A STOCHASTIC MODEL FOR

MANPOWER PLANNING

A thesis submitted by

JAMES I. J. MEDDINGS

to the University of Aston in Birmingham

for the Degree of Doctor of Philosophy

May 1979.

A STOCHASTIC MODEL FOR MANPOWER PLANNING

Submitted for the Degree of Doctor of Philosophy, 1979

by

JAMES IAN JOHN MEDDINGS

SUMMARY

The thesis deals with manpower planning problems within large organisations. The work was accomplished as a temporary employee of Dunlop U.K. Tyre Group under the auspices of the University of Aston I.H.D. scheme. The terms of reference for the project were;

a) To determine a methodology for estimating the requirements relating to the requisition and dispersal of manpower.
b) To produce a manpower plan for the short term (1-2 years), medium term (5 years) and the long term (10 plus years).

After an initial analysis of the problem faced by Dunlop and the state of available manpower records, it was concluded that classical statistical methods would be inappropriate. The aim was, therefore, to construct an estimation procedure which could handle; limited data, time-variant parameters and account for information gained only through noise corrupted observations.

Following a comprehensive and critical review of the current use of statistical techniques in manpower planning, a general stochastic model is formulated. The structure, solution and many applications to manpower planning of this general problem are examined. Consideration of the grade transitions in an organisation leads to a new probability distribution termed the 'Dirichlet-Multinomial', and the derivation of its properties.

On the completion of suitable Supply and Demand models, the question of controlling manpower systems is considered. A general cost function is constructed and algorithms for minimal cost control are given.

Finally, results obtained by the application of the stochastic models to Dunlop data over the period 1972-1977 are presented.

Keywords

Manpower Planning, Stochastic Estimation, Hilbert Space, Dirichlet.

ACKNOWLEDGEMENTS

I am grateful to Dunlop U. K. Tyre Group and the Science Research Council for sponsoring my research. I wish to thank my colleagues at Fort Dunlop and my tutors at Aston University for their advice; in particular I mention Phil Lavender of Dunlop, and Ray Loveridge and Alastair Cochran of the university. I am also indebted to Judith Shutler for her opinion of the work that has led to considerable improvements in the text.

CONTENTS

CHAPTER ONE

MANPOWER PLANNING AT DUNLOP U.K.T.G.

1.1. Introduction to the project

1.2. Historical background of the Company

1.3. Manpower information in Dunlop U.K.T.G.

1.3.1. Adaptation of existing manpower information

1.3.2. Design and implementation of a computerised manpower data system

1.3.3. Immediate uses of the data base

1.4. Review of the project and the way ahead

1.1. INTRODUCTION TO THE PROJECT

The work of the thesis was carried out as a member of the Interdisciplinary Higher Degrees (I.H.D.) Scheme of the University of Aston in Birmingham. Within the scheme a student becomes a temporary employee of an organisation, in order to solve an identified practical problem that exists in industry. Approximately 70% of the student's time is spent within the sponsoring organisation, which enables a thorough understanding of the problem to be obtained. The scheme underlines the difference between problems encountered within industry and those tackled at the undergraduate level. The solutions to 'well behaved' problems met in purely academic work are not necessarily directly applicable in the industrial environment. It is soon learnt that data upon which classical techniques are well proven are simply not available in the required form, may be limited in length, inaccurate or have just not been collected. This in itself is not at all surprising as the need for collecting data is not usually foreseen until the problem has been identified.

Nevertheless, limited data is a facet of industrial life and managers have to make decisions every day based on imperfect information, which itself presents many difficulties. The problem of manpower planning in Dunlop U.K. Tyre Group was of this category and the company was already aware of some of the difficulties associated with their current system. It was in view of this that the project was instigated with the terms of reference being stated initially as;

a) To determine a methodology for estimating the requirements relating to the requisition and dispersal of manpower

b) To produce a manpower plan for the short term (1-2 years), medium term (5 years) and long term (10 years plus).

The first objective was to gain a full understanding of the problem. In order to achieve this it was necessary to undertake a review of Dunlop U.K.T.G. with emphasis upon corporate strategy, employee structure, production processes, present personnel and employment practices and current management attitudes towards manpower planning. It was discovered that senior management held a definite commitment to the project but having limited knowledge of manpower planning techniques and their implications to the organisation they reserved judgement on the adoption on any system currently available.

In tandem with this induction to the company a thorough study of the academic work within this subject area was performed, the results of which are presented in the literature survey of Chapter 2.

1.2. HISTORICAL BACKGROUND OF THE COMPANY

In 1888 the first practicable pneumatic tyre was invented by J.B.Dunlop, a Scottish veterinary surgeon then practising in Ireland. The company was formed the following year in Dublin, assembling tyres from components supplied by other rubber manufacturers. However, Dunlop soon acquired its own rubber component manufacturer based in Birmingham. Demand for the product expanded necessitating the establishment of manufacturing and selling companies in Australia, South Africa, Canada, France and Germany. A further tyre factory was also needed in Britain and the natural choice for location was Birmingham where they already possessed the component works and in 1900 the tyre operation was moved to Aston. Dunlop then began to expand both its production and the extent of operations to adjacent industries, i.e.; the supply of raw materials and wheel production. In 1906 it took over a wheel manufacturing company in Coventry and 1909 bought some Malayan rubber plantations. By 1920 it was the largest single owner of plantations in Malaya and had established cotton mills in Rochdale for the manufacture of tyre cord and other fabric. In 1916 the Fort Dunlop site of nearly 400 acres was purchased and this was to become one of the largest tyre factories in Europe.

Dunlop suffered badly in the collapse of world trade in 1921, the drop in its raw material prices being catastrophic. This led to the decision of a new board to diversify the company's activities; acquisitions included C.Macintosh of Manchester (hose, belting and general rubber products), W.Bates of Leicester (cycle tyres and rubber thread) and a sports racket manufacturer at Waltham Abbey which became Dunlop International Sports. In 1929 Dunlop's laboratories invented 'Dunlopillo' latex foam, this formed a second important development in the operation. Further expansion took place overseas and Dunlop's tyre production was extended to Scotland with the purchase of the India Tyre and Rubber Company at Inchinnan near Glasgow.

In World War II Dunlop's research and manufacturing resources were devoted to the war effort. During this period the plantations fell under the control of the Japanese and its French and German factories were destroyed. It was at this time that American competition, which escaped the effects of the war, was able to attain a very strong commercial position. However, after the war Dunlop was soon able to react and resume its position as a world force with the commissioning of many new factories, including a tyre factory at Speke, Liverpool. Dunlop continued with its diversification policy, acquiring major interests in sports, hose, belting and fire protection equipment. In 1970 the company was the 39th largest outside the U.S.A., having 130 factories in 22 countries, providing employment for 108,000 over half of which were in the United Kingdom.

One of the most important advances in tyre technology was the development of the radial tyre but it could be said to be the instigator of the decline of the tyre industry. The main feature of the radial is its ability to last twice as long as the cross-ply tyre, resulting in over capacity within the world tyre market. In addition the policy of leasing factories and setting up tyre operations in the lower paid areas of Eastern Europe brought a flood of cheap tyres onto the market, which aggravated the situation.

As Dunlop had been one of the earliest manufacturers of tyres the plant and machinery were becoming old and outdated, thus productivity compared with competition from new sectors of the industry was low and over manning high. Dunlop decided to set up a small new tyre factory in Washington, County Durham to try to counteract this situation.

The labour force has subsequently decreased throughout the other U.K. tyre factories; Speke, Fort Dunlop and Inchinnan.

It is thought that two more factors will influence the decline of the tyre market;

i) The market price of the new steel radial

ii) The fail safe (Denovo) type tyre

The price of steel radials when compared with cross-ply does not reflect the added benefit to the customer, it is effectively under priced to the point of being a commercial disaster in an inefficient production unit. The over capacity in the tyre market creates a ruthless competitive situation and many older tyre factories must be doomed, unless they can be re-equipped and the majority of the workforce quickly trained in work other than tyre production. The fail safe tyre and wheel unit, 'Denovo', which enables control with a 'blow out' and continued travel for up to 100 miles at normal speed is perhaps Dunlop's main achievement of the seventies and as yet there is little competition in this area. Unfortunately it has been found that the public are reluctant to pay extra for this added safety feature and when competitors are able to react the market price might be affected in the same way as that of the steel radial in the past. The other important feature of this tyre is that it makes a spare tyre superfluous and thus reduces the world's car tyre market by a further 20% at once.

In 1971 Dunlop formed a business union with Pirelli. The main advantage of this was the complementary marketing areas of the two companies. Dunlop acquired a greater part of the markets of Italy, Southern Europe and Latin America and Pirelli; North America, Africa and Asia. The wider product range of the combined group was seen as an insurance against the effects of fluctuations in the fortunes of the existing products of each partner.

Nevertheless, the tyre industry is not in a state of expansion and the rate of dispersal of excess labour, whether through retraining, redundancy or natural wastage, will be a necessary consideration of any manpower system. The problem was involved with the maintenance of a viable workforce at minimum cost, given the present state of the industry. Thus, the desired manpower structure was not constant over time but was dynamic and subject to stochastic variation.

1.3. MANPOWER INFORMATION IN DUNLOP U.K.T.G.

Having gained an appreciation of the structure of U.K.T.G. it was considered sensible to ask what type of information was available at the employee level, as any manpower system could only be as good as the information upon which it was based. It was found that the main source of information for the operative level were manual records held in the Personnel Department at each of the four main factory units and that staff up to a certain salary level were again on manual records held at group and departmental level. Information on senior managers and employees with potential to reach this status was held at Dunlop H.Q. in London. In addition some valuable information for manpower planning purposes had been processed by the computerised salary suite in Coventry.

It was discovered that a Dunlop Computerised Personnel Information System (C.P.I.S.) had also recently been scrapped. The main reasons for its discontinuation was its inability to be kept up to date coupled with experiences of large time delay in information retrieval. This had meant that few requests were processed and hence the system proved to be cost ineffective. Various discussions took place when this system was to be terminated and Fort Dunlop thought that they could operate a C.P.I.S. more efficiently on their own. Unfortunately, due to plans to change both the computer payroll system and the present computer this scheme was never realised.

1.3.1. ADAPTATION OF EXISTING MANPOWER INFORMATION

My two immediate priorities were to decide whether useful information could be salvaged from the old system and to ensure that no further information was lost. A search through the 'dead' computer tapes, provided no useful data but 'year-end' tapes from the old salary suite had been kept for 'tax' purposes and these it was thought might contain useful information. These I.C.L. tapes for the years 1970-76 were then examined and 'cleaned' on the Coventry computer. Relevant data included;

Insurance number

Works number

Grade/Salary

Date of Birth

Date of Engagement

Date of Leaving

Sex

which were extracted for each year and merged into a new tape. This tape was converted to I.B.M. format for use on the Fort Dunlop computer.

Since Dunlop had a semi-structured salary payment scheme for its lower and medium salary range and with reference to location/year/ salary tables it was possible to achieve a graded data bank for the years 1970-76 on the computer. Any exceptions were traced through the manual records and graded taking account of standard yearly increases. The higher salaried managers up to board level were graded manually.

The whole package was then set up onto a 'QUEST' file. QUEST is the acronym for Query Extract Sort and Tabulate and is a programming language designed for this purpose. Its main attribute is the ability to handle large amounts of data rapidly and efficiently, although its mathematical ability is limited. However, it provided an excellent means of sorting and checking the salvaged data.

1.3.2. DESIGN AND IMPLEMENTATION OF A COMPUTERISED MANPOWER DATA SYSTEM

The recent failure of a Dunlop C.P.I.S. stimulated a thorough examination of the holding of manpower records. Manual personnel records were in existence for the 15,000 employees of Dunlop U.K.T.G.. Accessibility to data on an individual employee was fast and efficient if the name or department was known; but if information was required on points such as;

i) Number of engineers in the organisation aged over 30 years

ii) Number of retirements in 1980

iii) Number of employees working as accountants outside the Finance Division

the task was almost impossible as the design of current records only accommodated data on known individuals. Any additional demands, such as the profile on a group of employees led to unacceptable work loads, the cost of such a manual search operation being excessive due to the man-hours required.

Fortunately, Dunlop had recently purchased a new salary suite developed by Peterborough Data Processing and this had space for the inclusion of personnel information at any time on the Employee Master File. In addition to the usual data common to both salary and manpower requirements the suite had the following capabilities; a 200 character memoranda file and a 10 x 10 event-date automatic updating file which were available on the standard suite and could be increased if necessary. The advantage of having this type of arrangement of a C.P.I.S. intergrated with the payroll are well documented. It was also intended that the salary suite be extended to include the operative level sometime in the near future, thus making a unified career record available for the first time.

With the cooperation of U.K.T.G. management committee, Dunlop Central Personnel who were the custodians of the suite were approached in order that the possibilities of using the file for manpower planning could be examined. It was decided that the additional information that U.K.T.G. desired on the salary file for each employee would be initially limited to;

i) Job Title (abbreviated)

ii) Job Activity (coded)

iii) Job Evaluation Scheme (coded) and Points on the Scheme

iv) Date of Commencement of New Appointment

The job title was to be abbreviated and stored in the personnel memoranda file and could be overwritten at each job change. A full list of standard abbreviations were issued to the factory personnel departments so that the abbreviated job title could be consistant and easily understood. The job activity to be used as a basic unit for analysis was a three digit numerical code. This was very similar to the coding adopted on the Institute of Manpower Studies survey but was somewhat extended and tailored to meet Dunlop U.K.T.G. requirements. It was as a result of consultation with K.Smith of I.M.S. that a transformation between the codes was established.

Both the information on the job evaluation scheme/points and the job activity together with the salary were stored against the date of change in the personnel history file. In this way, the historical events covering up to nine changes could be stored. The file is constructed so that the new data automatically re-orders the file on a 'roll-over' basis in which the new entry goes to the top of the stack and the other entries move down one place. Updating was possible by a small change in an existing salary change form, used as an input document to the computer salary file, as it already contained all but the new job activity code. This form was redesigned (Appendix D) and unnecessary data removed.

The discussion of U.K.T.G. proposals with other Groups in Dunlop led to a committee being formed to examine the possibility of extending the system to all Groups. General agreement was reached with all participating Groups and permission to proceed was requested from the director responsible for both Group Management Services and Central Personnel. Permission was denied and concern for the security of information was mentioned as the reason. The committee reviewing manpower data was disbanded as a result of this decision.

U.K.T.G. were not satisfied with this decision and made several efforts to reverse it. It was explained that the security of information would be at no more risk with the new system than it was on the present salary suite and the additional benefits of the system were expounded; such as the assistance in the determination and formulation of procedures in the following areas;

a) Remuneration planning and control
b) Statutory returns
c) Staff development
d) Manpower planning

at little additional cost. The decision to proceed was eventually given for U.K.T.G. only. The Tyre Technical Group, also based on the Fort Dunlop site, were later allowed to join the Tyre Group scheme, as it was argued that much interchange existed with that Group and the Factory Technical Division of U.K.T.G..

Six months after the commencement of discussions the four Tyre Group factories were asked to provide information on a suitably designed input document for all current staff. This was then input to the computer and the information was made available through a monthly extracted tape covering all manpower information on the salary suite. Updating now, intimately connected with the salary data insured that the manpower data was the latest available. A simple merging operation with the salvaged data, discussed in the previous section, enabled a continuous but limited manpower data to be accessed for staff over the period 1971-77. As the manpower data base is updated this system will provide an increasingly more comprehensive source of information.

1.3.3. IMMEDIATE USES OF THE DATA BASE

At once vital information was available from this data bank, enabling staff previously occupied on the manual compilation of personnel statistics to be released from this work, so that they could concentrate on the analysis of the data. Basic programmes were quickly written providing useful information in the following areas;

i) Verification of job grade to salary levels

ii) Discrepencies in payments for similar jobs in different areas

iii) A tabulation of the grade/length of service/age distribution by department/job activity

iv) Average salary by department/job activity

A preliminary examination of 2880 staff based in England was undertaken, the results of this are presented in Appendix D. The data from programme iii) above provided the input to a 'camera' type analysis which will be discussed fully in Chapter 2. This was useful for obtaining a 'feel' for the manpower system, but this was only one observation of the employee structure at a single point in time and a deeper time dependent analysis was necessary.

1.4. REVIEW OF THE PROJECT AND THE WAY AHEAD

It was clear at this point that only limited data was going to be available and so a classical statistical analysis based on long time series would be impossible. An analysis of the yearly transitions and a monthly analysis of leavers with a low length of service revealed that not only was the employee structure changing but promotion rates and wastage rates were time variable. Also an observation of someone leaving is the only firm information that is available to demonstrate an employee's propensity to leave which means some observation error would occur in the manpower system. The objectives of the project were now more clearly seen to be;

a) To develop an estimation methodology that would be optimal given the limited time variable data available, which in turn is subjected to observation error.

b) Given these estimates of the future supply and demand structure of Dunlop U.K.T.G., determine an optimal control policy under a realistic cost function.

The general methodology which was developed to give optimal stochastic estimation is discussed at length in Chapters 3 and 4. The application of this method to the General Problem, along with a methodology based on a new 'Dirichlet-Multinomial' distribution is expounded for supply forecasting. The problems of demand forecasting are dealt with in Chapter 6 and in Chapter 7 the difficulties of controlling a system subject to stochastic variation are considered.

However before the development of a methodology can be attempted it is wise to make a critical review of the relevant literature in order to set the new work in the context of current methods.

