-217.

CHANAS, S. and KAMBUROWSKI, J. (1981). The use of Fuzzy variables in PERT, *Fuzzy sets and systems*, **5**, 11-19.

CHAPMAN, C. B. (1979). Large engineering project risk analysis, *IEEE Transactions of Engineering Management*, **76**, 78-86.

CHAPMAN, C. B., PHILLIPS, C. B., COOPER, D. F. and LIGHTFOOT, L. (1983). Selecting an approach to project time and cost planning. *International Journal of Project Management*, **3:1**, 19-26.

CHAPMAN, C. B., COOPER, D.F., and PAGE, M.J. (1987). *Management for engineers*, Wiley, Chichester.

CHARNES, A., COOPER, W. W. and THOMPSON, G. L. (1964). Critical path analysis via constrained and stochastic programming, *Opns. Res.*, **12**, 460-470.

CHENG, R. C. H. (1977). The generation of gamma variables with non-integral shape parameter. *Appl. Statist*, **26**, 71-76.

CHRISTENSEN, R. (1984). *Data Distributions: A Statistical Handbook*, Entropy Ltd, Massachusetts.

CLARK, C. E. (1961). The Greatest of a Finite Set of Random Variables, *Opns. Res.*, **3**, 145-162.

CLARK, C.E., MALCOLM, D. G. and FAZAR, W. (1959). Application of a technique for research and development program evaluation. *Opns. Res.*, **13**, 646-669.

CNUDDE, M. (1991). Lack of quality in construction - economic losses. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

COOK, T. M. and JENNINGS, R. H. (1979). Estimating a project completion time distribution using intelligent simulation methods, *J. Opl. Res. Soc*, **30:12**, 1103-1108.

COON, H. (1965). A Note on William A. Donaldson's 'The estimation of the mean and variance of a PERT activity time'. *Opns. Res.*, **13**, 386-387.

COOPER, D. F. and CHAPMAN, C. B. (1987). *Risk Analysis for Large Projecrs: Models, Methods and Cases*, Wiley, Chichester.

CORNICK, T. (1991). Construction quality and management - its delivery and discipline in housing and other building sectors. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

COX, D. R. (1962). *Renewal Theory.* Methuen, London.

COX, D. R. (1972). Regression models and life tables (with Discussion), *J. R. Stat. Soc. B*, textbf34, 187-220.

COX, D. R (1975). Partial likelihood. *Biometrika*, **62**, 269-276.

COX, D. R. and MILLER, H. D. (1965). *The theory of stochastic processes.* Methuen, London.

COZZOLINO, J. M.(1968). Probabilistic models of decreasing failure rate processes, *Nav. Res. Logist. Q.*, **15**, 361-371.

CURTIS, B., WARD, S. C. and CHAPMAN, C. B. (1991). Objectives and performance in construction projects. *Construction Management and Econ.*, **9**, 343-353.

DAGPUNAR, J. (1988). *Principles of Random Variate Generation.* Oxford Science Publication, Clarendon Press, Oxford.

DALE, B. G. and PLUNKETT, J. J. (1990). *Managing Quality*, Philip Allan, England.

DAVIS, E. W. (1974). CPM use in top 400 construction firms, *J. of the Construction Division*, **100**, 39-49.

DAVIS, C. S. and STEPHENS, M. (1983). Approximate percentage points using Pearson curves. Algorithm AS192, *Applied Statistics*, **32**, 322-327.

DODIN, B. M. (1980). Approximating the probability distribution function of the project completion time in PERT networks. *OR Report No. 153, North Carolina State University*, North Carolina.

DODIN, B. M. (1984a). Bounding the project completion time distribution in PERT networks. *Opns. Res.*, **3:4**, 862-881.

DODIN, B. M. (1984b). Determining the k most critical paths in PERT networks. *Opns. Res.*, **32:4**, 859-877.

DODIN, B. M. and ELMAGHRABY, S. E. (1985). Approximating the criticality indices of the activities in PERT networks. *Management Science*, (32:2), 207-223.

DONALDSON, W. A. (1965). The estimation of the mean and variance of a PERT activity time. *Opns. Res.*, **13**, 382-385.

DOWNTON, F. (1970). Bivariate exponential distributions in reliability theory. *J. of Royal Stats. Soc.*, **32**, 408-417.

DREGER, G. T. (1991). Contemporary quality control systems: case studies and lessons learned. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

DUBOIS, D and PRADE, H. (1988). *Possibility Theory: an approach to computerized processing of uncertainty*, Plenum Press, New York.

DUPONT (1958). CPM: Critical Path Method, *Technical Report*, E. I DuPont de Nemours Company.

ELMAGHRABY, S. E. (1995). Invited review, Activity nets: A guided tour through some recent developments. *European Journal of Opn. Res.*, **82**, 383-408.

ESHBACH, O. W. and SOUDERS, M. (1975). *Handbook of Engineering Fundamentals*, page 339, Wiley, London.

FARNUM, N. and STANTON, L.W. (1987). Some results concerning the estimation of Beta distribution parameters in PERT. *J. Opl. Res. Soc.*, **88:3**, 287-290.

FELLER, W. (1966). *An Introduction to Probability Theory and its Applications*, Vol. 2, 2nd ed., Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications [2 vols.]*, Vol.2, Chapter 2, Section 2, 'Gamma Distributions', Wiley, New York.

FENNEL, C. H. (1989). Network Risk Analysis. Paper presented at the Royal Aeronautical Society Conference 1989.

FORSYTHE, G. E. (1972). Von Neumann's comparison method for random sampling from the normal and other distributions, *Math. Comput.*, **26**, 817-26.

FREUND, J. (1961). A fundamental bivariate exponential extension, *J. Amer. Stat. Ass.*, **56**, 971-977.

FULKERSON, D. R. (1962). Expected critical path length in PERT networks, *Operations Research*, **10**, 808-817.

GOLENKO-GINZBURG, D. (1988). On the distribution of activity time in PERT, *J. Opl. Res. Soc.*, **39**, 767-771.

GOLENKO-GINZBURG, D. (1989). A new approach to activity-time distribution in PERT, *J. Opl. Res. Soc*, **40**, 389-393.

GRADSHTEYN, I. S. and RYZHIK, I. M. (1980). *Table of Integrals, series and products.* Academic Press, London.

GRATT, L. B. (1987). *Risk Analysis or risk assessment: a proposal for consistent definitions, in Uncertainty in Risk Assessment, Risk Management, and Decision Making.* Plenum Press. New York, 241-249.

GRUBBS, F.E. (1962). Picking on PERT. *Opns. Res.*, *10*, 912-915.

GUMBEL, E. J. (1960). Several early bivariate exponential distributions. *J. Amer. Stat. Ass.*, **55**, 698-707.

HAMMARLUND, Y. and JOSEPHSON, P. E. (1991). Sources of quality failures in buildings. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

HARPER, W. M., and LIM, H. C. (1982). *Operational Research*, 2nd ed., Macdonald & Evans Ltd, The Chaucer Press, Suffolk.

HARTLEY, H. D. and WORTHAM, A.W. (1966). A statistical theory for PERT critical path analysis. *Management Science*, **12:10**, 469-481.

HASTINGS, N. A. J. and PEACOCK J. B. (1974). *Statistical Distributions*, Butterworths, London.

HERBSMAN, Z. and ELLIS, R. D. (1991). The cost/time/quality integrated bidding system - an innovation in contract administration. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

HERTZ, D. (1964). Risk analysis in capital investment, *Harvard Business Review*, **42**, 95-106.

HERTZ, D. and THOMAS, H. (1983). *Risk Analysis and its Applications*, Wiley, Chichester.

ISO 8402 *Quality management and Quality Assurance*, International Standardisation Organistation.

ISO 9001-9004 *Quality Systems, Quality Assurance*, International Standardisation Organistation.

JACKSON, N. and CARTER, P. (1992). *The Perception of Risk*, Ansell and Wharton (1992), Chapter 3, 41-54.

JAAFAR, T. (1992). *A Unified Methodology for Project Planning Risk*, PhD thesis, The Nottingham Trent University, 1992.

JOHNSON, N. L. and KOTZ, S. (1972). *Distributions in Statistics*, see Chapter Continuous Multivariate Distributions, Wiley, New York.

JOHNSON, N. L., KOTZ S. and BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, Vol. 1, John Wiley Sons, Inc., New York.

KEDAR, B. Z. (1970). Again: Arabic Risq, Medieval Latin Risicum, Studi Medievali, Centro Italiano Di Studi Sull Alto Medioevo, Spoleto.