CHAPTER TWO

A CRITICAL REVIEW OF MANPOWER PLANNING MODELS

2.1. General introduction

2.2. Labour wastage

2.2.1. Monitoring the labour process

2.2.2. Completed length of service distribution

2.2.3. Parametric estimation of the wastage curves

2.2.4. Non-parametric methods

2.3. Manpower supply forecasting

2.3.1. The Stationary model

2.3.2. The Cambridge model

2.3.3. Simulation models

2.3.4. Renewal models

2.3.5. Markov models

2.4. Manpower demand forecasting

2.5. Control of manpower systems

2.1. GENERAL INTRODUCTION

Manpower planning may be considered at four levels; national, industrial, organizational and individual. At the national level the main determinants of future manpower requirements are the levels of economic activity and the future state of technology (Leicester (1969,1971)). In fact, national manpower forecasting becomes an integral part of the total economic forecasting of the country, enabling national policies to be determined. It should be noted that planning for large regions (i.e. a state in the U.S.A.) may be considered in an analogous way to national planning.

At the second level, that is the industrial and occupational level, the main forecasting tools are the econometric models and again changes in technology are considered important. Much of the work carried out at this level has been documented by the O.E.C.D., particularly in the area of labour mobility. Recently, the Industrial Training Boards have become more aware of the benefits that come from monitoring and planning the human resource, and have encouraged the gathering of basic personnel information from the industries that fall under their jurisdiction. The state of the art at this level was first reveiwed by Yewdall (1969) and his work is an adequate introduction to studies concerning industries and occupations.

The major part of this thesis is aimed primarily at the third level- that of the organization. It could be said that the foundations of manpower planning were laid more than 300 years ago. It was then that the Admiralty started to hold records on the throughput of naval officers, and indeed the Navy has one of the longest documented data banks of any hierarchical organization. The scope of the monitoring was extended in the reign of Queen Victoria to include a 'central personnel information system' to help with career development. About this time the first 'actuarial' model for career planning was developed.

The second world war saw the rebirth of manpower planning (Seal (1945), Jones (1946,1948) and Vajda(1947,1948)), sponsored to a large extent by the Royal Navy and again actuarial models were prominent. Following on from this the N.A.T.O. conferences of the late sixties (Jessop (1966), Wilson (1969), Smith (1971) and Bartholomew and Smith (1971)) provide a useful insight into the systems that are in operation within a strict hierarchical organization. It is within these papers that statistics affirm their position as the core discipline of manpower planning. Authors, in these collections of papers, used the concepts of demographic and actuarial statistics, deterministic and stochastic models and because of the form of military organizations, linear programming models (Morgan (1971a,1971b) and Charnes and Cooper (1971)) were also proposed.

On the civilian side, the initial catalyst of manpower planning, as it was for much of management science, was the work of the Tavistock Institute on the Glacier Project (Rice, Hill and Trist (1950,1951,1952), Brown and Jaques (1965,1971) and Jaques (1958)). These papers generated a lot of interest into the phenomenon of labour turnover (Silcock (1954, 1955), Lane and Andrew (1955) and Bartholomew (1959)).

In the early sixties work began on the 'university manpower system' and it was then that two very important papers were published, (Young and Almond (1961) and Gani (1963)). The former covered the topic of 'a' labour turnover index very well. Papers continued to appear in learned journals at a steady rate throughout the remainder of the sixties, then dramatically increased in the early seventies (Hyman (1970), Bartholomew (1969,...), Forbes (1971,...),...). Since there is now a vast amount of literature written on manpower planning at the organizational level, only references to works that are particularly important to the development of the thesis will be made. These will be found in the following four sections. It is noted that Huber (1976) gives extensive references concerning both the quantitative and qualitative approaches to the measurement of labour turnover.

The final level is that of the individual. Here is entered the disciplines of Psychology and the Behavioural Sciences, but statistics is still a basic tool of analysis. The literature in this area developed much later than that of the other levels. The works of Herbst(1963) and Clowes (1972) should be mentioned, both of which are based on Markov transition matrices. These papers are mainly concerned with an individual's perception of the firm, the elements of the matrices being the probability that an individual moves from one state to another. States such as 'temporarily committed', 'permanently committed' and 'decided to leave' are introduced. The identification of attributes that are correlated with promotability, and propensity to leave are important (Timperley (1971)). These attributes have been included in computer simulation models, (Walmsley(1971), Wishart and Ko (1973) and Weber (1971)) where the approach is to mirror the manpower system by probable individual movements. Most of these tackle the concept of promotability, by age, service and some measure similar to Jaques' (1958) salary progression index.

Returning to the organizational level, the models that have been proposed for analysing the manpower system at this level are now critically reviewed. It is convenient to discuss this topic in four stages; the measurement and prediction of labour wastage, supply forecasting, demand forecasting and control of the manpower system. It is remarked upon that this partition is for ease of exposition only, as each part of the manpower system interacts with the others.

2.2 LABOUR WASTAGE

Labour turnover refers to the flow of employees throughout organisations. It is readily seen then that labour wastage is only a part of labour turnover. It is not unusual, however, to see the terms 'wastage' and 'turnover' used synonymously. The problem of wastage has been examined and various indices have been proposed to measure this leaving process. A selection of these are considered below:-

2.2.1. MONITORING OF THE LABOUR PROCESS

The British Institute of Management (1949) proposed an index of labour turnover, commonly known as the 'crude' turnover rate;

$$I = \frac{\text{Number who leave in a given interval} \times 100}{\text{Average number employed during the same interval}} \%$$

Usually the time interval is that of the calendar year, but any time period may be used. The main characteristic of this index is its ease of calculation - perhaps its only recommendation. It is believed that the B.I.M. index is the most widely misused index in Manpower Studies and its shortcomings are well documented, (Silcock(1954, 1955), Bowey(1969) and Bartholomew(1959)). The main criticism of the index stems from its inability to take account of the length of service of an employee. This is important as it has been observed by many researchers that length of service correlates very highly with propensity to leave, (Rice,Hill and Trist (1950), Young and Almond(1961), and Lane and Andrew (1955)). There are other objections, though to be fair to the B.I.M. these were not apparent at the time;

(i) It has been impossible to find a single parameter distribution that adequately follows the form of the 'completed length of service' distribution This is further explained in the sequel. Considering this, it is clear that a single index that would summarise the wastage process is an unattainable goal.

(ii) Less numerate managers have corrupted the index taking 'average' to mean the half sum of the number of employees at the beginning and the end of the interval. In a relatively unstable work force this is an unrealistic approximation and can lead to misleading results.

Criticisms such as these led Silcock(1954) to reason that:

'the difficulties of interpreting the crude rate are such that it would probably be better to abandon it altogether and to seek an alternative measure of the employer's power to retain labour'

This led to several proposals of different types of stability indices:

Duncan's Index (1955) was;

$$I = \frac{\text{Number of employees with more than 1 year service now}}{\text{Total employed 1 year ago}} \times 100\%$$

The only way in which this index takes account of the length of service of employees is by its avoidance of the most volatile group - the new recruits.

Bowey's Index (1969) was;

$$I = \frac{L_n}{\sum_i n_i N_i} \times 100\%$$

where L_n is the sum of the length of service (in months) of all employees with less than two years service plus 24 times the number of employees with two years or more service. The denominator is the sum of the sum of the steps in the expansion of the labour force, a step being measured by the total size of the labour force at time i,(N_i) multiplied by the number of months it remained at that size (n_i).

This index does not discriminate adequately between a labour force with constant propensity to leave across all length of service categories and one with an unchanging core of workers but also having quick 'turn round' employees.

It is worthwhile to consider one further single index. It was conceived by Rice et al (1950) and later Bartholomew (1973 b.) that the leaving rates of employees from organisations could be considered in an analogous way to mortality rates. Demographers and actuaries both know and understand the limitations of a crude death rate index and therefore proposed the use of standardised mortality rates. Here a standard age (length of service) distribution is assumed. The standard index is simply the value that the crude turnover rate would have been if the length of service distribution were standard. In this way the dependency of severance on service is discounted. Other correlates of severance can be standardised in a similar way. Bartholomew (1973 b.) shows that if the standard length of service distribution were chosen to be the steady state distribution (see 2.3.1.), this index would be equivalent to the mean of the completed length of service distribution suggested by Lane and Andrew (1955). Most organisations do not, however, operate in the steady state and hence this distribution would change over time, altering the base of the index.

At present only single indices have been noted. A set of indices, known as the wastage rates (Forbes (1971) and Bell (1974)) will now be considered. The following information is required for each manpower group:

(i) The length of service distribution of the group.

(ii) The number of leavers from each length of service class of the year.

The series of wastage rates are then calculated loosely as:

$$w_i = \frac{l_{i,i+1}}{l_i} \qquad i \in \mathbb{N}$$

where w_i represents the wastage rate of the manpower group l_i which had length of service i at the start of the period, given that $l_{i,i+1}$ were the number of leavers from this group in the next time period. The wastage rates will be more rigorously defined and developed in Chapter 3 where they will be allowed their true dynamic (time-dependent) structure.

2.2.2. COMPLETED LENGTH OF SERVICE DISTRIBUTION

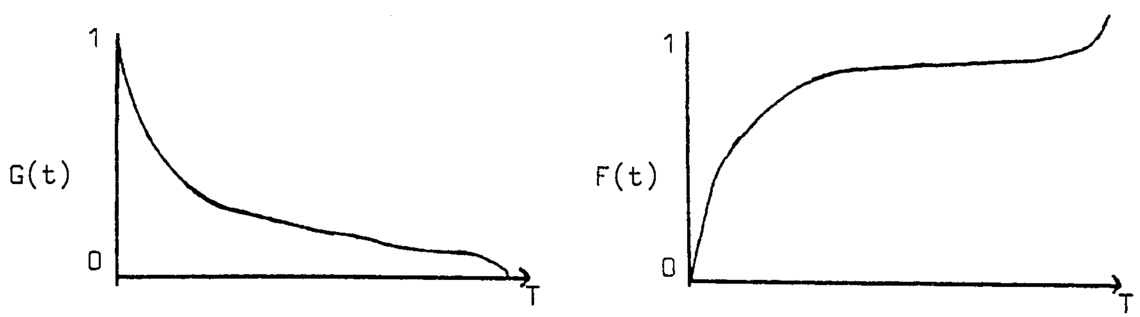
It is convenient here to introduce the notion of the completed length of service (C.L.S.) distribution as it forms the basis for the development of much of the recent work. The terminology and notation stems from Bartholomew (1973 b.) but its origins were much earlier, (Rice et al (1950) and Lane and Andrew (1955)).
The probability that an individual survives in the organisation for a length of time t is denoted by G(t) and known as the survivor function. Its compliment F(t) is called the completed length of service distribution function. The density of the C.L.S. distribution function is represented in the usual way by f(t).

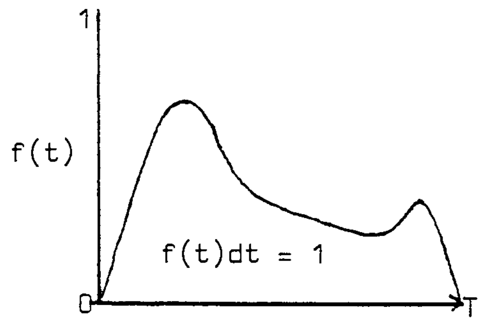
The propensity to leave function $\lambda(t)$ can now be defined as;

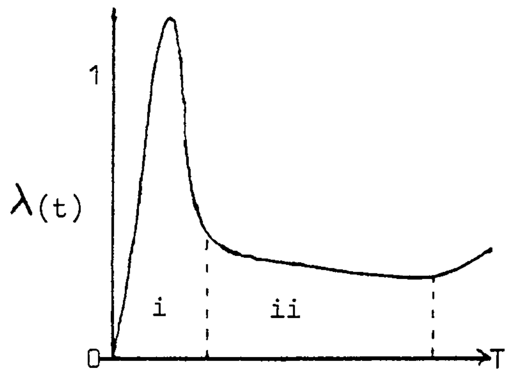
$$\lambda(t)dt = \frac{f(t)dt}{G(t)}$$

Typical graphs of these four functions are shown below;

<u>Figure 2.2.1.</u>







Clearly, all these functions are equivalent under transformation by the following equations;

$$F(t) = 1 - G(t)$$

$$f(t) = - dG(t) / dt$$

$$\lambda(t) = - d \log G(t) / dt$$

Both parametric and non-parametric methods have been proposed for estimating these functions, developments in both occuring in tandem. The parametric method will be considered first.

2.2.3. PARAMETRIC ESTIMATION OF THE WASTAGE CURVES

There were two important contributions from the Glacier Project (1965);

i) The finding that:

'wastage was a function primarily of length of service'

ii) A remark by one of the authors that:

'there seemed to be an underlying pattern in human behaviour with respect to labour wastage, independent of social and economic factors outside the organisation'

Silcock (1954) understood this to mean a constant propensity to leave function that is; $\lambda(t) = \lambda$. As can be seen from Figure 2.3.1., this is a reasonable approximation over the settled period (ii) but not over the induction period (i). This led to a simple survivor function;

$$G(t) = e^{-\lambda t} \qquad \lambda > 0,\ t \geq 0$$

and to an exponential C.L.S. density;

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

The exponential distribution was, however, not a satisfactory fit to the observed data.

Silcocks second approach was to assume that an individual's propensity to leave was constant but admitted the possibility that it might vary between individuals according to a distribution function $H(\lambda)$. Bartholomew (op. cit.) shows that any non-degenerate distribution function $H(\lambda)$ will introduce more skewness into f(t), leading to a better fit to the data. He uses a discrete density function concentrated on two points whilst Silcock uses the usual gamma distribution. When fitted to data from the Glacier Metal Company the fit is equally as good for both models. Even though the 2-term mixed exponential model fits the data well it is not, however, correct to conclude that the assumptions of the model are true, as authors Herbst(1963) and Clowes (1972) starting from different hypotheses have obtained the same functional form for f(t). The main criticism of this distribution is that the density function decreases monotonically with increasing time, which typically is not the case.

Lane and Andrew (1955) proposed an alternative form for f(t) which was the lognormal distribution. This distribution has the desired property of increasing to a mode then slowly dying away. Considerable success has been achieved using the form of the curve (Young(1971)). It has two distinct advantages over the 'exponential' models;

(i) It is closely related to the normal distribution and therefore all inference about the data can be carried out by 'Normal' Theory.

(ii) When plotted on readily available log-cycle paper the resultant is a straight line.

The second point is very important practically as straight lines are easily extended (for prediction) and the process is easy to communicate to non-numerate personnel managers. There are many explanations as to why the completed length of service distribution should be lognormal,(Young(1971) , Aitchison(1955) and Marshall (1974)), these include the assumption that 'an individual's propensity to leave varies with the number of jobs he has held and his length of service in each'. This can be seen to be lognormal in the limit by taking the logarithmic transform of the related distributions and applying the Central Limit Theorem. Unfortunately this does not explain why the lognormal distribution should be applicable to first job employees.

In conclusion, a rigorous argument to validate the use of a particular 'parametrised' distribution is still required. Only then can the power of classical statistical theory be brought to bear on labour wastage analyses. It should be noted, however, that because of the number of cases in which the lognormal curve has been found to fit the data, an approach along these lines would be more than satisfactory in an initial examination of labour wastage.

2.2.4. NON-PARAMETRIC METHODS

Non-parametric methods; that is, methods estimating the C.L.S. distribution without parametric assumptions are now considered. A general reference is Forbes (1971a).

There are two main approaches; those of cohort and census analysis. A cohort is defined as a group of people that enter an organisation in a given period, each manpower group being as homogeneous as possible with respect to age, sex, marital status, et cetera.. The fundamental principle of this method is to eliminate all attributes correlated highly with wastage other than length of service, thus leaving length of service as a good predictor of wastage.

In practice, five to ten years data would have to be collected before meaningful results were obtained. This is far too long, as changes in employee attitude would take place in this period, so invalidating the severance propensity of the original cohort as a useful predictor of the behaviour of new recruits. This problem always occurs when 'static' models are used to estimate truly 'dynamic' systems.

The second approach is that of census analysis. This is a 'snap shot' of the wastage profile at a point in time. The ratio of the number of leavers with a certain length of service over the number at risk in the group is chosen as an estimator of $\lambda(t)$, the propensity to leave function. Forbes (1971a) gives an excellant summary of current non-parametric methods for estimating the survivor function and also shows that the previous estimator is a minimum variance unbiased estimator of $\lambda(t)$. He does, however, persist in calling $\lambda(t)$, his $\phi(x)$, a probability which it is not. This point is discussed further in Section 2.3.1. .

As stated above, the wastage is dependent on length of service but this dependence also varies with real time. A model using this as a basic assumption is proposed in Chapter 3 Section 1 and then further developed in Chapter 3 Section 2 and 3 and also in Chapter 4.

2.3. MANPOWER SUPPLY FORECASTING

Supply forecasting deals with the prediction of the future states of the labour force under assumptions about the flows that occur in the manpower system. Such assumptions might be; constant wastage across all groups, promotion by seniority, recruitment allowed only at the lowest level or compulsory retirement at a specified age.

The simplest of these developments is the stationary model, (Forbes, Morgan and Rowntree (1975)) which can be useful in getting a feel for the problem. A more visual interpretation is provided by the Camel/Cambridge models, (Morgan, Keenay and Ray (1974)), where the promotion prospectus is the main feature of the analysis. Warmsley (1971) uses a form of the promotion prospectus in his paper on simulation models.

When considering the flows that occur in a manpower system, it is worthwhile distinguishing between push flows and pull flows. A push flow is appropriate when there is a constant probability of promotion between grades, that is the flow is in essence from the source. A pull flow describes a process such as if a vacancy occurs, a replacement is selected from levels below, the flow can therefore be considered as initiating at the sink. It is noted that both push and pull flows may be present in the same manpower system. The pull flows lend themselves to analysis by Renewal theory and the push flows by Markov theory.

Most of the work using the Markov approach has considered prediction, assuming the promotion flows follow their expected stochastic path. In Chapter 5 of this thesis the probability transition matrix is assumed to be an observation from a probability distribution. Its distribution and corresponding moments are then derived.

In summary, there has been five main approaches to supply forecasting:

(1) Stationary models

(2) The Cambridge models

(3) Simulation models

(4) Renewal models

(5) Markov models

Each of the above models are now critically examined.

2.3.1. <u>THE STATIONARY MODEL</u>

Under the assumptions of constant recruitment, retirement and expansion policies and given that wastage rates are constant, there is a unique age (length of service) distribution that is time invariant (Forbes, Morgan and Rowntree (1975)). This distribution has been termed the stationary (age) distribution.

A simple stationary model with constant wastage rates is initially considered.

Let the propensity to leave be constant and independent of time;

$$\lambda(T) = \lambda$$

this yields the familiar exponential decay curve for the survivor function;

$$G(T) = e^{-\lambda T}$$

Over the unit time period (T, T + 1) the proportion who leave the system is;

$$\frac{G(T) - G(T+1)}{G(T)} = \frac{e^{-\lambda T} - e^{-\lambda(T+1)}}{e^{-\lambda T}} = 1 - e^{-\lambda}$$

This is clearly independent of T and is the unit (annual) wastage rate w. Therefore;

(1) $w = 1 - e^{-\lambda}$ $\quad w \in (0,1)$, $\lambda \in (0,\infty($ and so $w \neq \lambda$ if $w \neq 0$

If the system is such that the maximum length of service is K years and the structure is maintained by a recruitment R, then N the number in the system is given by;

$$N = R\int_0^k e^{-\lambda T}dT = R\lambda^{-1}(1 - e^{-\lambda K}) \qquad \lambda > 0 \tag{2}$$

It is noted that λ may take all values along the positive axis, so it is not a probability as some authors have called it.

In the stationary model authors have not stressed the difference between λ and w strongly, and some have misunderstood them as being identical. As can be seen from (1) this is not the case.

The first four rows of the Table 2.3.1. compare values of N, obtained under a constant recruitment of 100 for w = 0.1 (0.1)1.0, for the correct values N_c and the false values N_f . N_c and N_f being defined as;

$$N_c = \frac{R(1 - (1 - w)^K)}{\log_e(1 - w)^{-1}} \quad \text{and} \quad N_f = \frac{R(1 - e^{-wK})}{w}$$

The final row gives the appropriate value of λ to be used in (2) where the wastage rates are as stated.

TABLE 2.3.1.

K	W	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	N_c	94.9	89.6	84.1	78.3	72.1	65.5	58.1	49.7	39.1	0
1	N_f	95.1	90.6	86.4	82.4	78.7	75.2	72.0	68.8	65.9	63.3
40	N_c	935	448	280	196	144	109	83	62	43	0
40	N_f	981	500	333	250	200	166	143	125	111	100
R=100		.105	.223	.357	.511	.693	.916	1.20	1.61	2.30	∞

From the table it is readily seen that a constant propensity to leave of 1.2 is approximately equivalent to a wastage rate of .7. It is, perhaps, fortunate that the error in the sector of the table most widely used, i.e.; $K = 1 \mapsto 40$, $W = 0 \mapsto 20\%$, is usually less than 10%.

Earlier in this chapter it was shown that to consider wastage independent from length of service was inappropriate. Even if this assumption were valid much care needs to be taken in choosing the correct formula. This method, used solely to gain an insight into the problem has a limited value; to make decisions based on the results of this model is considered unreasonable.

The next stage in the development of the model was to introduce grades. Dealing with the simpler case of grade invariant wastage rates first, a simple formula relating the fixed proportion of those promoted with a length of service A to the proportion of the total work force is easily derived;

$$p = \frac{q\int_A^K G(u)du}{\int_0^K G(u)du} = \frac{q(1 - e^{-\lambda(A-K)})}{1 - e^{\lambda K}} = \frac{q(1 - (1 - w)^{A-K})}{1 - (1 - w)^{-K}}$$

It is seen then that there are four degrees of freedom in the equation. Fixing a value to any four degrees of the parameters automatically fixes the fifth. Again this method might be useful in obtaining an insight into the structure of the system.

The model can then be extended to admit differing wastage rates between grades, but as before the same criticisms about the assumption of wastage being independent of tenure hold. The Camel/Cambridge models of Morgan and Keenay allow for the changing wastage rate over the age structure. This is more by considering promotion rates of the survivors only than by actually including the wastage rate in the model. These models are considered in the following sub-section.

2.3.2. THE CAMBRIDGE MODEL

The Camel/Cambridge approach requires only stock data to produce results that are useful to management. The necessary input being the age - grade matrix of the organisation. The model calculates the age structure and outputs the proportion of the survivors that are in each age and grade.

FIGURE 2.3.1.

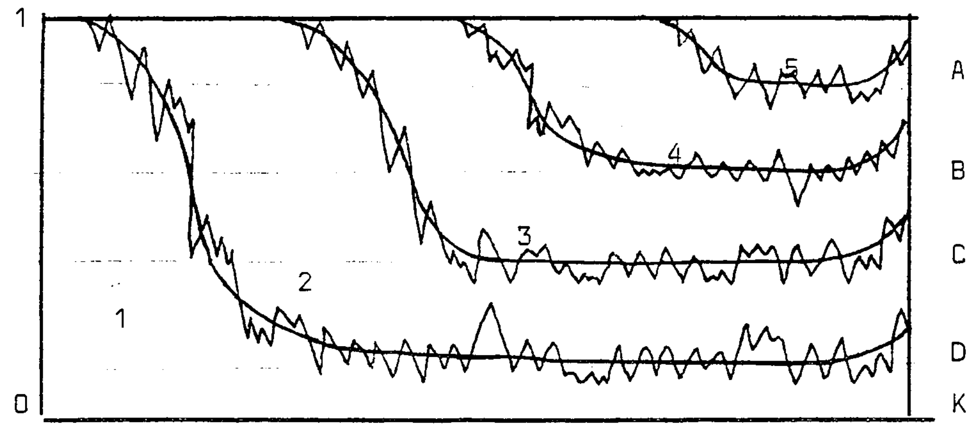
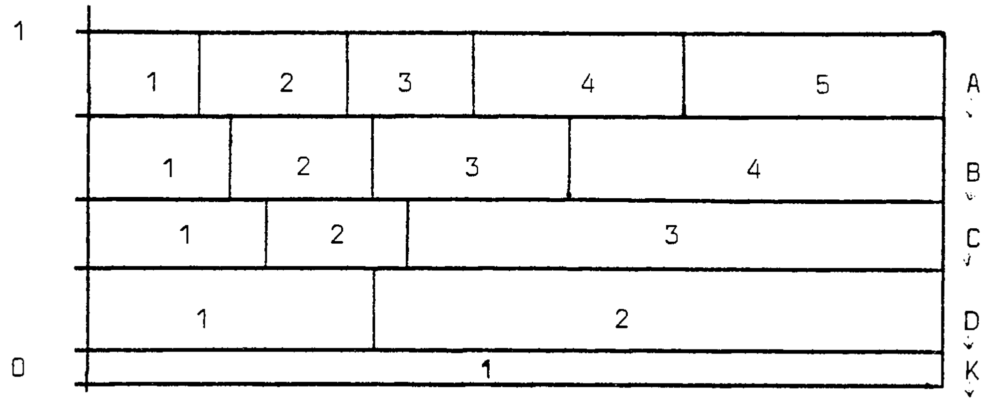


Figure 2.3.1. shows a typical unsmoothed and smoothed output. The output reflects the results of the promotion and recruitment policies of many years. If this promotion prospectus was now adopted as policy and could be assumed roughly constant over time, then the approximate age on promotion for those people ending their careers in any of the grades is easily calculated. A quick method of doing this is to project backwards the points where the smoothed grade lines cross the maximum age axis, and then by eye, draw in the 'average' age on promotion. Figure 2.3.2. shows the relevant constructed

horizontal and vertical lines.

FIGURE 2.3.2.



The horizontal lines divide the figure into the various career streams and the vertical lines the average promotion points within the career streams.

The model can also be used in 'reverse', that is given the above career prospectus (Figure 2.3.2.) and wastage/expansion rates it is possible to calculate the stationary age grade distribution. It is noted that although the 'Quick' method of estimating the promotion points has been shown above, a more rigorous mathematical method based on the logistic curve is used on computer to estimate the average promotion points. The simplicity of the assumptions means that this method lends itself to valuable sensitivity analysis.

In conclusion, the Camel/Cambridge models provide worthwhile information about the manpower system, and from very little, usually available data. In considering a new manpower system with management, an analysis similar to that of the above model is always carried out, gaining insight into the changing age structure that might occur under hypothesized promotion policies under certain wastage functions.

2.3.3. SIMULATION MODELS

These models represent the first step in the direction of a full stochastic treatment of manpower supply forecasting. Until now the models discussed have carried the assumption that the expectation of a variable may be considered as a deterministic input.

All stochastic models may be simulated on computer by the selection of random numbers as determined by the probabilistic assumptions of the model to represent the manpower system. In this way realisations of possible future states of the manpower stocks may be obtained. A number of different results of the simulation could then be averaged and a range for the future number of employees in each grade calculated. A theory of order statistics gives the probability that a future trial will result in an outcome outside the current range of (n) trials is $2(n + 1)^{-1}$. Thus, 39 simulations are required to achieve ninety five percent confidence limits for the future state of the manpower system.

Warmsley (1970) produced a simulation model where promotion for employees aged thirty or more was simulated according to an index similar to Jacques' (1958) salary progression curve. Here the claim of Jacques is the existence of a set of age-salary curves and that an employee remains on one of these curves throughout their working life. This hypothesis is not universally accepted. In present industrial circumstances the salary curves would have to be corrected for inflation. It is noted that this is a very similar concept to that of the promotion prospectus discussed in the last subsection. It has been explained in this context that individuals may change grade bands, but that the prospectus represents the net result of these changes. The problem is only significant then when the simulation is related to specific individuals and not treated as the net result of individual movements.

Wishart and Ko (1973) give a more extensive simulation model known as MANSIM, an acronym for MANpower Simulation by Individual Movements. This model is used extensively for small group analyses in the Civil Service. As the authors point out MANSIM is not so much a model, more of a structure in which the user can build a model that reflects his own personal view of the manpower system. Wastage can therefore be simulated according to several attributes that might be deemed as useful predictors, including length of service.

This model would then seem to be the answer to the prediction of manpower supply. However, as the data storage and computational time requirements of the model are large, such a simulation model is limited to groups containing around 500 employees.

2.3.4. RENEWAL MODELS

In renewal models the grade sizes are predetermined by the modeller. Here employees are said to be 'pulled' through the manpower system to fill vacancies, in this way a loss from the highest grade will permeate the system resulting in a recruitment into the lowest grade and promotions in the intermediate grades to maintain stock values. Wastage rates therefore actuate the system causing vacancies in the grades and so determine the number of promotional opportunities. The way in which those eligible for promotion or the order of promotion, is decided constitutes a second input to the model.

Bartholomew (1959, 1963a, 1963b, 1967, 1969, 1970b, and 1973b) can justifiably claim much credit for the development of the theoretical aspects of renewal models. In his book (1967, 1973b) he examines four different policies, which are the permutations of the following variables:

1) Wastage as a function of length of service.

2) Wastage depending on seniority within each grade.

3) Promotion by seniority.

4) Promotion at random.

Forbes has applied these ideas to the much acclaimed Kent model used by the Civil Service, (Forbes, Morgan and Rowntree (1975), Hopes (1973) and Smith (1976)).

The way in which 1-step promotion and wastage are intimately related in an n-grade manpower system can be seen by observing that;

$$P_K(t) = P_{K+1}(t) + W_{K+1}(t)$$

$$= \sum_{j=1}^{n-K} W_{K+j}(t)$$

where $P_K(t)$ is the number promoted from grade K and $W_K(t)$ is the number of leavers from grade K, both within (t, t+1). The output of the renewal type model is usually the promotion vector and recruitment vector needed to maintain the grade size. Easily interpreted results being excessive promotion rates which may result in inexperienced personnel reaching the higher grades without spending ample time in each grade and secondly the converse of this situation when the opportunities of promotion are low due to lack of turnover in the higher grades.

Noting the above, a good method of estimating the wastage rates is deemed important. This problem is considered in Chapter 3, where a method that gives optimal estimates of the time-dependent wastage rates is presented. This will make the current renewal models 'dynamic' and thus they will be able to reflect the true structure of organisations more accurately.

2.3.5. MARKOV MODELS

Consider a process which can exist in one of an at most countable number of states $(s_1, s_2, \ldots)$. Labelling the sequence of states occupied by the process by $x_0, x_1, \ldots$ where x_0 is the initial state, then if;

$$P(x_{n+1}/x_n \ldots x_0) = P(x_{n+1}/x_n) \quad , \quad n = 0,1,\ldots$$

the process is said to constitute a Markov chain.

The elements of the transition matrix P are then defined as;

$$p_{ij} = P(x_{n+1} = s_j/x_n = s_i)$$

If n(T) represents the vector of the number of employees in the various states at time T it is found that;

$$E(n(T + 1)/n(T)) = n(T)P$$

This type of model has been used frequently in the manpower supply forecasting field. Young and Almond (1961) and Gani (1963) have documented uses in universities, Sales (1971) and Hopes (1973) in the Civil Service and Forbes (1970) on data from the W.R.N.S..

Bartholomew (1975) has shown that errors may occur in the 'deterministic' assumption that the process follows its expected stochastic path. He points out the well known fact;

> 'Conditional on $n_i(T)$ the flows from state i will be multinomial with probabilities $p_{i1}, p_{i2} \ldots p_{iK}$.'

Pollard (1966 and 1973) gives formulae for the computation of the moments of predicted stock vectors subject to this type of error. The next generalisation is the assumption that the p's are sampled from a parent time invariant distribution. A further error is incurred by the estimation of the parameters of this distribution.

Bartholomew (op. cit.) assumes the p_i.'s follow Dirichlet distributions and gives an equation for the second order moments of the grade size at time T as; (his equation)

$$M_2(T+1) = M_2(T)(EP \times EP) + \{En(T) - En(T)D^{-1} + (En(T) \times En(T))J'D^{-1}\}R_E$$

This is incorrect as, from equation 7 of Bartholomew (1975a) it can be shown th

$$M(T+1) = M(T)\Pi_E + M^*(T)A$$

expanding yields;

$$(M_1(T+1)\ M_2(T+1)) = (M_1(T)\ M_2(T))\begin{pmatrix} EP & RE \\ 0 & EP \times EP \end{pmatrix}$$

$$+ (M_1^*(T)\ M_2^*(T))\begin{pmatrix} 0 & -JC \\ 0 & C \end{pmatrix}$$

so that

$$M_2(T+1) = M_1(T)R_E + M_2(T)(EP \times EP) - M_1^*(T)JC + M_2^*(T)C$$

$$= M_2(T)(EP \times EP) + \{En(T) - En(T)D^{-1} + M_2^*(T)J'D^{-1}\}R_E$$

but

$$M_2^*(T) = En(T) \times En(T) + M_2(T)$$

This means the last term is in error by;

$$M_2(T)J'D^{-1}R_E$$

The error, however, does not significantly detract from an otherwise forward thinking paper.

In Chapter 5 of this work the Markov approach is further developed. The assumption of the Dirichlet distribution for the transition probabilities is initially relaxed and a general covariance structure is given for the prediction of the grade sizes one stage ahead. It is shown that for <u>any</u> distribution compounded with the multinomial distribution, the moments of $n_{jk}(T)/n_k(T-1)$ approach the moments of the arbitrary compounded distribution. Further, methods for estimating parameters under a Dirichlet assumption are proposed. It is remarked upon, that the assumption of a Dirichlet distribution for the transition matrix, up until now thought to be a very general assumption, introduces a nested assumption of a very specific form for the stock distribution.

2.4. MANPOWER DEMAND FORECASTING

It should be made clear from the very outset that the forecasting of manpower demand relates directly to that demand determined by the corporate plan of an organisation. It should not be assumed from this, however, that manpower demand is a secondary effect of the corporate plan but rather an intimately laced element of it. Bartholomew, Hopes and Smith (1976) state:

> 'Demand Forecasts have to be made using a blend of statistical analyses and management judgement'

It is within this context that the various methods that have been proposed for the purpose of manpower demand forecasting are now reviewed. Whilst pursuing this objective it should be noted that although the number of publications on supply forecasting are voluminous the converse applies to demand forecasting.

It is considered that all demand models published so far have in fact one underlying theme; the prediction of future 'production' levels in their widest sense and the method of relating these levels to the manpower requirements. The only difference within the various models is the mix between management information/experience and the statistical techniques employed.

The simplest method is to ask managers their future manpower requirements. In this case the process is carried out by purely 'mental' calculation. It has the advantage of being cheap having small data collection requirements and subjective factors which may sometimes not be confided to the modeller are incorporated. Of course, when many time-dependent variables have to be considered which are governed by stochastic rules this mental exercise is unlikely to yield usable bounds for the estimates produced.

The second method, at the other extreme, is the simple statistical process of extrapolating the curve of the number of employees in the organisation. This is thought to be a dangerous method to adopt as there is usually a great deal of available data on the nature of the employee structure which has a causal relation to the size of the labour force and may enable a far more accurate method of demand forecasting than extrapolation to be used.

This brings us to the most fruitful approach which encompasses the methods of work study and time series analysis. The basic proposal is that various measures of the workloads in an organisation can be determined. Once identified these may constitute inputs into the corporate plan, in which case they may be assumed to be given or, they may have to be predicted from past measurements. These workloads must then be converted into manpower requirements. As before, there are two methods; the first relates to Industrial Engineering and the second to Regression Analysis. Further explanation of these four methods now follows. It is not, however, intended to expound the methods of corporate planning. It is enough to say that a more professional view of the market in which statistical methods play their part would often reap benefits when attempting to determine the company's market share.