KIBBLE, W. F. (1941). A two variate gamma-type distribution, *Sankhyā*, **5**, 137-150.

KIDD, J. B. (1987). A comparison between the VERT program and other methods of project duration estimation. *OMEGA*, **15:3**, 129-134.

KLEIN, J. H. (1993). Modelling risk trade-off. *J. Opl Res. Soc.*, **44:5**, 445-460.

KLEIN, J. H., POWELL, P. L. and CHAPMAN, C. B. (1994). Project Risk Analysis Based on Prototype Activities. *J. Opl Res. Soc*, **45:7**, 749-757.

KLIR, G. J. and YUAN, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, New Jersey.

KOTTAS, J. F. and LAU, H. (1978). Stochastic breakeven analysis, *J. Opl. Res. Soc.*, **29:3**, 251-258.

KOTZ, S., JOHNSON, N. L. and READ, C. B. (1983). *Encyclopedia of Statistical Sciences*, Vol. 6, Chapter Multivariate Distributions, Wiley.

KULKARNI, V. E. and ADLAKHA, V. G. (1986). Markov and Markov-regenerative PERT networks. *Opns. Res.*, **34:5**, 769-781.

LEE, S. M., MOELLER, G. L. and DIGMAN, L. A. (1982) *Network Analysis for Management Decisions*, Kluwer-Nijhoff, Boston.

LOWRANCE, W. W. (1976). *Of Acceptable Risk.* William Kaufmann, Los Altos, California.

LUKASZEWICZ, J. (1965). On the estimation of errors introduced by standard assumptions concerning the distribution of activity duration in PERT calculations. *Opns. Res*, **13**, 326-327.

MacCRIMMON, K. R. and RYAVEC, C. A. (1964). An analytical study of PERT assumptions. *Opns. Res.*, **12**, 16-37.

MacCRIMMON, K. R. and WEHRUNG, D. A. (1986). *Taking Risks: The Management of Uncertainty*, Collier Macmillan, London.

MACLAREN, M. D., MARSAGALIA, G., and BRAY, T. A. (1964). A fast procedure for generating exponential random variables. *Commum. Assoc. Comput. Mach.*, **7**, 298-300.

MARES, M. (1989). Network analysis of Fuzzy technologies. *Applications of Fuzzy Set Methodology in Industrial Engineering*, 115-125.

MARES, M. and HORAK, J. (1983). Fuzzy quantiles in networks. *Fuzzy Sets and Systems*, **12**, 123-134.

MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Ass*, **62**, 30-44.

MARTIN, J. J. (1965). Distribution of time through a directed acyclic network. *Opns. Res.*, **13:1**, 46-66.

MATHAI, A. M. (1982). Storage capacity of a dam with gamma type inputs, *Annals of the Institue of Statistical Mathematics*, **34**, 591-597.

MATHUR, K. and McGEORGE, D. (1991). An integrated decision making environment for cost vs quality control. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

MIL-STD-721C (1981). Definitions of Terms for Reliability and Maintainability, Commanding Officer, Engineering Specifications and Standards Department (Code 93), Naval Air Engineering Center, Lakehurst, NJ 08733.

MODER, J. J., PHILLIPS, C.R., and DAVIS, E. W. (1983). *Project Management with CPM, PERT and Precedence Diagramming*, Van Nostrand Reinhold, London.

MOELLER, G. L. (1972). VERT. Technical papers presented at the *Twenty-Third Institute of Industrial Engineers*, September 20-23, Anaheim, California.

MOSCHOPOULOS, P. G. (1985). The distribution of the sum of independent gamma random variables, *Annals of the Institute of Statistical Mathematics*, **37**, 541-544.

MURDOCH, J. and BARNES, J. A. (1986). *Statistical Tables for Science, Engineering, Management, and Business Studies*, 3rd ed., Macmillan, London.

MURRAY, J. E. (1962). Consideration of PERT assumptions. *Technical Report*, Condructron Corporation, Ann Arbor.

NAG (1990). *The NAG Fortran Library Manual, Mark 14*, The Numerical Algorithms Group Limited, Oxford.

NERO, J. M. G. (1991). A quantitative model for the evaluation of construction quality. *Trans. of the European Symposium on Management, Quality and Economics in Housing and other Building sectors*, Lisbon, Portugal.

NICHOLAS, J. M. (1990). *Managing Business and Engineering Projects*, International Editions, Prentice Hall.

OBERHETTINGER, F. and BADII, L. (1973). *Tables of Laplace transforms*, Springer-Verlag, New York.

O'CONNOR, P. (1991). *Practical Reliability Engineering*, Wiley, Chichester.

PAYNE, B. J. (1993). *Risk Analysis in Project Planning*, BSc Dissertation, The Nottingham Trent University.

PAYNE, B. J. (1994). Project Risk Analysis: A Discussion. *Royal Statistical Society*, International Conference, Newcastle Upon Tyne.

PEARSON, E. S. and HARTLEY, H. O. (1976). *Bimetrika Tables for Statisticians*, Vol. 1 and 2, Charles Griffin.

PRITSKER, A. A. B., and BURGESS, R. R. (1970). *The GERT simulation programs*, Dept. of Industrial Engineering, Virginia Polytechnic, Blacksburg.

RAGSDALE, C. (1989). The current state of network simulation in project management theory, *OMEGA*, **17:1**, 21-25.

RESCHER, N. (1983). *A Philosophical Introduction to the Theory of Risk evaluation and management*, University Press of America.

RINGER, L. G. (1969). Numerical operators for statistical PERT critical path analysis. *Management Science*, 136-143.

ROWE, W. D. (1977). *An Anatomy of Risk.* Wiley.

SCOTT, G. (1976). *Building Disasters and Failures*, The Construction Press Ltd., The Blackburn Times Press.

SCULLI, D. (1983). The completion time of PERT networks. *J. Opl. Res. Soc.*, **34:2**, 155-158.

SCULLI, D. and WONG, K. L. (1985). The Maximum and Sum of Two Beta Variables and the Analysis of PERT Networks, *OMEGA Int. J. of Mgmt Sci.*, **13:3**, 233-240.

SEARS, R. W. Jr (1991). A model for managing the cost of unreliability. *Proceedings of Annual Reliability and Maintainability Symposium.*

SESHADRI, V. and PATIL, G. P. (1964). A characterisation of a bivariate distribution by the marginal and conditional distributions of the same component. *Annals of the Institute of Statistical Mathematics*, **15**, 215-221.

SIM, C. H. (1992). Point process with correlated gamma interarrival times, *Statistics & Probability Letters*, **15**, 135-141.

SMITH, K. and Keenan, D. J. (1979). *English Law*, Sixth Edition, Pitman, Bath.

SOROUSH, H. (1993). Risk Taking in Stochastic PERT networks. *European J. of Opl. Res.*, **67**, 221-241.

TAGUCHI, G. and WU, Y.-I. (1979). *Introduction to Off-Line Quality Control*, Central Japan Quality Control Association.

THE ROYAL SOCIETY (1992). *Risk: analysis, perception, management*, Report of Royal Society study group, The Royal Society, London.

THOMAS, L. (1992). Financial Risk Management Models. Chapter four of Ansell and Warton (1992).

TURNER, J. R. (1994). Editorial, *Int. J. of Project Manager*, **12:4**, 203-204.

UK Defence Standard 00-40: *The Management of Reliability and Maintainability*, HMSO.

US MIL-STD-1629: *Failure Mode and Effects Analysis.* Available from the National Technical Information Service, Sprongfield, Virginia.

VAN SLYKE, R. M. (1963). Monte Carlo methods and the PERT problem, *Opns. Res.*, **11**, 839-860.

WARD, S. C., CURTIS, B. and CHAPMAN, C. B. (1991). Objectives and performance in construction projects. *Construction Mgmt and Econ*, **9**, 343-353.

WARNER, F. (1992). *Risk; analysis, perception, management.* Study group report, The Royal Society.

WHARTON, F. (1992). *Risk Management: Basic Concepts and General Principles*, Ansell and Wharton (1992), Chapter 1, 1-14.

WELSH, P. J. A. (1965). Errors introduced by a PERT assumption. *Opns. Res.*, **13**, 141-143.

WIGHTMAN, D. and BENDELL, A. (1986). The Practical Application of Proportional Hazards Modelling, *Reliability Engineering*, **15**, 29-53.