If there is no knowledge of future workloads other than that they will continue to follow a pattern that can be inferred from past values, time series analysis is appropriate. Drui (1963) assumes that the explanatory variables used in his regression analysis remain unchanged, however Livingstone and Montgomery (1966) show several shortcomings of this type of hypothesis.

Cameron and Nash (1974), in the first full analytic paper on Manpower Demand Forecasting, use Box and Jenkins (1970) techniques to evaluate future workloads. Young and Vassiliou (1974) have used exponentially weighted moving averages to predict these variables. Cameron and Nash have mentioned defects in their method; firstly that over fifty observations are usually required and secondly, errors will occur if any change takes place in the parameters of the time series. The same faults are applicable to the E.W.M.A. model. In Chapter 6 an adaptive model for the prediction of seasonal and trend inherent time series is developed.

Once the workloads have been established there are two prominent methods that have been employed to convert these into manpower requirements. The method used in Industrial Engineering is considered first. As identified by Purkiss (1976), many firms have a standard manning level; that is a standard output per man hour figure for their operations. Often these are related to the British Standard scale. Given these figures it is a simple matter to compute the manpower demand. However, these standards should be recalculated whenever operational changes occur and small but important changes in efficiency and technology occur almost continuously. This is the reason why productivity bonus' tightly related to these scales at their conception often become very loose over time and prove to be a plague to management. New schemes related to the 'new' standard must then be cleverly negotiated and introduced to maintain the effort of the worker at a desired level.

The second method is that of Regression Analysis. Drui (1963) made the first attempts along these lines using 'independent' variables to try and predict the manpower requirements. Again, Livingstone and Montgomery criticise his failure to account for autocorrelation of the regression residuals. Cameron and Nash (1974) also use regression analysis to construct a model that relates the manpower demand to the workloads. They further investigated the time stability of their model amd found that the 'model was not stable over time', and that 'a major change in the parameters had taken place during the period covered by the data'. They then tried to make the series time stable. By finding the first three principle components from the correlation matrix, they achieved a correlation coefficient of 0.928, 'but once again... tests suggested that the model was not stable over time'. After searching the data for a time stable period they eventually found one in the last four years of data. They then made the tenuous assumption that the regression coefficients for the last 14 quarters will be time invariant for the forecasting period and further, that the factored variables continue to explain future variation in the workload series. Surely by now they should have discontinued the approach of trying to achieve time stability for the time series when it is, by their own admittance, in essence time instable and tried to analyse it as such. A method of calculating adaptive regression coefficients for this type of series is explained in Chapter 6.

In summary it is thought that the best approach for Demand Forecasting lies in a subtle interplay between management foresight and statistical analysis. We should tend towards the estimation of desired workloads by the Planning Department of an organisation and the analysis of the complex and time variable relationships between these variables by statistical techniques. In each case the management plans should be constructed in the light of the results from both approaches.

2.5. CONTROL OF MANPOWER SYSTEMS

In simple terms the control of a manpower system refers to the method of matching the supply and demand forecasts. This process is achieved by the introduction of some 'control' input into the manpower system. In a typical representation of a manpower model;

$$n_{t+1} = P_t n_t + R_t$$

where

n_t = stock vector at time t

P_t = transition matrix at time t

R_t = recruitment vector at time t

there are three implicit variables: wastage, promotion(demotion) and recruitment(redundancy). Of these control exercised over the latter variable is the most acceptable, although to some extent both promotional and wastage control are available.

There have been many papers on the solution of this problem in its deterministic formulation including: Charnes, Cooper and Niehaus (1968), Grinold and Marshall (1977), Grinold and Stafford (1974) and Bartholomew (1973). In real life however the deterministic model, although providing some insight into the problem, cannot be considered a good approximation. The reasons for this are twofold; wastage is inherently stochastic and even if a probability transition matrix were known for all time, then the distribution of the grade sizes would be multinomial, once again making the system stochastic. The only variable that can be taken as known (deterministic) is the recruitment/redundancy vector. It is for this reason that the main problem considered in Chapter 7 is the control of a stochastic manpower system, when the only known input that may be used to exercise control is recruitment/redundancy.

Bartholomew (1976, 1977, 1976b) has made many advances in the development of control strategies for the stochastic model. He advocates a 1-step control strategy as being the most worthwhile. Although at first his arguments are very persuasive, a deeper analysis of results presented in Bartholomew (1976b) reveals an error in his thinking which will be discussed in Chapter 7 Section 1. The 1-step strategy is clearly analogous to the 'fire-fighting' methods in which British management have needed to excel over recent years. It is generally accepted that this is not a form of good management, 'by doing something at this point in time we might ruin our chances of achieving a desired future goal', but an alternative strategy is not immediately obvious. The reason for this lies in the measure chosen to evaluate differing control strategies. The measuring function is therefore of paramount importance and a suitable form for it is discussed in Chapter 7 Section 2. It is sufficient to state here that such a function should represent the minimising of some cost over a planning period not just from one time instant to the next; as the latter will clearly not be optimal over the complete planning horizon.

In Chapter 7 Section 3, having made a very general choice of a 'cost' function to be minimised over the total planning period, the solution of the deterministic problem is then obtained. This may initially appear to be somewhat contradictory following the earlier criticisms of this approach, however there are two strong reasons for its inclusion:

1) The solution is certainly valid when considering deterministic systems.
2) The notation introduced in arriving at the deterministic solution can be carried over to the stochastic problem in such a way as to ease the exposition of its solution.

Chapter 7 Section 4, deals with the solution of this stochastic control problem.

In the last four sections, various manpower planning models and approaches have been reviewed, some have already been rejected, some have been critically examined in depth and others have only been mentioned. In doing this it was hoped to bring out two very important facts that underlie the philosophy of this thesis. Firstly, industrial structures are not static, they are organic and therefore change over time. Secondly, manpower systems have variables which are not deterministic but stochastic. These two remarks constitute the main assumptions of the models that are developed over the next five chapters. These models are then brought together into a total manpower system in Chapter 8.

The objective is to add further realism to manpower planning. Usually this would entail highly complex statistical treatment and at times this may be true of our approach. However, after spending some time defining the problem with acceptable mathematical rigour, the rewards are extremely worthwhile. The form of the solution is concise and easily interpreted and in addition is computationally efficient.

Throughout the following chapters the intention is to further investigate as mentioned previously, the basis of those models that have been useful in the formulation of a stochastic time variable manpower model. This will then be succeeded by establishing such a model and exploring its applications. Those problems which are difficult solely because of their obvious notational difficulties will be mainly appended, as they do not provide significant insight into the structure of the manpower system. The same applies to the results produced on trial data by the new manpower models.

The next chapter commences with a discussion of the wastage process and develops into the problem of determining the solution to a general time variable stochastic system. This model and its solution is then seen in the following chapters to be a most useful tool in the analysis of manpower systems.

CHAPTER THREE

THE SOLUTION OF A GENERAL FORECASTING AND FILTERING PROBLEM

3.1. INTRODUCTION

In this chapter the problems of estimating future labour wastage are considered. Section 1.1. includes a discussion of Forbes (1971a). a paper that has numerous non-parametric methods of estimating the survivor function. The relevance of one of the basic methods of the paper is questioned. Various elements of this article are then brought together to form a new model of the wastage process that allows the possibility of the true wastage probabilities changing over time. This model is just a special case of a more general problem, that of finding optimal forecasts for dynamic systems only observable through noise corrupted measurements.

This general problem is stated more precisely in Section 2, where an example is presented and its method of solution discussed. It is found that the general problem carries a natural mathematical structure that of the Hilbert Space. The structure and properties of the Hilbert Space relevant to the solution of the general problem are given in Appendix A. To estimate a vector in this space, given a cost (error) function that it also an inner product, there is a unique vector that minimises this cost criteria. Moreover, this vector is just the orthogonal projection of the estimated vector into the subspace generated by the observations.

The final section of this chapter is devoted to the solution of the General Problem. Here the optimal filtered and predicted estimates of vectors which follow processes that are encompassed by the General Problem are given. Further the error covariance matrix of these estimates are derived. The solution is presented as a set of recurrsive equations from which given initial estimates, the new estimates converge automatically and optimally to the true value of the estimated vector.

3.1.1. A DYNAMIC MODEL OF THE WASTAGE PROCESS

This section begins with a discussion of Forbes (1971a) and in particular his section 3(b) on the use of the Cohort service table for estimating the survivor function.

Defining the mean force of separation over the ith interval by;

$$\phi_i = \int_{x_i}^{x_{i+1}} \phi(u)du/c_i$$

Forbes shows,(his equation 8);

$$\phi_i = (\ln G_i - \ln G_{i+1})/c_i$$

This is readily seen to be equivalent to;

$$\phi_i = - \ln p_i/c_i$$

He also gives, (equation 9);

$$\phi_i = (1 - p_i)/(1 - (1 - a_i)(1 - p_i))c_i$$

from which a_i;

> 'the average fraction of the ith interval completed before leaving for those who leave during the interval'

can be evaluated as;

$$a_i = (\ln(1/p_i))^{-1} - p_i(1 - p_i)^{-1}$$

Forbes suggests $a_i \simeq \frac{1}{2}$. Equality only holds when there is zero severance propensity, this is a rare occurence, but certainly for $p_i > \frac{1}{2}$ the approximation is reasonable. The values of the force of separation for the true ϕ_{iT} and the approximation ϕ_{iA} can be compared for p_i = .1(.1).4 on the unit time period by reference to Table 3.1.1. below.

TABLE 3.1.1.

p_i	a_i	ϕ_{iT}	ϕ_{iA}
.4	.424	0.916	0.857
.3	.402	1.204	1.077
.2	.371	1.609	1.333
.1	.323	2.303	1.636

Taking $a_i = \frac{1}{2}$ clearly underestimates ϕ_i and care must therefore be taken when applying this approximation in areas of high wastage. In his section 4(a) the use of the 'mean of the numbers in service at the beginning and the end of the calendar time interval' is equivalent to setting $a_i = \frac{1}{2}$ and consequently, the use of this will underestimate ϕ_i. It is noted that his value R_i will overestimate ϕ_i. This can easily be seen by observing that under a constant force of separation, the number of employees who will eventually leave during the interval who are present at the midpoint of the interval, is less than half of the total leavers in the interval.

However, over the wastage rates normally experienced in industry the approximation proposed by Forbes is quite acceptable. What is not clear is the reason for the introduction of such an approximation. This seems to be superflous as his equation 8 affords a simple equation for ϕ_i.

As pointed out by Forbes;

$$G(x_r) = \exp(-\sum_{i=0}^{r-1} \phi_i c_i)$$

substituting the derived equation for ϕ_i gives;

$$G(x_r) = \exp(\sum_{i=0}^{r-1} \ln p_i)$$

$$= \prod_{i=0}^{r-1} p_i$$

which is a much simpler formulation - this is equation 13 of the paper. In order to estimate $G(x_r)$ from the above it is necessary to find estimates of the p_i.

Now as $l^o_{i+1} \sim$ binomial $(n = l^o_i,\ p = p_i)$

The minimum variance unbiased estimator of p_i is;

$$\hat{p}_i = l^o_{i+1}/l^o_i$$

and an estimate of its variance is;

$$\hat{V}(\hat{p}_i) = \hat{p}_i(1 - \hat{p}_i)/l_i^o$$

Result 1 of the paper can now be employed to find an estimate of the variance of the estimate of the survivor function;

$$\hat{V}(\hat{G}(x_r)) = \sum_{i=0}^{r-1} \hat{V}(\hat{p}_i) \left(\frac{\partial \hat{G}(x_r)}{\partial \hat{p}_i} \right)^2$$

$$= \sum_{i=0}^{r-1} \frac{\hat{p}_i(1 - \hat{p}_i)}{l_i^o} \left(\frac{\hat{G}(x_r)}{\hat{p}_i} \right)^2$$

$$= \hat{G}(x_r)^2 \sum_{i=0}^{r-1} \left(\frac{1}{l_{i+1}^o} - \frac{1}{l_i^o} \right)$$

$$= \hat{G}(x_r)^2 \left(\frac{1}{l_r^o} - \frac{1}{l_o^o} \right)$$

$$= \hat{G}(x_r)\ (1 - \hat{G}(x_r))/l_o^o$$

which is in form the same as Forbes' equation 5 for Cohort analysis. It is not in fact the same equation. The confusion arises as $\hat{p}_i$ has in fact different values dependent on whether it is obtained from cohort or census analysis.

Assuming that time is counted from zero; for cohort analysis:

$$\hat{p}_i = l_{i+1}^o(i + 1)/l_i^o(i)$$

and for census analysis:

$$\hat{p}_i = l_{i+1}^o(1)/l_i^o(0)$$

equality holding only when i = 0.

To overcome this a change of notation is introduced. The problem of forecasting the wastage rates is then formulated in a dynamic sense, encompassing elements of both census and cohort analysis.

Before doing this it is noted that the maximum of the variance of the estimated survivor function occurs at the half life which is often used as a measure of wastage.

Let the number of employees observed with length of service i at real time t be represented by $l_i^o(t)$. The proportion of these leaving in the next unit of time will be denoted $O_{i,i+1\ t+1}$, where;

$$O_{i,i+1\ t+1} = \frac{l_{i+1}^o(t+1) - l_i^o(t)}{l_i^o(t)}$$

Similarly the true probability of <u>leaving</u> over the interval (i,i + 1) will be symbolised $p_{i,i+1\ t+1}$. In this way, we say O is an observation of p, the true probability of leaving. It is readily seen that O is an unbiased estimator of p.

An observation system is now postulated as below;

$$O_{i,i+1\ t+1} = p_{i,i+1\ t+1} + dO_{i,i+1\ t+1}$$

where $dO_{i,i+1\ t+1}$ is the total (stochastic) disturbance in the system. In order to ease the exposition matrix notation is introduced. If there are n intervals (i, i+1), the column vector of the observations, $O_{i,i+1\ t+1}$, where i = 0(1)n - 1, will be denoted O_{t+1}. The 'observation' system is now;

$$O_{t+1} = p_{t+1} + dO_{t+1}$$

where $E(dO_{t+1}) = 0$ and $E(dO_{t+1}, dO_{t+1}^T)$ is finite.

Now it is recognised that the true wastage probabilities might vary over time. In order to incorporate this into the model, dp_t is chosen to represent a small random additive disturbance that transforms p_t to p_{t+1}. Therefore the 'system' equation can be written as;

$$p_{t+1} = p_t + dp_t$$

Further it is assumed that the expectation of dp_t is zero and that it has finite variance.

The task can be summarised as finding the optimum estimate of p_{t+1} denoted $\hat{p}_{o,t+1,t}$ from the set of observations $O_t, O_{t+1}, \ldots, O_1$. This problem is of great importance in the prediction of manpower supply. Most manpower supply models have not treated p_{t+k} as a time variant vector input, until now the modeller being asked to state a scalar to represent this quantity. Clearly, if the optimal estimates of these wastage probabilities could be derived 'on line' it would give additional creditability to any manpower supply model using them, given the same information.

3.2. FORMULATION OF A GENERAL STOCHASTIC PROBLEM

3.2.1. THE GENERAL PROBLEM

The problem stated above belongs to a wider class of problems; that of finding optimal forecasts for dynamic systems only observable through noise corrupted measurements. It is the more general problem that will be solved in this chapter. More strictly the problem to be solved is stated below.

Given the:

System Equation

$$x_{k+1} = \phi_{k+1,k}\, x_k + dx_k$$

where

x_k is the state of the system at time k (n x 1)

$\phi_{k+1,k}$ is a state transition matrix known at time k (n x n)

dx_k is the random error vector at time k (n x 1)

It is assumed that;

$$E(dx_k) = 0 \qquad \forall k$$

and that

$$E(dx_k dx_j^T) = R_k\, \delta_{kj}$$

where δ_{kj} is the 'kronecker delta', i.e.;

$$\delta_{kj} = \begin{cases} 1 & k=j \\ 0 & k \neq j \end{cases}$$

and R_k is a real symmetric positive definite (n x n) matrix.

Observation Equation

$$z_{k+1} = \psi_{k+1} x_{k+1} + dz_{k+1}$$

where

z_k is the observation vector at time k (m x 1)

ψ_k is the transition matrix that relates the observations at time k to the system parameters (m x n)

dz_k is the observation error vector at time k (m x 1)

It is further assumed that;

$$E(dz_k) = 0 \qquad \forall k$$

and

$$E(dz_k dz_j) = S_k \delta_{kj}$$

where S_k is a real symmetric positive definite (m x m) matrix.

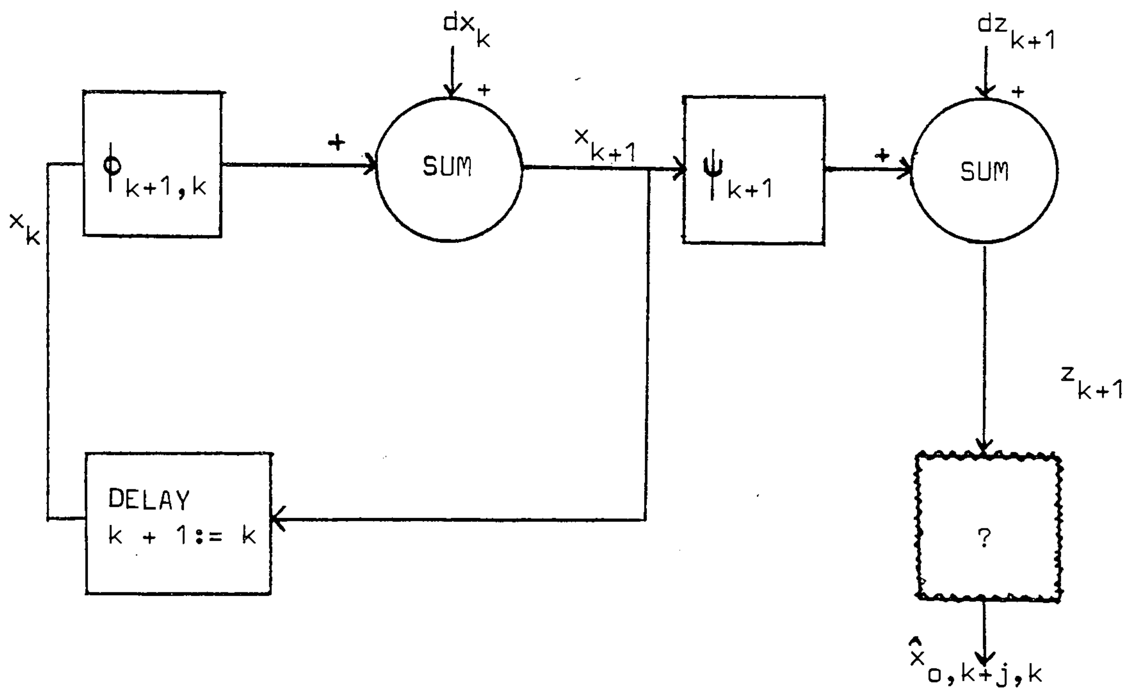
Covariance Restrictions

The observation noise, system noise and any estimate of the initial state ($x_{0,0}$) are mutually uncorrelated for all time.

The task is to find the optimal estimate of x_{k+j} given the observations up to z_k.

Further, it is proposed to determine the error covariance matrix of the optimal estimate of x_{k+j}.

The systems model of this general problem is shown below:



The task is to find the form of the 'fuzzy box' that yields the optimum estimate.

The class of problems included as subsets of the General Problem is very wide. By an appropriate choice of the system and observation matrices, least squares, ARIMA, linear and polynomial regression and EWMA analyses can be performed, in a dynamic sense.

The General Problem can be further extended to include correlated disturbances or functions of known error sources. This extension is fairly easy but cumbersome and hence the results are given in Appendix B. The General Problem is naturally complex and its solution somewhat lengthy. Before embarking on its solution we will motivate the necessary theory by giving a short example taken from minimum variance estimation.

3.2.2. AN EXAMPLE

Define x_t, z_t and x_t as nx1 matrices and $V_{.}$, w_t as nxn matrices.

Let x_t be a forecast of X_t just before observing z_t - a measurement of X_t being the system parameter.

x_t and z_t being independent unbiased estimates of X_t are deemed to have variances V_{xt} and V_{zt} respectively.

It is required to find the best estimate $x_{o,tt}$ of the process X_t after the observation.

Let $x_{o,tt}$ be a weighted combination of x_t and z_t, w_t being a linear weighting factor.

So

$$x_{o,tt} = (1 - w_t)\, x_t + w_t z_t$$

where

1 = nxn matrix.

The expectation of $\hat{x}_{o,tt}$ is readily seen to be X_t, and its variance given by (dropping the t suffices);

$$V_{\hat{x}_o tt} = E((x - E(x))\ (x - E(x))^T)$$
$$= (1 - w_t)V_{xt}(1 - w_t)^T + w_t V_{zt} w_t^T$$
$$+ \text{(Covariance terms} = 0)$$

It is now required to minimise this expression with respect to w_t.

$$V_{\hat{x}_o tt} = V_{xt} + w_t(V_{xt} + V_{zt})\ w_t^T - w_t V_{xt} \qquad - V_{xt} w_t^T$$

As $V_{xt} + V_{zt}$ is real symmetric there exists C such that (see footnote)

$$V_{xt} + V_y = CC^T$$

Therefore

$$V_{\hat{x}_o tt} = V_{xt} + w_t CC^T w_t^T - w_t V_{xt} - V_{xt} w_t^T$$
$$(+V_{xt}(CC^T)^{-1} V_{xt}^T - V_{xt}(CC^T)^{-1} V_{xt}^T)$$
$$= V_{xt} + \left[w_t C - V_{xt}(C^{-1})^T\right]\left[w_t C - V_{xt}(C^{-1})^T\right]^T - V_{xt}(CC^T)^{-1} V_{xt}^T$$

Clearly the middle term is either positive or zero and the other terms do not include w_t, $V_{\hat{x}_o tt}$ is at a minimum when;

$$w_t C - V_{xt}(C^{-1})^T = 0$$

that is

$$w_t = V_{xt}(CC^T)^{-1} = V_{xt}(V_{xt} + V_{zt})^{-1}$$

A well known theory of linear algebra states: (Hoffman & Kunze(1961));

If A is a symmetric nxn matrix over the Reals, then there exists an invertible nxn matrix P such that $P^T AP$ is diagonal and further P may be chosen to be orthogonal i.e. $P^T = P^{-1}$

Let $V_{xt} + V_{zt} = A$, then , $P^T AP = BB^T$ for some complex B

Therefore

$$A = (P^T)^{-1} BB^T P^{-1}$$
$$= PBB^T P^T$$
$$= PB(PB)^T$$
$$= CC^T, \text{ say}$$

and $V_{xt} + V_{zt}$ may be written as CC^T.

3.2.3. DISCUSSION OF THE EXAMPLE

It was desired to find an estimate for the state of the system at time t. The best estimate was, in this case, defined as the estimate with minimum variance. A linear combination of the best forecast up to the time of observation and the new observation, was formed and the weighting factor was found to be a function of these two estimates' variances. In the sequel the methodology for obtaining the best forecast up to the time of the new observation is given, this is then combined in a similar way, as above, with the new observation in order to give the optimum filtered estimate.

3.2.4. THE IMPORTANCE OF THE HILBERT SPACE TO THE STRUCTURE OF THE GENERAL PROBLEM

As stated in the introduction a natural topological structure in which to solve the General Problem is the Hilbert Space. It's structure is reviewed in Appendix A. Section 1. Included in this discussion is sufficient theory to develop an understanding of the important properties of this space as related to the General Problem only. Definitions of a metric, norm, complete space and convexity are worked through fairly quickly. In particular, one of the main tasks of the Appendix is the proof that the Lesbesgue space of dimension two; i.e. the space generated by random functions with finite second moments, is indeed a Hilbert Space.

In the second section of Appendix A. the importance of the study of Hilbert Spaces is shown. The structure allows the concept of orthogonal projections and an orthogonal projection lemma is proved. Further, it is shown that under certain conditions the optimum estimate $\hat{x}_t$ of x_t is simply the orthogonal projection of the x_t into the subspace generated by the observations.

The General Problem is easily solved using this mathematical structure. A similar problem founded in engineering has been tackled by Kalman and Bucy (1971). Their solution, an attempt to solve the Wiener 1949 problem, was based on conditional expectations, assumed the specific form of the Gaussian distribution for the system noise and error free observation. The solution presented in the following section requires only that the first two moments of the system disturbance and the observation error distributions are finite. The next chapter deals with the simplified solution of the General Problem that relates to wastage forecasting.

3.3. SOLUTION OF THE GENERAL FORECASTING AND FILTERING PROBLEM

The general problem was stated precisely in Section 3.2., here the salient points are reviewed and placed within the mathematical structure developed in Appendix A.

3.3.1. REVIEW OF THE GENERAL PROBLEM

The state of the system is governed by the equation;

(1) $x_{k+1} = \phi_{k+1,k}\, x_k + dx_k$

where,

$$E(dx_k) = \underset{\sim}{0}\ ,\quad E(dx_k\, dx_j^T) = R_k\, \delta_{kj}$$

The only information that is available are the noise corrupted observations;

(2) $z_{k+1} = \psi_{k+1}\, x_{k+1} + dz_{k+1}$

where,

$$E(dz_k) = \underset{\sim}{0}\ ,\quad E(dz_k\, dz_j^T) = S_k\, \delta_{kj}$$

Further it is assumed that;

(3) x_0, $\{dx_k\}$ and $\{dz_j\}$ are stochastically independent.

The task is to find the optimal estimate $(\hat{x}_{o\ k,j})$ of x_k given the observations $z_1,\ldots,z_j = \{z_j\}$, the constants $\{\phi_{k,k-1}\}$ and $\{\psi_k\}$ and the noise variances $\{R_k\}$ and $\{S_k\}$.

The problem is one of:

'Smoothing' when $j > k$

'Filtering' when $j = k$

'Forecasting' when $j < k$

The last two of these are discussed in this section.

3.3.2. MATHEMATICAL FRAMEWORK

As shown in Appendix 1.1 the state space X, the set of all mappings from $\Omega \mapsto \mathbb{R}^n$ that are measurable with respect to a σ-field $\mathcal{R}$ of Ω, such that $|x(\omega)|^2$ is integrable, with the inner product;

$$\langle x_1, x_2 \rangle = \int_{\Omega} x_1^T(\omega)\, x_2(\omega)\, d\mu(\omega)$$

and associated norm;

$$\|x\| = \left(\int_\Omega |x(\omega)|^2 d\mu(\omega)\right)^{\frac{1}{2}} < \infty$$

forms a Hilbert Space.

Since both the system noise $d\,x_k$ and the observation noise $d\,z_k$ have finite second moments, $\|d\,x_k\|^2$, $\|dz_k\|^2 < \infty$, for all k.

3.3.3. THE SOLUTION OF THE GENERAL PROBLEM

The only knowledge of the system must be gained through the observations $\{z_k\}$. The subspace generated by these observations is defined as;

$$\hat{X}_j = \left\{ \hat{x}_j : \hat{x}_j = \sum_{i=1}^{j} A_i z_i \quad \forall A_i \in \mathbb{R}^n \times \mathbb{R}^m \right\}$$

and called the Estimation Space as the optimal estimate $\hat{x}_{o.,j}$ is a member. The orthogonal complement of this subspace, $\hat{X}_j^{\perp}$, will be denoted $\tilde{X}_j$ and called the Estimation Error Space.

It was proven in THEOREM A.2.1. that within the subspace $\tilde{X}_j$ there exists a unique element of the smallest norm. This element will be denoted $\tilde{x}_{o\ k,j}$ where the symbols $_{o\ k,j}$ represent the error in the optimum (o) estimate of the kth.(k) state given the observation set $\{z_j\}$.

Using THEOREM A. 2.2. it was shown that any vector $x_k \in X$ can be represented uniquely as;

$$x_k = \tilde{x}_{o\ k,j} + \hat{x}_{o\ k,j} \tag{4}$$

where

$$\hat{x}_{o\ k,j} = x_k - \tilde{x}_{o\ k,j} \in \hat{X}_j$$

and is the orthogonal projection of x_k into $\hat{X}_j$.

It is desired to minimise the error norm.

$$\|\tilde{x}_{k,j}\|^2 = \|x_k - \hat{x}_{k,j}\|^2 = \int_\Omega (x_k(\omega) - \hat{x}_{k,j}(\omega))^T (x_k(\omega) - \hat{x}_{k,j}(\omega) d\mu(\omega) \tag{5}$$

Substituting (4) into (5):

$$\|\tilde{x}_{k,j}\|^2 = \|\tilde{x}_{o\ k,j} + \hat{x}_{o\ k.j} - \hat{x}_{k,j}\|^2$$

$$= \|\tilde{x}_{o\ k,j}\|^2 + 2\langle \tilde{x}_{o\ k,j}, \hat{x}_{o\ k,j} - \hat{x}_{k,j}\rangle + \|\hat{x}_{o\ k,j} - \hat{x}_{k,j}\|^2$$

The second term is zero since $\tilde{x}_{o\ k,j} \in \tilde{X}_j$ and $(\hat{x}_{o\ k,j} - \hat{x}_{k,j}) \in \hat{X}_j$ which are by definition orthogonal subspaces.

Therefore

$$\|\tilde{x}_{k,j}\|^2 = \|\tilde{x}_{o\ k,j}\|^2 + \|\hat{x}_{o\ k,j} - \hat{x}_{k,j}\|^2$$

and the minimum of this expression is attained when $\hat{x}_{k,j}$ is chosen to be the orthogonal projection of x_k into $\hat{X}_j$, $(\hat{x}_{o\ k,j})$.

The object of this section is to seek the solution of the General Problem. This solution will be given in the form of a set of recurrence relationships. The proof of these equations is quite lengthy so for ease of exposition each of them will be derived separately and then brought together in the final THEOREM (3.3.6.) of this Section.

Before embarking on this task it is necessary to define one further subspace - the subspace generated by the observation errors.

If $\hat{x}_{o\ k,j}$ is the optimum estimate of x_k given $\{z_j\}$, then the optimum estimate of z_k is given by;

$$\hat{z}_{o\ k,j} \triangleq \psi_k \hat{x}_{o\ k,j}$$

The corresponding error in this estimate is,

(6) $$\tilde{z}_{o\ k,j} = z_k - \hat{z}_{o\ k,j}$$

and for any estimate $\hat{z}_{k,j}$

$$\tilde{z}_{k,j} = z_k - \hat{z}_{k,j}$$

The subspace of the observation errors is now defined as;

$$\tilde{Z}_k = \{\tilde{z}_k : \tilde{z}_k = B_k \tilde{z}_{k,k-1} \quad \forall B_k \in R^n \times R^m\}$$

In THEOREM 3.3.3. it is shown that $\hat{X}_{k+1} = \hat{X}_k \oplus \tilde{Z}_{k+1}$ where $\oplus$ denotes the direct sum. From this identity the value of $\hat{x}_{o\ k,k}$ can be found. For the next theorem it is assumed that $\hat{x}_{o\ k,k}$ has been obtained by some method.

THEOREM 3.3.1.

The optimum estimate of the state vector x_{k+j} given $\{z_k\}$, $\hat{x}_{o\ k,k}$ is

(A) $$\hat{x}_{o\ k+j,k} = \phi_{k+j,k}\hat{x}_{o\ k,k}$$

where

$$\phi_{k+j,k} = \phi_{k+j,k+j-1} \cdots \phi_{k+1,k} = \prod_{i=1}^{j} \phi_{k+i,k+i-1}$$ ▼

PROOF

We know that $x_{k+j} = \hat{x}_{o\ k+j,k} + \tilde{x}_{o\ k+j,k}$

So it is enough to prove that;

$$x_{k+j} - \phi_{k+j,k}\hat{x}_{o\ k,k} \in \tilde{X}_k$$

Let $\hat{x}_{k,k} \in \hat{X}_k$ and consider;

(7) $$\langle x_{k+j} - \phi_{k+j,k}\hat{x}_{o\ k,k}, \hat{x}_{k,k} \rangle$$

Now

$$x_{k+j} = \phi_{k+j,k+j-1}x_{k+j-1} + d\,x_{k+j-1}$$

$$= \phi_{k+j,k+j-1}\phi_{k+j-1,k+j-2}x_{k+j-2} + \phi_{k+j,k+j-1}\,d\,x_{k+j-2} + dx_{k+j-1}$$

$$= \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

(8) $$= \phi_{k+j,k}x_k + \sum_{i=1}^{j} \phi_{k+j,k+i}\,d\,x_{k+i-1} \qquad (\phi_{l,l} = 1)$$

Substituting into (7) gives;

$$\langle \phi_{k+j,k}x_k + \sum_{i=1}^{j} \phi_{k+j,k+i}\,dx_{k+i-1} - \phi_{k+j,k}\hat{x}_{o\ k,k}, \hat{x}_{k,k} \rangle$$

$$= \langle \phi_{k+j,k}\tilde{x}_{o\ k,k} + \sum_{i=1}^{j} \phi_{k+j,k+i}\,dx_{k+i-1}, \hat{x}_{k,k} \rangle$$

(9) $$= \langle \phi_{k+j,k}\tilde{x}_{o\ k,k}, \hat{x}_{k,k} \rangle + \sum_{i=1}^{j} \langle \phi_{k+j,k+i}\,dx_{k+i-1}, \hat{x}_{k,k} \rangle$$

The first term is clearly zero, as;

$$\tilde{x}_{o,k,k} \in \tilde{X}_k \quad \text{and} \quad \phi^T_{k+j,k}\,\hat{x}_{k,k} \in \hat{X}_k$$

which are orthogonal subspaces. It is only necessary then to consider the second term. As $\hat{x}_{k,k} \in \hat{X}_k$,

$$\hat{x}_{k,k} = \sum_{i=1}^{k} A_i z_i \quad \text{for some } A_i$$

and z_i can be expressed as;

$$z_i = \psi_i x_i + d z_i \qquad \forall_i ; \quad 1 \leq i \leq k$$
$$= \psi_i \phi_{i,i-1} x_{i-1} + \psi_i \, dx_{i-1} + dz_i$$

This expansion does not have any of the terms $dx_k \ldots dx_{k+j-1}$, and noting (3), the second term vanishes.

Therefore

$$\langle x_{k+j} - \phi_{k+j,k} \hat{x}_{o\;k,k} , \hat{x}_{k,k} \rangle = 0 \qquad \forall \hat{x}_{k,k}$$

and

$$x_{k+j} - \phi_{k+j,k} \hat{x}_{o\;k,k} \in \tilde{X}_k$$

so that

$\phi_{k+j,k} \hat{x}_{o\;k,k}$ is the orthogonal projection of x_{k+j} into $\hat{X}_k$. ■

We now define $C_{o\;k,j}$ the error covariance matrices;

$$C_{o\;k,j} = \int_{\Omega} \tilde{x}_{o\;k,j} \tilde{x}^T_{o\;k,j} \, d\mu(\omega) \triangleq \rangle \tilde{x}_{o\;k,j} , \tilde{x}_{o\;k,j} \langle$$

where $\rangle \; \langle$ defined above is called an outer product.

THEOREM 3.3.2.