WILLIAMS, T. M. (1992). Practical use of distributions in network analysis, *J. Opl Res. Soc.*, **43**, 265-270.

WOLFRAM, (1988). *A system for doing mathematics by computer.* Wolfram Research Inc. (Software)

WOLFRAM, S. (1991). *Mathematica: A system for Doing Mathematics by Computer.* Addison-Wesley, Wokingham, (User Guide)

ZADEH, L. A. (1965). Fuzzy Sets, *Information and Control*, **8:3**, 338-353.

APPENDIX 1

Review of Risk Analysis Software

Risk Analysis Software: Summary and Comments

Summary

There are essentially two types of software available for performing a risk analysis, namely stand alone packages and add-ons. The majority of stand alone packages, such as Risknet are the result of specialised research and tend to be for specific needs. On the whole stand-alone packages are relatively expensive compared to add-on packages as support and training is often necessary.

On the other hand add-on (or add-in) packages are relatively cheap and provide the facility of modelling uncertainty with Monte Carlo simulation in project planning software or for spreadsheets.

We provide a list with details of common software utilised in industry which is a result of personal communication with suppliers and where possible an evaluation of the software. For further details of software refer to the internet address given in the Introduction of Chapter 1.

Add-On Packages to Perform a Risk Analysis

Risk (spreadsheet add-on) There are currently two versions available where there are over 37 probability distributions to choose from.

1. Version RISK 3.5 for MS Project. Results may be customised in EXCEL
2. Version RISK 3.5 (32 bit) runs in EXCEL or 123 for Windows 95.

CRYSTAL BALL 4 (spreadsheet add-on) Uses EXCEL or Lotus 123. Choice of distributions. Dependency is possible using the correlated assumptions tool i.e. profit is dependent on raw material. Analysis of percentiles is performed on p.d.f. plot and not on c.d.f. See review of Levy (1998). Users include Shell and British Petroleum.

RISK+ (spreadsheet/MS Project-add-on) Enables milestones points to be analysed, criticality indices to be calculated. Reviewed by 'Project Manager Today' (March 1996), as an 'excellent introduction to dealing with estimating risk'.

Stand Alone Packages

Predict! General purpose risk analysis package. A spreadsheet front end compatible with lotus 123. Choice of 26 probability distributions and own 'wobbly house' distribution. Output as p.d.f and c.d.f. Uses include, financial investment risk. Construction users include Bovis, AMEC and Balfour Beattie.

Predict! Risk Analyser (P ! RA) Monte Carlo analysis tool with spreadsheet functionality.

Predict! Risk Controller Is a risk register that provides for the identification, assessment and management of risks. (Not a simulation package)

Project Risk Not based on a spreadsheet, but requires a profile risk profile to be constructed from a series of elicit questions regarding strategic, business, project size, planning, user, technical development and operations issues. The user is asked to rank the risks as avoidable, manageable, unavoidable and not known. Net result is that the program calculates the risk for specific issues and for the overall project. Aimed at analysing the risks associated with developing bit IT projects

Risknet Produced by the consultancy YARD of BAeSema in analysing the risks of developing military aircraft (originated as in-house). Spreadsheet modelling environment. Uncertainty is modelled as triangular distributions. Geared towards large development projects of the government type contract.

RiskMaster Addresses the quantitative and qualitative aspects of risk. Allows conditional branching.

Primavera Project Planner (P3) version 2.0 with Monte Carlo 3 P3 is a schedule and resource resource control package. Designed to handle large scale projects up to 100,000 activities. Uses Monte Carlo to perform an in-depth project risk analysis. Results are based on plots as used by the consultancy group EUROLOG where time-cost probability charts are presented. Concentrates on time and cost aspects. It is a widely used professional package within the UK and was recently used by Bektol to manage the construction of Hong Kong airport. Other users include:- British Airways, Barclays technical services, GPT, Railtrack, Post Office Counters and Robert McAlpine.

REMIS Enables the identification, assessment and management of risk. It is referred to as a Top Down Risk Model (TDRM). Allows consequences to be modelled as impact scales for time, cost and performance. Does not use simulation to compute overall risks.

Open Plan A project management system that allows for Monte Carlo simulations.

Comments

Our main area of concern was that many of the packages, specifically the add-ons for spreadsheet packages give the impression of being able to do all types of risk analysis. This is not true, as most packages are only suited to modelling problems with few variables, such as the risk model of Hertz (1964). Representation of a project network is possible however we found no standard method to model risks associated with project networks. From experience and our investigation we identify the advantages and limitations of using a spreadsheet environment.

Advantages of Spreadsheets

- It is not necessary to have a detailed knowledge of simulation or modelling techniques.
- Software is cheap and requires minimum computer hardware and training.

Limitations of Spreadsheets

- The screen bears little relation to the problem under analysis (PC 1992)
- A tricky aspect is that the user must be careful in location of calculations as affects the order of evaluation, however one can specify to have the model evaluated by row or column first.
- In order to reduce simulationtime we found that some suppliers of software recommend to simulate 1000 cycles, which is contrary to the belief of Van Slyke (1963) who recommends 10,000 cycles. In such cases, accuracy may be lost, however it is unclear if stratified sampling is used.

With regards to stand-alone packages we found that some packages provide the facility to model forms of dependency such as consequence analysis of costs and reaction to external states. However we found no software that analyses the dependency between the operation of the project outcome and the performance of the project. Also dependency between time, cost and quality aspects of a project are sometimes mentioned however no common model is apparent.

APPENDIX 2

Network Reduction Algorithm

Running the program

The data requirements are a network structure of the k events and Special Erlangian parameters, (ρ_i, α_i) for the q activity time distributions.

The network is represented as a $k-1 \times k$ matrix where entry $n_{i,j}$ denotes the activity number (Note. A zero indicates no activity) connecting event i to event j. Thus,

$$network = \{\{n_{1,1}, n_{1,2}, \ldots, n_{1,k}\}, \{n_{2,1}, \ldots, \{n_{k-1,1}, \ldots, \{n_{k-1,k}\}\}$$

The q activities are entered as a 2 dimensional vector.

$$activity = \{\{\rho_1, \alpha_1\}, \{\rho_2, \alpha_2\}, \ldots, \{\rho_q, \alpha_q\}\}$$

The Network reduction procedure is executed by typing,

report[network, activity]

For further information concerning the commands and data structures within Mathematica see Wolfram (1991).

```
(*: Network Collapsing algorithm: 1.1 *)
(*: Author: Brad Payne *)
(*: Date: 7th May 1996 *)

(* Determines the Laplace transform of a special erlangian with parameters
 p,m where m is a positive integer *)

conv[activity_List]:=
Block[{p,m,return},
p=activity[[1]];
m=activity[[2]];
return=(p^m)/(s+p)^m
]


(* Inverts the result of the Laplace Transform of the previous node
(if exists) multiplied by the Laplace transform of the connecting activity
i.e. performs the convolution of two random variables *)

sum[input_List,act_List,output_List]:=
Block[{loutpt,lerlang,return},
If[output[[input[[1]]]]==0,loutpt=1,
loutpt=dummy,loutpt=LaplaceTransform[output[[input[[1]]]],t,s]];
lerlang=conv[act[[input[[2]]]]];
return=InverseLaplaceTransform[loutpt*lerlang,s,t]
]



(* Returns the pdf of the greatest of 'len' paths leading into a node *)

max[input_List]:=
Block[{i,j,len,return},
Clear[i];Clear[j];
len=Length[input];
i=1;j=1;
return=Sum[input[[i]]*Product[If[j!=i,(Integrate[input[[j]],{t,0,t}])
,1],{j,1,len}],{i,1,len}]
]

comp[l_,cdf_]:=
Block[{t,output},
x=0;
output=Table[i-i,{i,32},{i,2}];
While[x<=32,t=x*0.25;output[[x]]={t,cdf};
x++;
]
```

```
Print["output is ",output]
]
```

(* Prints the pdf and cdf, plots the cdf, and calls a tabulation procedure (comp) to generate cumulative probability values for time points at equal intervals between 0 and 8, (specific to Dodin[4] example where all time distributions are of identical exponential). The exponential parameter is set within the procedure by the variable 'l'. *)

```
result[pdf_]:=
Block[{cdf},
Print["pdf is ",Simplify[pdf]];
cdf=Integrate[pdf,{t,0,t}];
Plot[cdf,{t,0,200},PlotRange->All];
Print["cdf is ",Simplify[cdf]];
l=2;
comp[l,cdf]
]
```