The covariance error matrix of the optimal estimate of x_{k+j}, given observations $\{z_k\}$ and $\hat{x}_{o,k,k}$, is related to $C_{o\;k,k}$ by;

(B) $$C_{o\;k+j,k} = \phi_{k+j,k} C_{o\;k,k} \phi^T_{k+j,k} + \sum_{i=1}^{j} \phi_{k+j,k+i} R_{k+i-1}$$

▼

PROOF

It has been shown that;

$$\tilde{x}_{o\;k+j,k} = x_{k+j} - \hat{x}_{o\;k+j,k}$$

So using the result derived in THEOREM 3.3.1. this may be re-written as;

$$\tilde{x}_{o\;k+j,k} = x_{k+j} - \phi_{k+j,k} \hat{x}_{o\;k,k}$$

The expansion (8) is now substituted into the above to yield

$$\tilde{x}_{o\ k+j,k} = \phi_{k+j,k}\tilde{x}_{o\ k,k} + \sum_{i=1}^{j} \phi_{k+j,k+i}\ dx_{k+i-1}$$

Now

$$C_{o\ k+j,k} = \rangle \tilde{x}_{o\ k+j,k}, \tilde{x}_{o\ k+j,k} \langle$$

$$= \phi_{k+j,k} \rangle \tilde{x}_{o\ k,k}, \tilde{x}_{o\ k,k} \langle \phi^{T}_{k+j,k}$$

$$+ \sum_{i=1}^{j} \sum_{l=1}^{j} \phi_{k+j,k+i} \rangle dx_{k+i-1}, dx_{k+l-1} \langle \phi^{T}_{k+j,k+l}$$

$$+ \sum_{l=1}^{j} \phi_{k+j,k} \rangle \tilde{x}_{o\ k,k}, dx_{k+l-1} \langle \phi^{T}_{k+j,k+l}$$

$$+ \sum_{i=1}^{j} \phi_{k+j,k+i} \rangle dx_{k+i-1}, \tilde{x}_{o\ k,k} \langle \phi^{T}_{k+j,k}$$

From the workings in THEOREM 3.3.1., the last two terms are identically zero. Also from (1) $E(dx_k\ dx_j^T) = R_k\,\delta_{k,j}$ - this means that all constituents for which $i \neq l$, vanish in the second term.

Therefore

$$C_{o\ k+j,k} = \phi_{k+j,k} \rangle \tilde{x}_{o\ k,k}, \tilde{x}_{o\ k,k} \langle \phi^{T}_{k+j,k}$$

$$+ \sum_{i=1}^{j} \phi_{k+j,k+i} \rangle dx_{k+i-1}, dx_{k+i-1} \langle \phi^{T}_{k+j,k+i}$$

$$= \phi_{k+j,k} C_{o\ k,k} \phi^{T}_{k+j,k} + \sum_{i=1}^{j} \phi_{k+j,k+i} R_{k+i-1} \phi^{T}_{k+j,k+i}$$

■

The forecasting side is now complete, that is to say, given the optimum estimate of x_k based on the observation set $\{z_k\}$ and its error covariance matrix $C_{o\ k,k}$ we have a system to generate the optimum estimates of $\{x_{k+j}\}$, j=1.... and the respective error covariance matrices of our forecasts $\{C_{o\ k+j,k}\}$. We have now to show how to obtain $\hat{x}_{o\ k,k}$ and $C_{o\ k,k}$. Their form is not immediate and it is necessary to investigate the composition of the Estimation Space $\hat{X}_{k+1}$. This we now give in the following theorem.

THEOREM 3.3.3.

The subspaces $\hat{X}_k$ and $\tilde{Z}_{k+1}$ are orthogonal,

Moreover their direct sum,

$$\hat{X}_k \oplus Z_{k+1} = \hat{X}_{k+1}$$

▼

PROOF

The orthogonal proposition is first proven.

Take any $\hat{x}_k \in \hat{X}_k$, $\tilde{z}_{k+1} \in \tilde{Z}_{k+1}$ is now well defined and consider the inner product;

$$\langle \tilde{z}_{k+1}, \hat{x}_k \rangle = \langle B_k \tilde{z}_{k+1,k}, \hat{x}_k \rangle \quad \text{for some } B_k$$

$$= \langle \tilde{z}_{k+1,k}, B_k^T \hat{x}_k \rangle$$

$$= \langle z_{k+1} - \hat{z}_{k+1,k}, B_k^T \hat{x}_k \rangle$$

$$= \langle \psi_{k+1} x_{k+1} + d\, z_{k+1} - \psi_{k+1} x_{k+1,k}, B_k^T \hat{x}_k \rangle$$

$$= \langle \psi_{k+1} \tilde{x}_{k+1,k} + d\, z_{k+1}, B_k^T \hat{x}_k \rangle$$

$$= \langle \psi_{k+1} \tilde{x}_{k+1,k}, B_k^T \hat{x}_k \rangle + \langle dz_{k+1}, B_k^T \hat{x}_k \rangle$$

The first term is clearly zero as the elements of the inner product belong to orthogonal subspaces. The second term is now expanded;

$$\langle dz_{k+1}, B_k^T \hat{x}_k \rangle = \langle dz_{k+1}, B_k^T \sum_{i=1}^{k} A_i z_i \rangle \qquad \text{for some } A_i$$

$$= \langle dz_{k+1}, B_k^T \sum_{i=1}^{k} A_i (\psi_i x_i + dz_i) \rangle$$

Noting that x_i is not a function of dz_{k+1}, and that;

$$\langle dz_k, dz_j \rangle = \delta_{kj} S_k$$

this second term also vanishes.

Therefore

$\hat{X}_k$ and $\tilde{Z}_{k+1}$ are orthogonal.

■ 1.

To prove $\hat{X}_{k+1} = \hat{X}_k \oplus \tilde{Z}_{k+1}$, we first prove:

i) $\hat{X}_{k+1} \subset \hat{X}_k \oplus \tilde{Z}_{k+1}$ and then

ii) $\hat{X}_{k+1} \supset \hat{X}_k \oplus \tilde{Z}_{k+1}$

i) Let $\hat{x}_{k+1}$ be any element of $\hat{X}_{k+1}$ then;

$$\begin{aligned}
\hat{x}_{k+1} &= \sum_{i=1}^{k+1} A_i z_i \qquad \text{for some } A_i \\
&= \sum_{i=1}^{k} A_i z_i + A_{k+1}(\hat{z}_{k+1,k} + \tilde{z}_{k+1,k}) \\
&= \sum_{i=1}^{k} A_i z_i + A_{k+1}\psi_{k+1}\hat{x}_{k+1,k} + A_{k+1}\tilde{z}_{k+1,k} \\
&= \sum_{i=1}^{k} (A_i + A_{k+1}\psi_{k+1}A_i^*) z_i + A_{k+1}\tilde{z}_{k+1\ k}
\end{aligned}$$

$$\in \hat{X}_k \oplus \tilde{Z}_{k+1} \qquad \therefore \hat{X}_{k+1} \subset \hat{X}_k \oplus \tilde{Z}_{k+1}$$

ii) Take any $\hat{x}_k + \tilde{z}_{k+1} \in \hat{X}_k \oplus \tilde{Z}_{k+1}$

$$\begin{aligned}
\hat{x}_k + \tilde{z}_{k+1} &= \sum_{i=1}^{k} A_i z_i + B_{k+1}\tilde{z}_{k+1,k} \qquad \text{for some } A_i, B_{k+1} \\
&= \sum_{i=1}^{k} A_i z_i + B_{k+1} z_{k+1} - B_{k+1}\psi_{k+1}\hat{x}_{k+1,k} \\
&= \sum_{i=1}^{k} A_i z_i + B_{k+1} z_{k+1} - B_{k+1}\psi_{k+1}\sum_{i=1}^{k} A_i^* z_i \\
&= \sum_{i=1}^{k+1} C_i z_i \in \hat{X}_{k+1} \text{ where } C_i = \begin{cases} A_i - B_{k+1}\psi_{k+1}A_i^* & i=1\ldots k \\ B_{k+1} & i=k+1 \end{cases}
\end{aligned}$$

$$\hat{x}_k + \tilde{z}_{k+1} \in \hat{X}_{k+1} \text{ and } \hat{X}_k \oplus \tilde{Z}_{k+1} \subset \hat{X}_{k+1}$$

Thus from i) and ii)

$$\hat{X}_{k+1} = \hat{X}_k \oplus \tilde{Z}_{k+1}$$

■

THEOREM 3.3.4.

The recurrence relationship for obtaining the filtered estimate of x_k given K observations is;

(C) $$\hat{x}_{o\ k+1,k+1} = \hat{x}_{o\ k+1,k} + B_{k+1}\tilde{z}_{o\ k+1,k}$$

where

$$B_{k+1} = C_{o\ k+1,k}\psi^T_{k+1}\left[\psi_{k+1}C_{o\ k+1,k}\psi^T_{k+1} + S_{k+1}\right]^{-1}$$ ▼

PROOF

Let $\hat{x}_k \in \hat{X}_k$ and $\tilde{z}_{k+1} \in \tilde{Z}_{k+1}$, so from THEOREM 3.3.3., $\hat{x}_k + \tilde{z}_{k+1} \in \hat{X}_{k+1}$.

Denote the orthogonal projection of x_{k+1} onto $\hat{X}_k$ by $\hat{x}_{o\ k+1,k}$ and similarly the orthogonal projection of x_{k+1} onto $\tilde{Z}_{k+1}$ by $\tilde{z}_{o\ k+1}$.

By forming the inner product;

$$\langle x_{k+1} - (\hat{x}_{o\ k+1,k} + \tilde{z}_{o\ k+1}),\ \hat{x}_k + \tilde{z}_{k+1}\rangle$$

and noting that

$$= \langle x_{k+1} - \hat{x}_{o\ k+1,k}, \hat{x}_k\rangle - \langle \tilde{z}_{o\ k+1}, \hat{x}_k\rangle$$

$$+ \langle x_{k+1} - \tilde{z}_{o\ k+1}, \tilde{z}_{k+1}\rangle - \langle \hat{x}_{o\ k+1,k}, \tilde{z}_{k+1}\rangle$$

which is clearly zero,

$$\hat{x}_{o\ k+1,k+1} = \hat{x}_{o\ k+1,k} + \tilde{z}_{o\ k+1}$$

Now

$$\tilde{z}_{o\ k+1} \in \tilde{Z}_{k+1} \text{ and therefore } \tilde{z}_{o\ k+1} = B_{k+1}\tilde{z}_{o\ k+1,k} \text{ for some } B_{k+1}$$

This may be restated as;

$$\hat{x}_{o\ k+1,k+1} = \hat{x}_{o\ k+1,k} + B_{k+1}\tilde{z}_{o\ k+1,k}$$ ■ 1.

and it only remains to find B_{k+1}.

B_{k+1} is such that;

$$\langle x_{k+1} - B_{k+1}\tilde{z}_{o\ k+1,k}, \tilde{z}_{k+1}\rangle = 0 \quad \forall\ \tilde{z}_{k+1} \in \tilde{Z}_{k+1}$$

Now let

$$\tilde{z}_{k+1} = B^{*}_{k+1}\tilde{z}_{o\ k+1,k} \qquad \text{for some } B^{*}_{k+1}$$

Thus

$$\langle x_{k+1} - B_{k+1}\tilde{z}_{o\ k+1,k}, B^{*}_{k+1}\tilde{z}_{o\ k+1,k}\rangle = 0$$

and this must be true for all B^{*}_{k+1}.

Hence

(9) $$\rangle x_{k+1}, \tilde{z}_{o\ k+1,k}\langle \; - B_{k+1}\rangle \tilde{z}_{o\ k+1,k}, \tilde{z}_{o\ k+1,k}\langle \; = 0$$

Considering the first term only:

$$\begin{aligned}
\rangle x_{k+1}, \tilde{z}_{o\ k+1,k}\langle &= \rangle \tilde{x}_{o\ k+1,k} + \hat{x}_{o\ k+1,k}, \tilde{z}_{o\ k+1,k}\langle \\
&= \rangle \tilde{x}_{o\ k+1,k}, \tilde{z}_{o\ k+1,k}\langle + \underset{\sim}{0} \\
&= \rangle \tilde{x}_{o\ k+1,k}, \psi_{k+1}\tilde{x}_{o\ k+1,k} + dz_{k+1}\langle \\
&= \rangle \tilde{x}_{o\ k+1,k}, \tilde{x}_{o\ k+1,k}\langle \psi^{T}_{k+1} + \underset{\sim}{0} \\
&= C_{o\ k+1,k}\psi^{T}_{k+1}
\end{aligned}$$

The second term is;

$$\begin{aligned}
B_{k+1}\rangle \tilde{z}_{o\ k+1,k}, \tilde{z}_{o\ k+1,k}\langle &= B_{k+1}\rangle \psi_{k+1}\tilde{x}_{o\ k+1,k} + dz_{k+1}, \psi_{k+1}\tilde{x}_{o\ k+1,k} + dz_{k} \\
&= B_{k+1}\,(\psi_{k+1}C_{o\ k+1,k}\psi^{T}_{k+1} + S_{k+1} + \underset{\sim}{0} + \underset{\sim}{0}) \\
&= B_{k+1}(\psi_{k+1}C_{o\ k+1,k}\psi^{T}_{k+1} + S_{k+1})
\end{aligned}$$

Substituting back into (9);

$$C_{o\ k+1,k}\psi^{T}_{k+1} - B_{k+1}(\psi_{k+1}C_{o\ k+1,k}\psi^{T}_{k+1} + S_{k+1})^{-1}$$

Hence

$$B_{k+1} = C_{o\ k+1,k}\psi^{T}_{k+1}(\psi_{k+1}C_{o\ k+1,k}\psi^{T}_{k+1} + S_{k+1})^{-1} \qquad \blacksquare$$

We now need only to determine the error covariance matrix of the filtered estimate.

<u>THEOREM 3.3.5.</u>

The outer product of $\tilde{x}_{o\ k+1,k+1}$ is

(D) $$C_{o\ k+1,k+1} = (I_n - B_{k+1}\psi_{k+1})C_{o\ k+1,k} \qquad \blacktriangledown$$

PROOF

From THEOREM 3.3.4.

$$\hat{x}_{o\ k+1,k+1} = \hat{x}_{o\ k+1,k} + B_{k+1}\tilde{z}_{o\ k+1,k}$$

$$= \hat{x}_{o\ k+1,k} + B_{k+1}(z_{k+1} - \hat{z}_{o\ k+1,k})$$

$$= \hat{x}_{o\ k+1,k} + B_{k+1}(\psi_{k+1}\tilde{x}_{o\ k+1,k} + dz_{k+1})$$

$$= x_{k+1} - (I_n - B_{k+1}\psi_{k+1})\tilde{x}_{o\ k+1,k} + B_{k+1}\ dz_{k+1}$$

This can now be substituted into

$$C_{o\ k+1,k+1} = > \tilde{x}_{o\ k+1,k+1}, \tilde{x}_{o\ k+1,k+1} <$$

$$= > (I_n - B_{k+1}\psi_{k+1})(\tilde{x}_{o\ k+1,k}) - B_{k+1}\ dz_{k+1}, \tilde{x}_{o\ k+1,k+1} <$$

$$= (I_n - B_{k+1}\psi_{k+1}) > \tilde{x}_{o\ k+1,k}, \tilde{x}_{o\ k+1,k+1} < - B_{k+1} > dz_{k+1}, \tilde{x}_{o\ k+1,k+1} <$$

$$= (I_n - B_{k+1}\psi_{k+1})C_{o\ k+1,k}(I_n - B_{k+1}\psi_{k+1})^T + B_{k+1}S_{k+1}B^T_{k+1}$$

$$= C_{o\ k+1,k} - C_{o\ k+1,k}(B_{k+1}\psi_{k+1})^T - B_{k+1}\psi_{k+1}C_{o\ k+1,k} + B_{k+1}\psi_{k+1}C_{o\ k+1,k}(B_{k+1}\psi_{k+1})^T + B_{k+1}S_{k+1}B^T_{k+1}$$

$$= C_{o\ k+1,k} - B_{k+1}\psi_{k+1}C_{o\ k+1,k} + (B_{k+1}S_{k+1} - C_{o\ k+1,k}\psi^T_{k+1} + B_{k+1}\psi_{k+1}C_{o\ k+1,k}\psi^T_{k+1})B^T_{k+1}$$

$$= C_{o\ k+1,k} + B_{k+1}\psi_{k+1}C_{o\ k+1,k} + (B_{k+1}(S_{k+1} + \psi_{k+1}C_{o\ k+1,k}\psi^T_{k+1}) - C_{o\ k+1,k}\psi^T_{k+1})B^T_{k+1}$$

Noting the expression for B_k in THEOREM 3.3.4., the last term above vanishes, so that;

$$C_{o\ k+1,k+1} = (I_n - B_{k+1}\psi_{k+1})C_{o\ k+1,k}$$ ■

The whole of this chapter can now be summarised in the following theorem.

THEOREM 3.3.6.

The solution of the General Problem (3.2.1.)may be obtained from the one step recurrence relationships;

(A) $\hat{x}_{o\ k+1,k} = \phi_{k+1,k}\hat{x}_{o\ k,k}$

(B) $C_{o\ k+1,k} = \phi_{k+1,k}C_{o\ k,k}\phi^{T}_{k+1,k} + R_k$

(C) $\hat{x}_{o\ k+1,k+1} = \hat{x}_{o\ k+1,k} + B_{k+1}(z_{k+1} - \psi_{k+1}\hat{x}_{o\ k+1,k})$

(D) $C_{o\ k+1,k+1} = (I_n - B_{k+1}\psi_{k+1})\ C_{o\ k+1,k}$

where

(E) $B_{k+1} = C_{o\ k+1,k}\psi^{T}_{k+1}\ (\psi_{k+1}C_{o\ k+1,k}\psi^{T}_{k+1} + S_{k+1})^{-1}$ ▼

PROOF

Letting j = 1, in THEOREMS 3.3.1. and 3.3.2., expanding the last term in THEOREM 3.3.4. and noting the result in THEOREM 3.3.5., the above equations are obtained. ■

The only data required to start the recurrence process are initial guesses of $\hat{x}_{o,o}$ and $C_{o,o}$, the known transformation matrices $\{\phi_{k+1,k}\}$ and $\{\psi_{k+1}\}$ and the error covariance matrices $\{R_k\}$ and $\{S_k\}$. It is possible to estimate all of the last four sets $\{\ \}$ 'on line'. The method of doing this is explained in the context of the 'Labour Turnover' problem in Chapter 4.

The algorithm for operating the system given $\hat{x}_{o\ k,k}$ and $C_{o\ k,k}$ is;

i) Use (A) to obtain $\hat{x}_{o\ k+1,k}$

ii) Substitute $C_{o\ k,k}$ into (B) to obtain $C_{o\ k+1,k}$

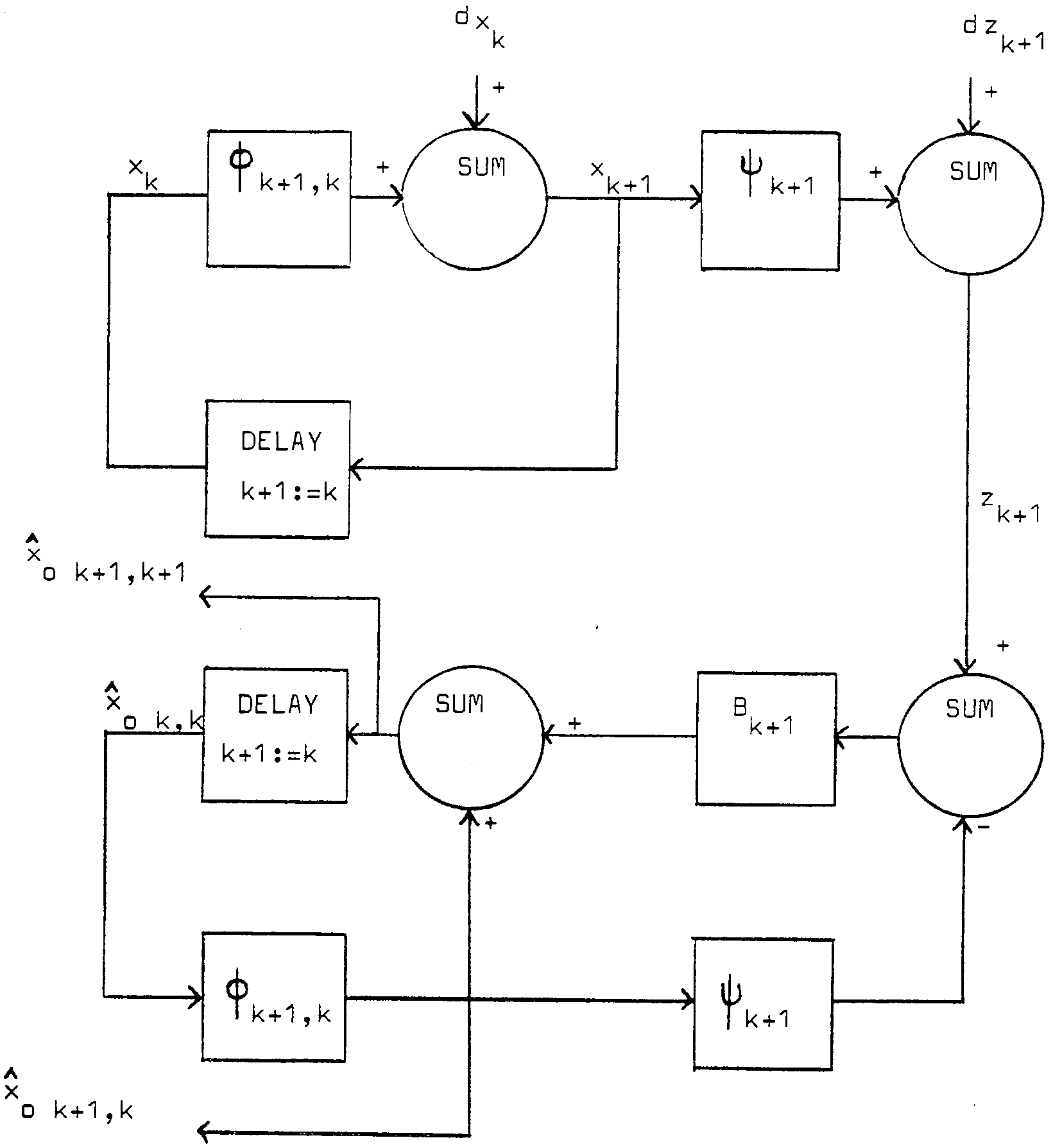
iii) B_{k+1} can now be obtained from (E)

iv) Note the new information z_{k+1}

v) The filtered estimate of x_{k+1} is now obtained using (C) and

vi) its variance using (D)

The systems model of the solution is given below:-



CHAPTER FOUR

THE PREDICTION OF LABOUR TURNOVER

4.1. INTRODUCTION

At the start of Chapter 3 a model of the wastage process was proposed. It was shown that this was just a special case of a wider class of models; for convenience these were known collectively as the General Problem. The General Problem was found to carry its own natural mathematical structure and within this structure the solution was accomplished (THEOREM 3.3.6.). By applying the constraints of the proposed wastage model to the solution of the General Problem it is then easy to solve the wastage problem. This optimal solution for known, time-dependent, system and observation, covariance matrices is given in Part 4.1.1..

It is recognised that these latter matrices may not be known for all time. This, however, is not a severe setback for, as shown in the last part of this section, all that is required for optimal asymptotic estimation is knowledge of RS^{-1}. Further, in the sequel three methods of estimating the covariance matrices $\{R_k\}$ and $\{S_k\}$ are proposed. The first method is based on Sampling Theory and the other two on Bayesian Statistical Theory.

4.1.1. SOLUTION OF THE WASTAGE PROBLEM

The model chosen to represent the wastage process was detailed in Chapter 3 Section 1.1.. Here the actual probability of an employee who had seniority i, breaking tenure with the company in the next unit of time, at time k+1, ($p_{i\ i+1,k+1}$), was related to that probability at time k by equality except for some small additive random disturbance $dp_{i\ i+1,k}$. If there are n intervals (i, i+1), this may be written in matrix notation as;

$$p_{k+1} = p_k + dp_k$$

the above being n-vectors.

It was also assumed that the actual wastage probability was not directly observable - but that the number of employees who leave is directly observable. Therefore, an observation process was proposed.

If O_{k+1} is the n-vector whose elements are the proportion of employees who left from service category i in the next time unit, at time k+1,

$$O_{k+1} = p_{k+1} + dO_{k+1}$$

where dO_{k+1} is the total (sampling) error in the observation system at time k+1.

By making the following transformation to the General Problem:

$$\phi_{k+1} = I \ , \quad \psi_{k+1} = I \ , \quad \hat{x}_{o\ k,j} = \hat{p}_{o\ k,j} \quad \text{and } z_k = O_k$$

the wastage problem is obtained.

Noting THEOREM 3.3.6. the corresponding solution to the wastage problem is:

1) $\hat{p}_{o\ k+1,k} = \hat{p}_{o\ k,k}$

2) $\hat{p}_{o\ k+1,k+1} = \hat{p}_{o\ k+1,k} + B_{k+1}\ (O_{k+1} - \hat{p}_{o\ k+1,k})$

3) $C_{o\ k+1,k} = C_{o\ k,k} + R_k$

4) $C_{o\ k+1,k+1} = (I - B_{k+1})\ C_{o\ k+1,k}$

where;

5) $B_{k+1} = C_{o\ k+1,k}\ (C_{o\ k+1,k} + S_{k+1})^{-1}$

Utilising any 'a priori' information to construct estimates $\hat{p}_{o\ o,o}$ and $C_{o\ o,o}$ the filtering and forecasting models are activated by the first observation. So, given the covariance matrices $\{R_k\}$ and $\{S_{k+1}\}$ the learning process begins. If they are not known they may be estimated. The treatment of this is deferred until Section 4.2. however, the variance structure is examined in the concluding part of this sub-section.

4.1.2. COVARIANCE STRUCTURE OF THE MODEL

A covariance matrix has a special structure called (semi) positive definiteness which will be formally defined below. An important aspect of any model must be its ability to mirror reality. Therefore, if a true covariance matrix is input at time zero ($C_{o\ o,o}$) it is valuable that any proposed model retains this special structure. After proving two alternative representations of Equation 4 Section 4.1.1. it is shown that the properties of covariance matrices are indeed preserved by the wastage equations.

The structure of the matrix B_k, which may be thought of as a weighting factor of the forecast and the new observation, is then examined. It is found that the determinant of this matrix always lies in the interval (0,1) and so has a sensible influence in the construction of the new filtered estimate. Lastly, an equation in B_k, B_{k+1}, S and R is given, from which, if B_k has a limiting form its value may be obtained.

THEOREM 4.1.1.

The covariance error matrix of the filtered estimate may be written as ;

1) $$C_{ok+1,k+1} = B_{k+1}S_{k+1}$$

2) $$C^{-1}_{o\ k+1,k+1} = C^{-1}_{o\ k+1,k} + S^{-1}_{k}$$ ▼

PROOF

1) $$B_{k+1} = C_{o\ k+1,k}\,(C_{o\ k+1,k} + S_{k+1})^{-1}$$

Post multiplying each side by $C_{o\ k+1,k} + S_{k+1}$ yields;

$$B_{k+1}(C_{o\ k+1,k} + S_{k+1}) = C_{o\ k+1,k}$$

and upon re-arranging the terms;

$$(I - B_{k+1})C_{o\ k+1,k} = B_{k+1}S_{k+1}$$

Noticing Equation 4;

$$C_{o\ k+1,k+1} = B_{k+1}S_{k+1}$$ ■1)

2) $$C_{o\ k+1,k+1} = (I - B_{k+1})C_{o\ k+1,k}$$

Post multiplying by $C^{-1}_{o\ k+1,k}$ gives;

$$C_{o\ k+1,k+1}C^{-1}_{o\ k+1,k} = I - B_{k+1}$$

Therefore

$$C_{o\ k+1,k+1}C^{-1}_{o\ k+1,k} + B_{k+1} = I$$

From the first part of the Theorem;

$$B_{k+1} = C_{o\ k+1,k+1}S^{-1}_{k+1}$$

Substituting this into the preceeding equation yields;

$$C_{o\ k+1,k+1}C^{-1}_{o\ k+1,k} + C_{o\ k+1,k+1}S^{-1}_{k+1} = I$$

and premultiplying each side by $C^{-1}_{o\ k+1,k+1}$ proves the second part of the Theorem;

$$C^{-1}_{o\ k+1,k+1} = C^{-1}_{o\ k+1,k} + S^{-1}_{k+1}$$ ■2)

DEFINITION: PRINCIPAL MINOR

A principle minor of a matrix A is a submatrix obtained by deleting certain rows and the same numbered columns of A. Thus, diagonal elements of a principal minor are diagonal elements of A. For notational ease A_{ijk} is chosen to represent the principal minor of A obtained by eliminating the ith, jth and kth rows and columns.

DEFINITION: POSITIVE DEFINITE

A matrix A is positive definite if all its principal minors are positive.

It is clear that covariance matrices are therefore positive definite since their principal minors are positive. Before proving that the wastage equations preserve the structure of the covariance matrices, two preliminary lemmas are given.

LEMMA 4.1.1.

If A and B are symmetric positive definite, then so is A + B.

PROOF

Since B is symmetric there exists an orthogonal matrix P such that $P^{-1}BP$ is diagonal. In fact, this diagonal matrix is the matrix of the characteristic roots of B and is therefore positive definite. As P is orthogonal $|P| = \pm 1$. This implies that;

$$|P^{-1}(A + B)P| = |P^{-1}|\,|A + B|\,|P| = |A + B|$$

Letting

$$Q = P^{-1}AP \qquad \text{and} \qquad \Lambda = P^{-1}BP$$

then

$$P^{-1}(A + B)P = Q + \Lambda$$

Now, a well known theory of Matrix Algebra (Ayres(1962)) states that:

A real symmetric matrix A is positive definite if, and only if, there exists a non singular matrix C; such that $A = C^TC$. So letting:

$$A = C^TC\ ,$$

$$Q = P^{-1}AP = P^{-1}C^TCP = P^TC^TCP = (CP)^TCP$$

Since P is orthogonal then Q is positive definite.

Now consider $|Q + \Lambda|$:-

$$\begin{vmatrix} q_{11}+\lambda_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22}+\lambda_{22} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1} & \cdots & \cdots & q_{nn}+\lambda_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & & & \\ \cdots & & (Q+\Lambda)_1 & \\ q_{n1} & & & \end{vmatrix} + \begin{vmatrix} \lambda_{11} & 0 & \cdots & 0 \\ q_{21} & & & \\ \cdots & & (Q+\Lambda)_1 & \\ q_{n1} & & & \end{vmatrix}$$

$$= \begin{vmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & & \\ q_{n1} & \cdots & (Q+\Lambda)_{21} & \end{vmatrix} + \begin{vmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ 0 & \lambda_{22} & \cdots & 0 \\ \cdots & & & \\ q_{n1} & \cdots & (Q+\Lambda)_{21} & \end{vmatrix}$$

$$+ \begin{vmatrix} \lambda_{11} & 0 & \cdots & 0 \\ 0 & 22 & \cdots & 0 \\ \cdots & & & \\ q_{n1} & \cdots & (Q+\Lambda)_{21} & \end{vmatrix} + \begin{vmatrix} \lambda_{11} & 0 & \cdots & 0 \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & & \\ q_{n1} & \cdots & (Q+\Lambda)_{21} & \end{vmatrix}$$

Continuing this process to completion, it is clear that $|Q + \Lambda|$ is a function in the λ_{ii} and the principal minors of Q only. Therefore $|Q + \Lambda|$ is positive. Further, as a consequence of the above decomposition, the principal minors of $Q + \Lambda$ are readily seen to be positive and so $Q + \Lambda$ is positive definite.

Now

$$|P^{-1}(Q + \Lambda)P| = |A + B| > 0$$

and by similar decomposition the principal minors can be shown to be positive. This implies that A + B is positive definite. ■

COROLLARY 4.1.1.

It is an immediate consequence of the above decomposition that;

$$|Q + \Lambda| > |Q| + |\Lambda|$$

Hence, for positive symmetric matrices A,B

$$|A + B| > |A| + |B|$$

LEMMA 4.1.2.

If A is symmetric positive definite then so is A^{-1}. ▼

PROOF

Since A is symmetric positive definite then, $A = C^T C$ for some non singular matrix C.

Now

$$A^{-1} = (C^T C)^{-1} = C^{-1}(C^T)^{-1}$$

and letting

$$B = (C^T)^{-1},$$

B is clearly non singular and therefore $B^T B$ is positive definite.

But

$$A^{-1} = B^T B$$

and therefore A^{-1} is positive definite. ■

It is now a simple matter to show that the wastage equations retain the covariance structure of symmetric positive definiteness.

THEOREM 4.1.2.

If $C_{o\ o,o}$, $\{R_k\}$ and $\{S_k\}$ are symmetric positive definite then $C_{o\ k,k-1}$ and $C_{o\ k-1,k-1}$ are symmetric positive definite for $k \geqslant 1$. ▼

PROOF

The proof will be by induction; let k = 1. $C_{o\ k-1,k-1} = C_{o\ o,o}$ and therefore $C_{o\ k-1,k-1}$ is symmetric positive definite.

Now

$$C_{o\ 1,o} = C_{o\ o,o} + R_o$$

which is the sum of two symmetric positive definite matrices and consequently symmetric positive definite. The Theorem is true then for k = 1, so assume that it is true for k = n, that is; $C_{o\ n,n-1}$ and $C_{o\ n-1,n-1}$ are symmetric positive definite.

Noting Theorem 4.1.1.;

$$C_{o\ n,n} = (C^{-1}_{o\ n,n-1} + S^{-1}_n)^{-1}$$

Envoking the Lemmas (4.1.1. and 4.1.2.) it is immediate that $C_{o\ n,n}$ is symmetric positive definite. Also;

$$C_{o\ n+1,n} = C_{o\ n,n} + R_n$$

and is clearly symmetric positive definite. The Theorem is, therefore, true for k = n + 1. By the principle of mathematical induction, the Theorem holds for all $k \geq 1$. ∎

Two theories of Matrix Algebra (Hoffman & Kunze(1961)) are;

i) If A is positive definite then it is self adjoint.

ii) If A and B are self adjoint, then AB is self adjoint AB = BA.

A necessary condition for the product of A and B, both positive definite, to be positive definite is that A and B must commute.

In Theorem 4.1.1. it was proven that;

$$B_{k+1} = C_{o\ k+1,k+1} S^{-1}_{k+1},$$

and in general there is no reason to believe that $C_{o\ k+1,k+1}$ and S^{-1}_{k+1} commute. Therefore, B_{k+1} will not normally be symmetric positive definite. As mentioned earlier B_{k+1} may be thought of as a 'weighting' matrix - it is in this context that its properties are examined below.

THEOREM 4.1.3.

$$0 < |B_k| < 1$$

PROOF

From Theorem .1.1. B_k may be written as;

$$B_k = C_{o\ k,k} S_k^{-1}$$

Therefore

$$|B_k| = |C_{o\ k,k}||S_k|^{-1} > 0$$

In the wastage equations (.1.1.) B_k is given in the form;

$$B_k = C_{o\ k,k-1}(C_{o\ k,k-1} + S_{k+1})^{-1}$$

Taking determinants, gives;

$$|B_k| = |C_{o\ k,k-1}||C_{o\ k,k-1} + S_{k+1}|^{-1}$$

Now, Corollary .1.1. shows;

$$|C_{o\ k,k-1}| + |S_{k+1}| < |C_{o\ k,k-1} + S_{k+1}|$$

and since

$$|S_{k+1}| > 0,$$

$$|C_{o\ k,k-1}| < |C_{o\ k,k-1} + S_{k+1}|$$

Substituting into the last expression for B_k yields the desired result; that is;

$$|B_k| < |C_{o\ k,k-1} + S_{k+1}||C_{o\ k,k-1} + S_{k+1}|^{-1} = 1$$ ■

Three other properties are noteworthy:

(i) $|1 - B_k| > 0$ $\forall k$ as $|1-B_k| = |C_{o\ k,k}||C_{o\ k,k-1}|^{-1} > 0$

(ii) $|C_{o\ k,k}| < |S_k|$ $\forall k$ as $|C_{o\ k,k}| = |B_k||S_k| < |S_k|$

(iii) $|C_{o\ k+1,k}| > |R_k|$ $\forall k$ as $|C_{o\ k+1,k}| = |C_{o\ k,k} + R_k| > |R_k|$

THEOREM 4.1.4.