(* Control of the algorithm, scans the network matrix from top to bottom and left to right. Counts the number of paths into a given node j then if only one performs call the 'sum' procedure otherwise generates the input of summations for the 'max' procedure. Gives an output vector of length N, where Output[j] is the realisation till node j. *)

```
report[net_List,act_List]:=
Block[{count,cdf,i,j,len,store1,input,output,maxcount,expand},
len=Length[net]+1;
count=Table[i-i,{i,len}];
store1=Table[i-i,{i,len-1},{i,2}];
output=Table[i-i,{i,len}];
j=2;
While[j<len+1,
i=1;
While[i<len,
If[net[[i,j]]>0,count[[j]]++;
store1[[count[[j]]]]={i,net[[i,j]]}];
i++];
If[count[[j]]==1,Print["rec sum"];
output[[j]]=sum[store1[[1]],act,output],
input=Table[sum[store1[[maxcount]],act,output],
{maxcount,1,count[[j]]}];
output[[j]]=max[input]];
cdf=Integrate[output[[j]],{t,0,200}];
Print["integral for output ",j," is ",cdf];
```

```
j++];
expand=Simplify[ExpandAll[output[[j-1]]]];
result[expand]
]
```

APPENDIX 3

Project Risk Dependency (PRD) Simulation Software

```fortran
      program PRDsim
c
c*********************************
c
c     Project Risk Dependency simulator
c     Version 1.18
c     Date: Sept 1997 (modifed August 1998)
c
c*********************************
c
c
c     ** sets the max size of network and characteristics **
c     ** 100 events, 3 characteristics, 1000 maxiters **
c     ** NB: parameters will need changing for larger configurations **
c
      integer        maxevent,maxchar,maxiters,maxact
      parameter     (maxevent=4,maxchar=3,maxiters=10000)
      parameter     (maxinter=10,maxnanaly=5)
c
c     provides space for the subroutine local vars.
c
      integer       count(maxevent)
      integer       store(maxevent,maxevent,2)
      integer       inrel(maxchar)
      double precision          result(maxchar)
      double precision          parinput(maxevent,maxchar)
      double precision          coldata(maxevent)
      double precision          shapea(maxchar)
      double precision          meana(maxchar)
      double precision          intraa(maxchar,maxchar)
      double precision          realisation(maxchar)
      double precision          columnlist(maxchar)
      double precision          ireal(maxchar)
      integer       depcheck(maxchar,maxchar)
      integer       columncheck(maxchar)
      integer       ncharr
c
c     Initialises variables for data input and output
c
      character*12  filename
      integer       nevents,nact,network(maxevent,maxevent)
      integer       neactd
      integer       nchar,parop(maxchar),iters,nanaly
      integer       fromto(maxnanaly,2)
      parameter     (maxact=maxevent)
      double precision          intra(maxact,maxchar,maxchar
      double precision          shape(maxact,maxchar)
```

```
      double precision          meantemp(maxact,maxchar)
      double precision          mean(maxact,maxchar)
      double precision          output(maxiters,maxevent,maxchar)
      double precision          target(maxact,maxchar)
      double precision          inter(maxinter,5)
c
c     initialises variables for functions and subroutines
c
c     **  external functions and NAG routines **
c
c     g05fff,g05ccf variate generation
c     g01aaf descriptive statistics
      external      g05fff,g05ccf,g01aaf,analysis,
     +              input,start,check1,check2,
     +              paramcollect,remain,compcontrol,sum,simulate,
     +              parallel,fnmax,fnplus,fnmin,fnavg,variate
c
c     **  common block definitions **
c
      common/countblk/count
      common/storeblk/store
      common/resultblk/result
      common/parinputblk/parinput
      common/coldatablk/coldata
      common/shapeblk/shape
      common/meantempblk/meantemp
      common/meanblk/mean
      common/intrablk/intra
      common/shapeablk/shapea
      common/meanablk/meana
      common/intraablk/intraa
      common/realisationblk/realisation
      common/columnlistblk/columnlist
      common/networkblk/network
      common/paropblk/parop
      common/irealblk/ireal
      common/depcheckblk/depcheck
      common/columncheckblk/columncheck
      common/outputblk/output
      common/ncharrblk/ncharr
      common/targetblk/target
      common/interblk/inter
      common/fromtoblk/fromto
      common/inrelblk/inrel
c
c
c************************************
```

```
c     MAIN PROGRAM
c
      read'(a12)',filename
      open(9,file=filename)
      read(9,*)nevents
      read(9,*)nact
      read(9,*)nchar
      read(9,*)neactd
      read(9,*)iters
c
c     ** analysis data **
c
      read(9,*)nanaly
c
c     ** input the remaining data **
c
      call input(nevents,nchar,nact,neactd,nanaly)
c
c     ** call  the control engine of the simulation **
c
      call simulate(nevents,nchar,iters,nact,neactd)
      write(*,*)'nanaly is ',nanaly
      if(nanaly.gt.0)then
         call analysis(iters,nchar)
      endif
c
      stop
c
c
      end
c
c
c     END OF MAIN
c***********************************
c***********************************
c     THE SUBROUTINES
c
      subroutine analysis(iters,nchar)
c     ** analysis of results**
c
      parameter(maxiters=10000,maxevent=4,maxchar=3,maxnanaly=5)
      double precision      output(maxiters,maxevent,maxchar)
      double precision      s2,s3,s4,wtsum,xbar,xmax,xmin,wt(maxiters)
      double precision      datac(maxiters)
      integer   loop,i,j,ifail,iwt,iters
      integer   fromto(maxnanaly,2)
      common/fromtoblk/fromto
```

```
      common/outputblk/output
      external g01aaf
c
c     ** checks requirement of analysis results **
c     ** constructs data set to call analysis routine **
c
      ifail=0
      iwt=0
c
c     ** opens file 'res' to save simulated characteristic values **
c
      open(2,FILE='res')
c
         do 10 i=1,nanaly+1
            write(*,*)'ANALYSIS OF CHOICE',i
    do 15 j=1,nchar
            do 20 loop=1,iters
               from=fromto(i,1)
              to=fromto(i,2)
              datac(loop)=output(loop,to,j)-output(loop,from,j)
                     wt(loop)=1
c
c   ** stores simulated characteristic data in file
c
                     write(2,*)datac(loop)
 20           continue
c
c  ** calls NAG routine - descriptive statistics **
c
                  call g01aaf(iters,datac,iwt,wt,xbar,s2,s3,s4,xmin,
     + xmax,wtsum,ifail)
                  write(*,*)'ifail',ifail
c
                  write(*,*)'Completion of analysis of component ',j
                  write(*,*)'Mean           ',xbar
                  write(*,*)'Std. devn      ',s2
                  write(*,*)'Skewness       ',s3
                  write(*,*)'Kurtosis       ',s4
                  write(*,*)'Minimum        ',xmin
                  write(*,*)'Maximum        ',xmax
 15          continue
 10          continue
       close(2)
       return
       end
c
c
```

```
c
c
      subroutine simulate(nevents,nchar,iters,nact,neactd)
c     ** input variables **
c
      integer   nevents,nchar,iters,nact,neactd
c
c     ** local variables **
c
      integer   maxevent,maxchar,maxiters
      parameter (maxevent=4,maxchar=3,maxiters=10000)
      parameter (maxact=maxevent,maxinter=10)
      integer   eventoutchk(maxevent)
      integer   count(maxevent),loop,i,j,x,y,k,l
      integer   store(maxevent,maxevent,2)
      double precision      result(maxchar)
      double precision      parinput(maxevent,maxchar)
      integer   event,compf,activity,compt,param
      double precision      comptarg,compreal
      double precision      exceed
c
c     ** comon block variables **
c
      integer   network(maxevent,maxevent)
      integer   parop(maxchar)
      double precision      output(maxiters,maxevent,maxchar)
      double precision      meantemp(maxact,maxchar)
      double precision      mean(maxact,maxchar)
      double precision      inter(maxinter,5)
      double precision      target(maxact,maxchar)
      common/networkblk/network
      common/paropblk/parop
      common/countblk/count
      common/storeblk/store
      common/resultblk/result
      common/parinputblk/parinput
      common/outputblk/output
      common/targetblk/target
      common/meanblk/mean
      common/meantempblk/meantemp
      common/interblk/inter
c
c     ** initialises network structure **
c     ** activity dependency and between **
c     ** activity dependency **
c
      do 100 j=1,nevents
```

```
        do 110 i=1,nevents-1
           if(network(i,j).gt.0)then
              count(j)=count(j)+1
              store(j,count(j),1)=i
              store(j,count(j),2)=network(i,j)
           endif
 110    continue
c
c     ** between activity dependency **
c     ** 1=dependency exists, 0=no dependency **
c
        do 105 l=1,neactd
           If(inter(l,1).gt.0)then
              eventoutchk(inter(l,1))=1
           endif
 105    continue
 100    continue
      do 120 loop=1,iters
c
c     **  resests all mean params to input data, (meantemp) **
c
        do 125 k=1,nact
           do 127 p=1,nchar
              mean(k,p)=meantemp(k,p)
 127       continue
 125    continue

        do 130 j=1,nevents
           if(count(j).eq.1)then
              call sum(store(j,1,1),store(j,1,2),nchar,loop)
           endif
           if(count(j).gt.1)then
              do 140 x=1,count(j)
                 call sum(store(j,x,1),store(j,x,2),nchar,loop)
                 do 150 y=1,nchar
                    parinput(x,y)=result(y)
 150             continue
 140          continue
              call parallel(count(j),nevents,nchar)
           endif
c
c     ** output assignment **
c
      if(count(j).gt.0)then
      do 160 y=1,nchar
         output(loop,j,y)=result(y)
 160  continue
```

```
c
c     ** performs parameter update for between event/activity **
c     ** dependency **
c
      if(eventoutchk(j).eq.1)then
         do 170 i=1,neactd
            event=inter(i,1)
            if(event.eq.j)then
               compf=inter(i,2)
               activity=inter(i,3)
               compt=inter(i,4)
               param=inter(i,5)
               comptarg=target(activity,compf)
               compreal=output(loop,j,compf)
               exceed=compreal-comptarg
c
c     ** if exceeded a component target we shall **
c     ** assume consequences are encountered in   **
c     ** future activities. Otherwise no change   **
c
               if(exceed.gt.0)then
                  mean(activity,compt)=mean(activity,compt)+
     + param*exceed
               endif
            endif
 170  continue
      endif
c
c     ** end of (eventoutchk(j).eq.1 **
c
      endif
c
c     ** end of count(j).gt.0 **
c
 130  continue
 120  continue
      return
      end
c
c
c
c
      subroutine parallel(count,nevents,nchar)
c    ** performs parallel operation **
c
      integer    count,nevents,nchar,maxevent,maxchar
      parameter  (maxevent=4,maxchar=3)
```

```
      integer    parop(maxchar)
      double precision       parinput(maxevent,maxchar)
      double precision       result(maxchar)
c
      integer   i,j
      double precision      coldata(maxevent)
c
c     ** external functions **
c
      external   fnmax,fnmin,fnplus,fnavg
      common/paropblk/parop
      common/parinputblk/parinput
      common/resultblk/result
      common/coldatablk/coldata
c
c     ** collects the inputs into an event into coldata **
c
      do 200 j=1,nchar
         do 210 i=1,count
            coldata(i)=parinput(i,j)
 210  continue
c
c     ** decides which parallel operation to perform **
c
      if(parop(j).eq.1)then
          result(j)=fnmax(coldata,count)
      endif
      if(parop(j).eq.2)then
         result(j)=fnplus(coldata,count)
      endif
      if(parop(j).eq.3)then
         result(j)=fnmin(coldata,count)
      endif
       if(parop(j).eq.4)then
            result(j)=fnavg(coldata,count)
         endif
 200   continue
       return
       end
c
c
c
c
      subroutine sum(from,act,nchar,loop)
c     ** performs the sum operation **
c
      integer   from,act,nchar,loop
```

```
c
      integer    maxevent,maxchar,maxiters,maxact
      parameter (maxevent=4,maxchar=3,maxiters=10000,
     + maxact=maxevent)
      double precision   shapea(maxchar)
      double precision   meana(maxchar)
      double precision   intraa(maxchar,maxchar)
      double precision   realisation(maxchar)
      double precision   result(maxchar)
      integer i,j,m
c
      integer ncharr
      integer inrel(maxchar)
      double precision   output(maxiters,maxevent,maxchar)
      double precision   intra(maxact,maxchar,maxchar)
      double precision   shape(maxact,maxchar)
      double precision   mean(maxact,maxchar)
      common/resultblk/result
      common/shapeablk/shapea
      common/meanablk/meana
      common/intraablk/intraa
      common/realisationblk/realisation
      common/outputblk/output
      common/intrablk/intra
      common/shapeblk/shape
      common/meanblk/mean
      common/ncharrblk/ncharr
      common/inrelblk/inrel
c
      do 300 i=1,nchar
         realisation(i)=0
         inrel(i)=0
         shapea(i)=shape(act,i)
         meana(i)=mean(act,i)
         do 310 j=1,nchar
            intraa(i,j)=intra(act,i,j)
 310     continue
 300  continue
      do 315 i=1,nchar
         do 317 j=1,nchar
            if(intraa(i,j).gt.0)then
               inrel(j)=1
            endif
 317  continue
 315  continue
      call start(nchar)
      if(ncharr.lt.nchar)then
```

```
         call remain(nchar)
      endif
c
      do 320 m=1,nchar
         result(m)=output(loop,from,m)+realisation(m)
 320  continue
      return
      end
c
c
c
c
      subroutine start(nchar)
c    ** simulates independent characteristics **
c
      integer   j,nchar,maxchar,ncharr
      parameter (maxchar=3)
      double precision      shapea(maxchar),meana(maxchar)
      double precision      realisation(maxchar)
      integer inrel(maxchar)
c
      external variate
c
      common/shapeablk/shapea
      common/meanablk/meana
      common/realisationblk/realisation
      common/ncharrblk/ncharr
      common/inrelblk/inrel
c
      ncharr=0
      do 100 j=1,nchar
         if(meana(j).gt.0.and.inrel(j).eq.0)then
            realisation(j)=variate(shapea(j),meana(j))
            ncharr=ncharr+1
         endif
 100  continue
      return
      end
c
c
c
c
      subroutine check1(nchar)
c    ** creates a check matrix to identify characteristics **
c    ** that need simulating **
c
      integer   i,j,nchar,maxchar
```

```
      parameter (maxchar=3)
      integer   depcheck(maxchar,maxchar)
      double precision      intraa(maxchar,maxchar)
      double precision      realisation(maxchar)
c
      common/depcheckblk/depcheck
      common/intraablk/intraa
      common/realisationblk/realisation
c
      do 105 i=1,nchar
 105  continue
      do 110 j=1,nchar
         do 120 i=1,nchar
            if(intraa(i,j).ne.0.and.realisation(j).eq.0)then
               depcheck(i,j)=1
               else
               depcheck(i,j)=0
            endif
 120     continue
 110  continue
      return
      end
c
c
c
c
      subroutine check2(nchar)
c    ** uses the result of check 1 and  realised values to identify **
c    ** which characteristics may now be simulated  **
c
      integer   i,j,nchar,maxchar
      parameter (maxchar=3)
      integer   depcheck(maxchar,maxchar)
      integer   count,columncheck(maxchar)
c
      common/depcheckblk/depcheck
      common/columncheckblk/columncheck
c
      do 200 j=1,nchar
         count=0
         do 210 i=1,nchar
            if(depcheck(i,j).gt.0)then
               count=count+1
            endif
 210     continue
         columncheck(j)=count
 200  continue
```

```
      return
      end
c
c
c
c
      subroutine paramcollect(columnlist,nchar,column,meanc)
c    ** calculates the characteristic mean value, incorporating **
c    ** all dependencies  **
c
      integer  nchar,i,column,maxchar
      parameter (maxchar=3)
      double precision     columnlist(nchar)
      double precision     meana(maxchar)
      double precision     realisation(maxchar),meanc
      double precision     intraa(maxchar,maxchar)
c
      common/intraablk/intraa
      common/realisationblk/realisation
      common/meanablk/meana
c
      meanc=meana(column)
      do 300 i=1,nchar
         if(columnlist(i).ne.0)then
            meanc=meanc+realisation(i)*intraa(i,column)
         endif
 300  continue
      return
      end
c
c
c
c
      subroutine remain(nchar)
c    ** controls the remaining simulations after initial simulations  **
c    ** have been performed with subroutine start **
c
      integer  nchar,maxchar
      parameter (maxchar=3)
      double precision     shapea(maxchar),intraa(maxchar,maxchar)
      double precision     meanc
      double precision     ireal(maxchar)
      double precision     realisation(maxchar)
      double precision     columnlist(maxchar)
      integer  columncheck(maxchar)
      integer  column
      integer  i,j,k
```

```
c
      common/shapeablk/shapea
      common/intraablk/intraa
      common/irealblk/ireal
      common/realisationblk/realisation
      common/columncheckblk/columncheck
c
      call check1(nchar)
      call check2(nchar)
c
c     ** new remain subroutine **
c
      do 400 j=1,nchar
         if(columncheck(j).gt.0)then
            k=0
            do 410 i=1,nchar
               if(intraa(i,j).ne.0.and.columncheck(i).eq.0)then
                  k=k+1
                  columnlist(k)=intraa(i,j)
               endif
 410        continue
            if(k.eq.columncheck(j))then
               column=j
               call paramcollect(columnlist,nchar,column,meanc)
               if(meanc.lt.0)then
                  meanc=0.1
               endif
               realisation(j)=variate(shapea(j),meanc)
               call check1(nchar)
               call check2(nchar)
            endif
         endif
 400  continue
      return
      end
c
c
c
c
      subroutine input(nevents,nchar,nact,neactd,nanaly)
c    ** inputs remaining data file values **
c
      integer     nevents,nchar,nact,neactd
c
      integer        maxevent,maxchar,maxiters,maxact
      parameter     (maxevent=4,maxchar=3,maxiters=10000)
      parameter     (maxact=maxevent,maxinter=10,maxnanaly=5)
```

```
      integer     i,j,k
c
c     ** block variables **
c
      integer  network(maxevent,maxevent)
      double precision     intra(maxact,maxchar,maxchar)
      double precision     shape(maxact,maxchar)
      double precision     meantemp(maxact,maxchar)
      double precision     target(maxact,maxchar)
      double precision     inter(maxinter,5)
      integer  parop(maxchar),nanaly
      integer  fromto(maxnanaly,2)
c
      common/networkblk/network
      common/shapeblk/shape
      common/meantempblk/meantemp
      common/intrablk/intra
      common/paropblk/parop
      common/targetblk/target
      common/interblk/inter
      common/fromtoblk/fromto
c
      do 10 i=1,nevents
         read(9,*)(network(i,j),j=1,nevents)
 10   continue
c
c     ** reads component coefficients for every activity **
c
      do 20 k=1,nact
         do 30 i=1,nchar
            read(9,*)(intra(k,i,j),j=1,nchar)
 30      continue
 20   continue
c
c     **  reads the shape parameters   **
c
      do 40 k=1,nact
         read(9,*)(shape(k,i),i=1,nchar)
 40   continue
c
c     **  reads the mean parameters   **
c
      do 50 k=1,nact
         read(9,*)(meantemp(k,i),i=1,nchar)
 50   continue
c
c     **  reads the parallel operations   **
```

```
c
         read(9,*)(parop(i),i=1,nchar)
c
c     **  reads the target values for each comp. of each activity  **
c
      do 60 k=1,nact
         read(9,*)(target(k,i),i=1,nchar)
 60   continue
c
c     **  If event dependency is set **
c     **  reads the event activity dependency data **
c     **  event,component,act,component,parameter   **
c
      if(neactd.gt.0)then
         do 70 k=1,neactd
            read(9,*)(inter(k,i),i=1,5)
 70      continue
      endif
c
c      ** if nanalysis has been set, **
c      ** reads the analysis from to events **
c
      if(nanaly.gt.0)then
         do 80 k=1,nanaly
            read(9,*)(fromto(k,i),i=1,2)
 80      continue
      endif
c
      endfile 9
      close(9)
      return
      end
c
c     END OF SUBROUTINES
c********************************
c********************************
c     THE FUNCTIONS
c
      function variate(shap,mu)
c     ** Generation of a gamma variate **
c
      integer   ifail
      double precision      shap,mu,scal,var
      external  g05fff,g05ccf
c
      scal=mu/shap
c      call g05ccf
```

```fortran
      call g05fff(shap,scal,1,var,ifail)
c
      variate=var
      return
      end
c
c
c
c
      function fnmin(coldata,count)
c     ** identifies the minimum of list of data, **
c     ** labelled coldata of length count **
c
      integer count,i
      double precision    coldata(count),temp
c
      temp=coldata(1)
      do 100 i=2,count
         if(coldata(i).lt.temp)then
            temp=coldata(i)
         endif
 100  continue
      fnmin=temp
      return
      end
c
c
c
c
      function fnmax(coldata,count)
c     **  identifies the maximum of list of data, **
c     **  labelled coldata of length count **
c
      integer count,i
      double precision    coldata(count),temp
c
      temp=coldata(1)
      do 100 i=2,count
         if(coldata(i).gt.temp)then
            temp=coldata(i)
         endif
 100  continue
      fnmax=temp
      return
      end
c
c
```

```
c
c
function fnavg(coldata,count)
c     **  identifies the average of list of data **
c
      integer count,i
      double precision    coldata(count),temp
c
      temp=0
      do 100 i=1,count
         temp=temp+coldata(i)
 100  continue
      temp=temp/count
      fnavg=temp
      return
      end
c
c
c
function fnplus(coldata,count)
c     ** sums a list of data  **
c
      integer count
      double precision coldata(count),temp
c
      temp=0
      do 100 i=1,count
         temp=temp+coldata(i)
 100  continue
      fnplus=temp
      return
      end
c
c
c   END OF FUNCTIONS
c************************************
```

APPENDIX 4

Project Risk Analysis: A Discussion

by

Payne, B., Carter, M., and Wightman, D. (1994).

Presentation at RSS Conference, Newcastle-Upon-Tyle, 1994.

Project Risk Analysis: A Discussion

BRAD J. PAYNE, J.MARK CARTER and DAVE WIGHTMAN
The Nottingham Trent University

INTRODUCTION

Project risk analysis is often viewed as an extension of PERT type techniques where the risk of a project overrunning can be assessed. PERT, a methodology developed in the sixties was developed for the problem of time scheduling in large projects; however time, is often not the only risk of interest associated with a project. Methodologies such as C/SPEC and C/SCSC[1] have been developed for modelling costs associated with a project which enables risks to be identified with project overspending. In practice the management of the measures time and costs are often treated separately which can limit the understanding and modelling of risks associated with the project. Although applied methods such as a I.P. formulation[2] are available for modelling the relationship and trade-offs between time and cost a general model has not been defined for performing a project risk analysis.

Another measure of interest often considered too complex to evaluate is *quality*. The outcome of the project whether it be a product, system or a service has to be of acceptable *quality*. Also it may have to perform the required task for the required period of time, i.e. it must have an acceptable *reliability*. The amount of effort in terms of time spent and money used in the project obviously affects these measures. Also the specified levels of quality and reliability can determine the project duration and the project cost. However the modelling of these and how the identification of the optimal levels can be recognised such that the risk is minimal for time, cost, quality and possibly reliability need to be analysed.

Measures of quality and reliability are not commonly used in the context of project management due to the complexity of their relationships to time and cost, unavailable data and lack of a common measure. The purpose of this paper is to review what a project is, what the risks are and indicate a common measure that will enable trade-offs to be evaluated between the risks associated with duration, cost, quality and reliability. The authors thus introduce their concept of a project life cycle that will enable modelling of certain factors that are often not considered in a project risk analysis.

PROJECT CHARACTERISTICS

A project can be thought of as an arrangement of tasks to achieve some predetermined end. It exists to develop a product, system or in general it is created in response to a problem. Nearly all projects can be divided into phases to indicate the type of tasks or activities to be conducted in some logical sequence.

Projects are performed in many businesses. They may involve plans and undertakings in the areas of Research & Development, Engineering programmes, Construction, or Production planning[3]. However each of these different project areas have a number of generic features and it is these common properties that we exploit in our formulation. Clearly, though, many features of a project are project specified; so we adopt as a common initial baseline the definition of a project given in BS6046[4].

- **No two projects are the same**

Projects between different business sectors will obviously be different, but it is also unlikely that any two projects performed in the same business area will be identical. For example a construction firm will experience different constraints for all jobs even though the required building is the same. Everything that could affect the outcome of the project has to be considered; the working conditions, the expertise, what can go wrong. All of these will vary for different projects.

- **A project is made up of phases**

For planning and control of a project, phases are identified which also enables monitoring of progress towards an often predetermined target.

- **A project contains uncertainty and risk**

Because of the one off nature of the project, and not being able to forecast future occurrences accurately a project will inevitably contain uncertainty. The risk we are concerned with is the risk of not meeting the targets for the project due to possible occurrences within the various project phases.

- **The defining, developing and manufacturing phases often affect the subsequent life of the product**

Often the project is considered complete when its outcome is achieved and released to the customer. However, in most cases maintenance and failure costs are still incurred after the hand over of the product itself. To include such costs when modelling and evaluating the project, it is necessary to model the entire span of the project from design to end of product life. We shall call this the *project life cycle.*

THE PROJECT LIFE CYCLE

Life Cycle modelling for products is well established[5] and the approach can be adapted to redefine the concept of a project. By considering a project life cycle to include the maintenance and support period, costs which are often overlooked can be included in analysis and decision making. In addition an overall picture is created where by all the important project management measures e.g duration, cost, quality, and reliability are incorporated. The modelling of the project life cycle can therefore help in the evaluation of design and development decisions at an early stage in the project.

Each project has a starting point and progresses towards a desired conclusion. We shall assume the project conclusion is reached when the product has survived the required period of time. This could be at the end of the life of a piece of machinery or even the life of a building. In almost all projects there are three measures of interest that can be identified. These are the time, cost and performance (quality/reliability), see Figure 1.

The Project Life Cycle Phases

As mentioned earlier all project life cycles will contain uncertainty, in terms of the duration, the costs and the quality and reliability. We shall indicate which measures can be controlled/traded off at the various project phases and the possible consequences of such trade-offs. Often there is overlap and interaction between phases, yet they are clearly definable. The four phases that make up the project life cycle are as follows:-

1. Conception & Definition Phase - WHAT? (is required)

2. Design & Development Phase - HOW? (is it to be achieved)

3. Acquisition Phase - ACHIEVING (the outcome)

4. Operation Phase - USING (the product)

With each of the phases we have attached a keyword.

The *Conception and Definition phase* includes deciding the objectives of the project. i.e. *What* needs to be achieved in order to satisfy the objectives. What reliability is required? What cost is acceptable? What is the time when the handover should be performed? Although it is anticipated some of the answers can only be estimated, a clear understanding of the project objectives at this phase will enable effective planning for the remaining project life cycle. The outcome of this phase is therefore to have targets for quality, reliability, handover time and the costs.

The Design and Development phase involves designing and planning the activities and establishing the project organisation that will achieve the project outcome i.e. How to achieve targets. It is at this stage where the trade-offs between the duration, cost, quality and reliability should be recognised and understood. We believe a risk analysis should be performed at this stage of the project life cycle since targets have already been established and an analysis of achieving the targets should be performed.

The *Acquisition phase* contains numerous activities associated with producing the product, and controlling its reliability and quality. The effort and resources used in performing each of these tasks will determine the cost, C_1, and the time, T_1, prior to handover. Plans designed in the design and development phase will determine the nature of the activities and the resources required to attain a quality, reliable product in the required period at the specified cost.

The *Operation phase* involves the outcome of the project, the product, being put into operation. It is assumed all costs incurred in this phase are borne by the project organisation. The project is only considered complete when the product has performed/survived the required period of time, T_2. Any costs incurred are classed as unreliability costs.

Description of Project measures

• Performance

Performance refers to product specification and requirements established for the output of the project. It can be expressed in terms of *product quality* and *product reliability* which are controlled during the phases before handover. Often a reliability programme is incorporated into quality costs[6] where the cost of achieving reliability consists of preventative and appraisal costs. In addition the cost of unreliability after product handover, once the product is in operation, is classed as an external cost of failure.

Because of the significance of reliability in a project life cycle, quality and reliability will be considered separately. Their associated costs will also be considered separately.

A quality product is one which meets the requirements/specification of the project outcome. To achieve quality, preventative, appraisal and failure costs will be incurred, which go to make up total quality costs.

$$\text{Quality} = f(\text{Quality Costs})$$

A reliable product is one which performs the required task for the required period of time.

Reliability is planned for and built into a product, therefore costs will be incurred in the development and testing of reliability. The cost of unreliability, C_2, will also have to be evaluated. This will be discussed further when determining the Project Life Cycle Cost.

$$\text{Reliability} = f(\text{Reliability Programme costs})$$

Although many of the activities associated with a reliability programme are similar in nature to ensuring a quality product, the reliability should be treated separately[7]. A quality product, in terms of meeting specification may not have the required reliability. For example the construction of a sea wall to combat land erosion for the next ten years may be made to specification in terms of size and location. However, if a reliability programme is not implemented in testing the durability and the required reliability is not attained the wall will require maintenance and/or replacement in the ten year period. Similarly a reliable product may not be a quality product. For example if much time and resources are utilised in ensuring a reliable product the quality of the product may suffer.

•Time

The duration of the project life cycle comprises of the time from the project start to the handing over of the product (time T_1) and the time the product is in operation (time T_2).

$$\text{Duration}_{PLC} = T_1 + T_2$$

Time T_1 depends on the completion of three project phases (Phases 1-3 above). These phases often consist of many activities including the quality and reliability testing which can be represented as activity networks[8].

- **Cost**

The cost is more complex to model compared to time as there are many factors which contribute to the cost at different times in the project life cycle. Prior to the product handover, there are numerous costs including production costs, quality costs and reliability programme costs. After the handover when the product is in use, costs can be incurred in maintenance and possibly replacement. These are the costs of unreliability and denoted as C_2. The cost associated with the project life cycle is given by,

$$Cost_{PLC} = C_1 + C_2$$

The elements of C_1, the costs prior to product handover, together with the costs associated with the risks of not meeting the targets set in the conception and definition phase will be discussed further in the context of a project risk analysis.

Control of duration, cost, quality and reliability

The purpose of splitting a project into phases is to identify the activities that influence the duration, cost, quality and reliability and how they can be controlled. The time and cost before handover, T_1 and C_1 respectively are now shown as,

$$T_1 = WHAT_{time} + HOW_{time} + ACHIEVEMENT_{time}$$

$$C_1 = WHAT_{cost} + HOW_{cost} + ACHIEVEMENT_{cost}$$

The times and costs associated with the WHAT?, HOW? and ACHIEVEMENT phases are determined from the duration and the costs associated with production, reliability and quality prior to product handover. All of these times and costs are assumed to be represented as known probability density functions, as uncertainty in their values exists. Therefore T_1 and C_1 will also be represented as probability density functions. In addition to the above times and costs we need to evaluate the consequences of not meeting the targets defined in phase 1. Hence in our extended definition of a risk analysis we can now evaluate the consequences of not meeting the targets planned for in phase 1.

PERFORMING A PROJECT RISK ANALYSIS

Once targets have been set for duration, quality, reliability and costs together with a plan of how the project is to be conducted a risk analysis can be performed to indicate the feasibility of meeting the targets and a consequence if the target is not met.

We shall use a definition of risk similar to those used by Rowe[9], Lowrance[10] and Rescher[11].

Rowe defines risk as '*the potential for unwanted negative consequences of an event or an activity*'.

Lowrance on the other hand defines risk as '*a measure of the probability and severity of adverse effects*'.

Rescher explains that '*risk is the chancing of a negative outcome*'.

Specific to a risk analysis of a project we shall first define the risk components as comprising of duration, cost, quality and reliability. This is where the risks associated with the unsuccessful completion of the project lie. Using the following formulation, based on Rowes's, Lowrances's and Rescher's definitions a risk level can be subjectively determined for each of the risk components.

Risk Level = f(Probability of **NOT** meeting target, Consequence)

The probability of not meeting the targets for the risk components can be determined from assessing the uncertainty in each component. The uncertainty for each component can be represented as a probability function. From the location of a target an indication of the probability of achieving the target can be established. For example if the uncertainty in the time before handover, T_1, is represented as a cumulative probability function and the target handover time is t_1, then P is the probability of meeting the target and (1-P) is the probability of *not* meeting the target (see Figure 2).

Modelling difficulties

A problem with comparing the risk components is that we need a common measure for the consequences. For example, if there is a possibility of the handover time of the product being late as the level of quality is less than required, how can the situation be evaluated? A common measure is required. Money is that common measure. If the handover of the product is delayed, how much will it cost? If an acceptable level of quality is not obtained how much will it cost? By analysing the risks in this manner cost effective trade-offs can be performed. A compromise may be necessary between the targets which minimise the risk of losing money. A difficulty that may arise is locating where in the project life cycle the consequential costs will be incurred.

For example the costs associated with poor quality can be experienced in rectifying problems in the *acquisition phase* of the project life cycle. Similarly costs can be incurred for taking longer than planned, causing increased production costs and a penalty cost may be enforced for late delivery. However more effort spent in the development phase planning for possible causes will increase the time and costs within the phase, but may reduce the risk associated with the project lifecycle cost in the long run. These consequential costs need to be incorporated into the Project Life Cycle cost. The task of obtaining the required data to model the relationships and possible consequences is also recognised as a modelling difficulty.

O'Conner[12] realised that when a reliability programme was performed the costs incurred from planning and building reliability into a product are often less than the costs of an unreliable product when in operation. We are applying this principle in the context of a project life cycle. If it is possible to represent the uncertainty associated with meeting targets as a pdf and estimate the cost associated with not meeting the target, then risk components can be traded-off with one another, to minimise the risk associated with the total project life cycle costs.

A consideration of the long term affects is therefore necessary when evaluating the consequences. By modelling the consequences using this approach the organisation will have an indication of how the time planned, and the costs expected for production, quality and reliability at the various project phases will affect the overall objective of the project, which is to make money.

It is also important to realise that the risk components are not independent[4]. Altering the uncertainty associated with any of the risk components may result in a change in one or more of the other components. The following diagram, figure 3, illustrates the two way relationship between the risk components. We expect time and cost to determine reliability and quality, or alternatively reliability and quality to determine time and cost. However since both reliability and quality affect time and cost directly we have left out the link between reliability and quality.

The magnitude of the effect between the components is uncertain as exact relationships between the risk components is often not established. However the realisation that a change in one component may affect the other components can provide an understanding of the mechanics of a project and aid in the process of decision making for minimising project risks.

CONCLUDING REMARKS

The purposes of this discussion paper was to identify common properties between most projects and provide a structure/framework for a project using the concept of a project life cycle which enables the control of duration, cost, quality and reliability in the early stages. To evaluate the consequences of not meeting targets for the risk components it is suggested to use money as a common measure. This can enable cost effective trade-offs to be performed between the risk components at an early stage in the project.

Ideas have been presented and discussed which with further work and investigation of the costs experienced will enable a mathematical model to be defined. This could then provide an essential evaluation of decisions associated with the resources required and how they should be allocated in controlling quality, reliability and time to minimise the overall project life cycle cost.

REFERENCES

1 **TURNER, J R** Editorial, *International Journal of Project Manager.* Vol 12 No4 (1994) pp 203 - 204.

2 **MODER J J, PHILLIPS C R and DAVIS E W,** *Project Management with CPM, PERT and Precedence Diagramming.* Van Nostrand Reinhold, London(1983).

3 **NICHOLAS J M,** *Managing Business and Engineering Projects.* Prentice-Hall International Editions (1990).

4 **BS 6046** : Part 1: (1984) Guide to the use of management, planning, review and reporting procedures.

5 **BLANCHARD B S,** *Design and Manage to Life Cycle Cost.* M/A Press, Oregan (1978).

6 BS 6143 : Part 2 : (1990) Prevention, appraisal and failure model, Section 2, Definitions.

7 **LLOYD D K and LIPOW M,** *Reliability: management, methods, and mathematics.* Prentice-Hall, Englewood Cliffs, NJ (1982).

8 **REISS G,** *Project Management Demystified.* E & FN SPON, London (1992).

9 **ROWE W D,** *An Anatomy of Risk.* John Wiley and sons, New York (1977).

10 **LOWRANCE W W,** *Of Acceptable Risk.* William Kaufmann, Los Altos, California (1976).

11 **RESCHER N,** *Risk: A Philosophical Introduction to the Theory of Risk Evaluation and Management.* University Press of America (1983).

12 **O'CONNOR P D T,** *Practical Reliability Engineering.* John Wiley and Sons (1991).

13 **KLEIN J H,** Modelling Risk Trade Off. *J.Opl Res. Soc.,* **44**,(1993) pp 445-460.

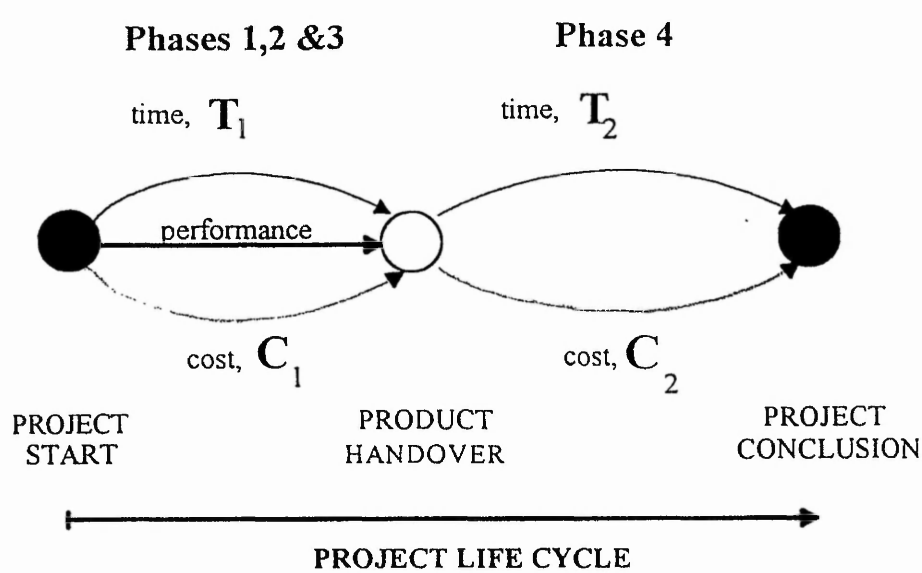


Figure 1. Project Life Cycle

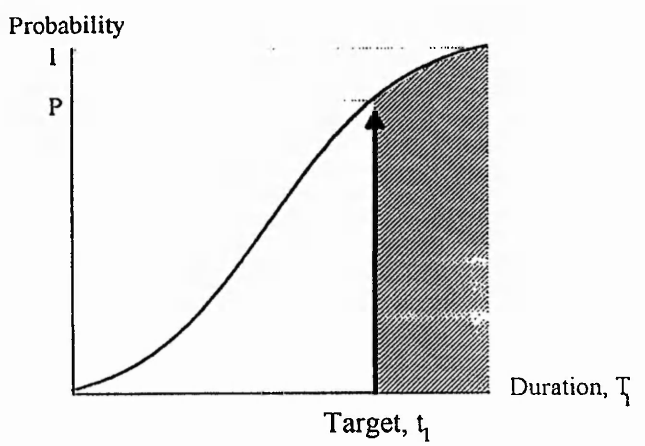


Figure 2. Assessing the chance of not meeting a target

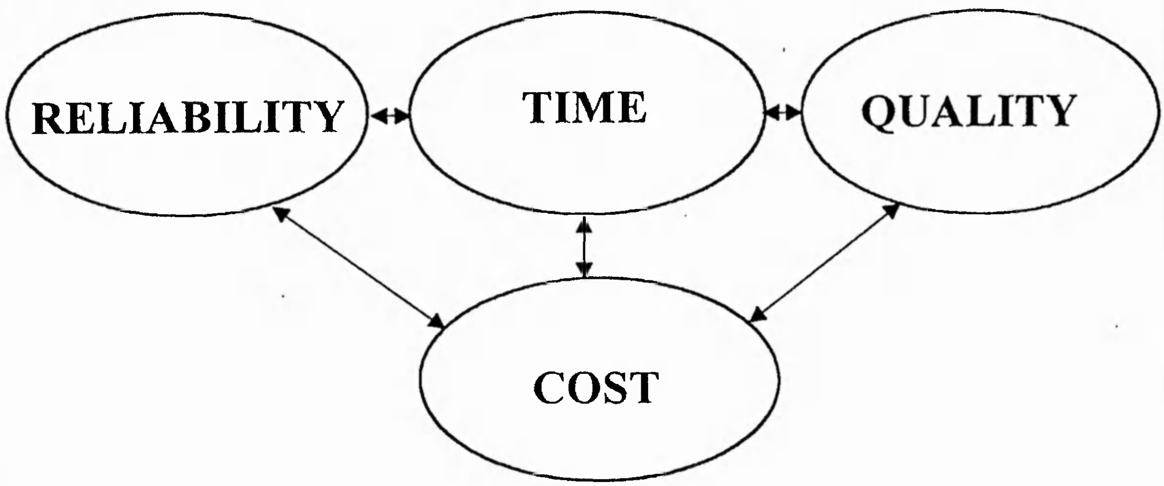


Figure 3. Relationship between time, cost, quality and reliability