If $R_k = R$ and $S_k = S$ for all k, then;

$$|B_{k-1} - B_k| \gtrless 0 \Rightarrow |B_k - B_{k+1}| \gtrless 0$$

PROOF

$$B_{k+1} = C_{o\ k+1,k}(C_{o\ k+1,k} + S)^{-1}$$

Expanding $C_{o\ k+1,k}$ and noting Theorem 4.1.1. gives;

$$B_{k+1} = (B_kS + R)(B_kS + R + S)^{-1}$$

Therefore

$$B_k(B_{k-1}S + R + S) - B_{k+1}(B_kS + R + S) = (B_{k-1}S + R) - (B_kS + R)$$

Collecting terms;

$$(B_kB_{k-1} - B_{k+1}B_k)S + (B_k - B_{k+1})(R + S) = (B_{k-1} - B_k)S$$

Adding $(B_kB_k - B_kB_k)S$ to the left hand side of the equation and simplifying yields;

$$B_k(B_{k-1} - B_k)S + (B_k - B_{k+1})(C_{o\ k,k} + R + S) = (B_{k-1} - B_k)S$$

So

$$(B_k - B_{k+1})(C_{o\ k,k} + R + S) = (1 - B_k)(B_{k-1} - B_k)S$$

Taking determinants proves the Theorem. ∎

COROLLARY 4.1.2.

$$|B_{k-1} - B_k| \gtrless 0 \Rightarrow \begin{cases} |C_{o\ k,k} - C_{o\ k+1,k+1}| \gtrless 0 \\ |C_{o\ k+1,k} - C_{o\ k+2,k+1}| \gtrless 0 \end{cases}$$ ∎

Another interesting fact is revealed by writing;

$$|B_k| = |B_{k-1}||B_{k-1}|^{-1}|B_k| = |B_{k+1}||C_{o\ k+1,k+1}|^{-1}|C_{o\ k,k}|$$

Proving

$$|B_k| \gtrless |B_{k+1}| \Rightarrow |C_{o\ k,k}| \gtrless |C_{o\ k+1,k+1}|$$

If there is no ' a priori ' information about the system then it is sensible to take the first observation vector as a starting point. This amounts to setting B_1 the weighting factor equal to the n-identity matrix. The recursive equations are now active;

$$B_1 = 1_n \quad \text{which implies} \quad |B_1| = 1$$

Therefore

$$C_{o\ 11} = B_1 S = S$$

and

$$C_{o\ 21} = S + R$$

The second weighting factor is now computed as;

$$B_2 = S + R(S + 2R)^{-1}$$

So clearly

$$|B_1| > |B_2|$$

But, further

$$|B_1 - B_2| = |1 - B_2| > 0$$

which implies $|B_k - B_{k+1}| > 0$ for all k.

THEOREM 4.1.5.

If R,S and $C_{o\ o,o}$ are diagonal and $B_k \longmapsto B$, then the limiting matrix B is diagonal and has elements;

$$b_{ii} = \frac{(1 + 4r_{ii}^{-1}s_{ii})^{\frac{1}{2}} - 1}{2r_{ii}^{-1}s_{ii}} \qquad \text{for } i = 1 \ldots n$$

▼

PROOF

Starting as in the previous Theorem with;

$$B_{k+1} = (B_k S + R)(B_k S + R + S)^{-1}$$

and post multiplying by $B_k S + R + S$ we obtain;

$$B_{k+1}(B_k S + R + S) = B_k S + R$$

or

$$B_{k+1}B_k + B_{k+1}(RS^{-1} + 1) - B_k - RS^{-1} = 0$$

It is noted that this is a function of RS^{-1} only; the importance of this is explained immediately after this Theorem.

So if $B_k \longmapsto B$ then;

$$B^2 + BRS^{-1} - RS^{-1} = 0$$

Since B_k is the product of diagonal matrices, B must be diagonal and therefore its elements are given by the solution of;

$$b_{ii}^2 + b_{ii}r_{ii}s_{ii}^{-1} - r_{ii}s_{ii}^{-1} = 0$$

and so

$$b_{ii} = \frac{-r_{ii}s_{ii}^{-1} \pm (r_{ii}^2 s_{ii}^{-2} + 4r_{ii}s_{ii}^{-1})^{\frac{1}{2}}}{2}$$

Algebraic manipulation proves the Theorem. ■

If the weighting matrix converges to a limit B, it is clear that the wastage model tends to a model that might be described as the matrix equivalent of the Exponentially Weighting Moving Average method - often employed by forecasters. As noted in the last theorem, the limiting matrix B is a function of RS^{-1} only. The choice of the correct value for B (in a E.W.M.A.) is often a problem for the inexperienced forecaster. The latter property of the B matrix implies that knowledge of the covariance matrices R and S is sufficient to ensure convergence to the optimal B, automatically.

Now, it might be thought that the problem has been complicated; the practitioner being asked to supply two covariance matrices instead of one weighting factor. The problem has not in fact been complicated but simplified. This statement is justified by noting that the structure introduced into the model has a physical meaning; whereas the weighting factor of classical methods is <u>just</u> an abstract matrix.

The structural model is therefore easier to interpret in the real world. Often, even inexperienced forecasters will be able to supply reasonable bounds for the error in any estimate they have put forward but will be unable to justify the guess of a weighting factor B, let alone give usable confidence limits for its value.

The complication; simplification question can, however, be conclusively resolved by noting a statement made in Chapter 3 Section 3:

> "It is possible to estimate the covariance matrices R_k and S_k on line"

So values of the covariance matrices are not asked for which must hold for all time but only initial estimates (R_0 and S_0) of them are required. These can then be automatically updated within the model so that they converge to the true values R_k and S_k. In the next section three methods of updating these covariance matrices are proposed.

4.2 COVARIANCE ESTIMATION

In this sub-section three methods of estimating the covariance matrices R_k and S_k are considered. The first method examined is that of unbiased estimation. Unbiasedness can be achieved by the evaluation of the expectation of certain differences in the temporal observations. The latter two methods employ techniques from Bayesian Statistical Theory. Here the procedure is to set up a class of alternative values for each R_k and S_k and to assign a prior probability to them. These prior probabilities are then modified in view of the observations, to give posterior probabilities of each element of the class attaining. The second method chooses the model with the highest posterior probability of attaining at time k as the model to be employed for the next time period. The third method uses a reduction process to obtain a model whose parameters are a weighted estimate according to the posterior probabilities of the elements obtained in the second method.

4.2.1. UNBIASED ESTIMATION OF R_k AND S_k.

After defining unbiasedness, unbiased estimators of $\{R_k\}$ and $\{S_k\}$ are presented. If $\{R_k\}$ can be considered time-invariant, the problem of its estimation is shown to be yet another special case of the General Problem. In this instance the sample mean of classical statistics is seen to be equivalent to the optimal estimator given by the particular solution of the General Problem. The variance of this estimator is then discussed qualitatively. It is seen that although the estimator is unbiased and the covariance matrix always (semi) positive definite, the estimator can in fact be negative in the early time periods.

DEFINITION: UNBIASED

An estimator $\hat{x}_{k,j} = f(z_1,\ldots,z_j)$ of x_k is defined as unbiased if, and only if;

$$E(\hat{x}_{k,j}) = x_k$$

The jth difference in the observations at time k will be denoted ${}^{j}Dz_k$. With this notation the expectation of ${}^{j}Dz_k$ can be seen below to be zero for all k and j, such that $k \geqslant j + 1$,

$$
\begin{aligned}
{}^{j}Dz_k &= z_k - z_{k-j} \\
&= x_k + dz_k - x_{k-j} - dz_{k-j} \\
&= \sum_{i=1}^{j} dx_{k-i} + dz_k - dz_{k-j}
\end{aligned}
$$

Since the expectation of all the terms on the right hand side of the last equation is zero it is clear that;

$$E({}^{j}Dz_k) = 0 \qquad \forall j,k \quad \text{such that } k > j > 0$$

Its covariance can now be considered, for $k > j > 0$;

$$
\begin{aligned}
&\rangle\, {}^{j}Dz_k,\ {}^{j}Dz_k \,\langle \\
&= \rangle \sum_{i=1}^{j} dx_{k-i} + dz_k - dz_{k-j}\ , \sum_{i=1}^{j} dx_{k-i} + dz_k - dz_{k-j} \langle \\
&= \rangle \sum_{i=1}^{j} dx_{k-i}, \sum_{i=1}^{j} dx_{k-i} \langle \ + \rangle\, dz_k,\ dz_k \langle \ + \ \rangle dz_{k-j},\ dz_{k-j} \langle \\
&\quad + (4j + 2 \text{ terms whose value is zero}) \\
&= \sum_{i=1}^{j} R_{k-i} + S_k + S_{k-j} + (j^2 + 3j + 2 \text{ terms whose expectation is zero})
\end{aligned}
$$

Similarly the expectation for two adjacent jth differences can be computed for all $k \geqslant j + 2$.

$$
\begin{aligned}
&\rangle\, {}^{j}Dz_k,\ {}^{j}Dz_{k-1} \langle \\
&= \rangle \sum_{i=1}^{j} dx_{k-i} + dz_k - dz_{k-j},\ \sum_{l=1}^{j} dx_{k-1-l} + dz_{k-1} - dz_{k-j-1} \langle \\
&= \rangle \sum_{i=1}^{j} dx_{k-i},\ \sum_{l=1}^{j} dx_{k-1-l} \langle - \rangle\, dz_{k-j}, dz_{k-1} \langle \\
&\quad + (4j + 3 \text{ terms} = 0) \\
&= \sum_{i=2}^{j} R_{k-i} - S_{k-1}\, \delta_{j1} + (j^2 + 3j + 4 \text{ terms equal to zero})
\end{aligned}
$$

where δ_{j1} is the Kronecker delta.

If j is set equal to 1 in the last two derivations then the following is obtained;

$$\rangle {}^1Dz_k, {}^1Dz_k \langle \; = R_{k-1} + S_k + S_{k-1} \quad \text{(+ 6 terms equal to zero)}$$

and

$$\rangle {}^1Dz_k, {}^1Dz_{k-1} \langle \; = - S_{k-1} \quad \text{(+ 8 terms equal to zero)}$$

Noticing the above it is now easy to see how unbiased estimators may be constructed. This is given in the following Theorem.

THEOREM 4.2.1.

i) $$- \rangle {}^1Dz_k, \; {}^1Dz_{k-1} \langle \; = S_{k-1}$$

ii) $$\rangle {}^1Dz_k, \; {}^3Dz_{k+1} \langle \; = R_{k-1}$$

PROOF

i) This has already been proved above, ■ i

ii) $\rangle {}^1Dz_k, \; {}^3Dz_{k+1} \langle$

$$= \rangle {}^1Dz_k, \; {}^1Dz_{k+1} + {}^1Dz_k + {}^1Dz_{k-1} \langle$$

$$= \rangle {}^1Dz_k, \; {}^1Dz_k \langle + \rangle {}^1Dz_{k+1}, {}^1Dz_k \langle + \rangle {}^1Dz_k, \; {}^1Dz_{k-1} \langle$$

and noting the arguments made above;

$$= R_{k-1} + S_k + S_{k-1} - S_k - S_{k-1}$$

$$= R_{k-1}$$ ■ ii

Unbiased estimators of R_{k-1} and S_{k-1} are then;

$$\hat{R}_{k-1} = (z_k - z_{k-1})(z_{k+1} - z_{k-2})^T$$

$$\hat{S}_{k-1} = (z_k - z_{k-1})(z_{k-2} - z_{k-1})^T$$

If R_k is time invariant the following is easily seen to be a specific form of the General Problem;

$$R_k = R$$

$$\hat{R}_k = R_k + d\hat{R}_k$$

where $d\hat{R}_k$ represents the estimation error.

As $\hat{R}_k$ is an unbiased estimator of R it is clear that $E(d\hat{R}_k) = 0$. Let $\rangle d\hat{R}_k, d\hat{R}_k \langle = T$ and the first observation $\hat{R}_1$ be the starting point. Then the algorithm of Chapter 4 Section 4, with $\overset{*}{R}_k = \hat{x}_{o\ k,k}$, gives the following;

$$B_1 = 1 \qquad B_2 = 2^{-1} \qquad B_3 = 3^{-1}$$

$$\overset{*}{R}_1 = R_1 \qquad \overset{*}{R}_2 = 2^{-1}\overset{*}{R}_1 + 2^{-1}R_2 \qquad \overset{*}{R}_3 = 2 \times 3^{-1}\overset{*}{R}_2 + 3^{-1}R_3$$

$$C_{o\ 1,1} = T \qquad C_{o\ 2,2} = 2^{-1}T \qquad C_{o\ 3,3} = 3^{-1}T$$

and in general;

$$B_k = k^{-1}, \quad C_{o\ k,k} = k^{-1}T \quad \text{and} \quad \overset{*}{R}_k = \sum_{i=1}^{k} R_i k^{-1} = \bar{\hat{R}}_k$$

Thus the mean value of the $\{\hat{R}_k\}$ is the optimal estimate of R. Further, $B_k = k^{-1}$ implies $B \mapsto 0$ as $k \mapsto \infty$. Since $C_{o\ k,k} = B_k T$, it also vanishes with increasing observations - this shows that not only is $\overset{*}{R}_k$ an unbiased optimal estimate for R_k but that its variance tends to zero in the limit. The same argument can clearly be repeated for an unbiased estimator S_k of a system with time invariant S_k. Results of these unbiased estimators are presented in Appendix 3

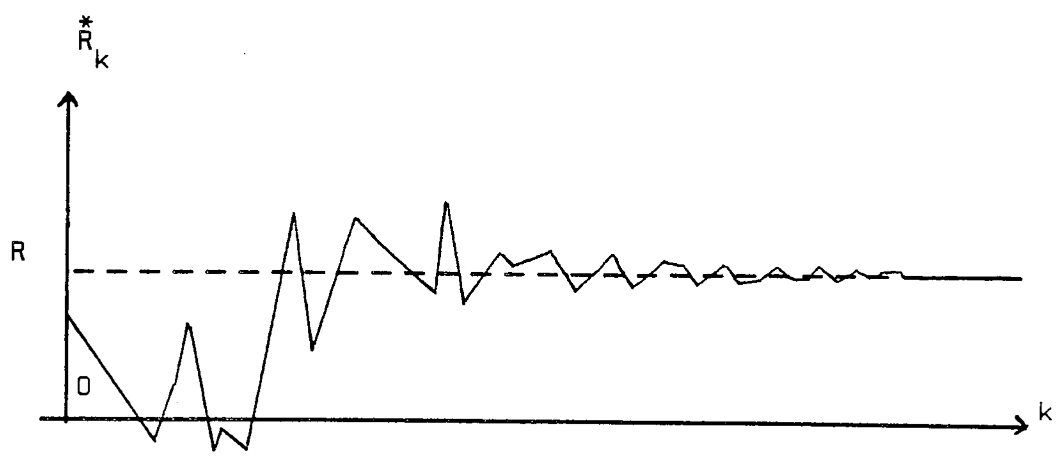
In practice the number of observations required to start the estimation process depends on the dimension of the system. Inspection of the 'raw' estimators, $\hat{R}_{k-1}$ and $\hat{S}_{k-1}$, reveals four observations are needed for the former and three for the latter. This number of observations will give a sample of one (dimension n) for each estimator. The dimension of a n x n covariance matrix is, however, $\sum_{i=1}^{n} i = \frac{1}{2}n(n + 1)$. This means that degeneration, (singularity in the covariance matrix) will occur if $n \neq 1$. Therefore, to avoid singularity the number of 'samples' needed (N) is given by the solution of the following equation;

$$Nn \geqslant \tfrac{1}{2}n(n + 1)$$

If n is odd this amounts to setting $N = \frac{1}{2}(n + 1)$ and if n is even $N = \frac{1}{2}(n + 2)$. Coupled with the number of observations needed to give the first sample it is seen that $N = \frac{1}{2}(n + 7)$ is sufficient to expect that the estimated covariance matrices would be non singular.

Earlier in this subsection whilst deriving the covariance of certain difference combinations it was noted that there were terms whose expectation were zero. The covariance of these terms, however, do not vanish. For this reason the unbiased estimators have significant variance. This, of course, is true for most covariance estimators. So, although the estimators are unbiased and covariance matrices cannot have negative diagonal elements - when the true covariance matrix has diagonal elements close to zero, negative estimates can arise (See Figure 4.2.1.). This amounts to loss of positive definiteness of the covariance estimators in the model.

FIGURE 4.2.1.



A typical graph illustrating the convergence of $\overset{*}{R}_k$ to R for the univariate case and the fact that $\overset{*}{R}_k$ may be negative in the early stages of prediction.

Many practical procedures can be suggested to overcome the loss of positive definiteness in the estimated covariance matrix. The solution given below is neither the only nor possibly the optimal one. It has, however, worked well in practice when absolutely no knowledge is available to the magnitude of the estimated matrix. Of course, if knowledge exists to the nature of this matrix, it could be used in the early stages of prediction, until positive definiteness is achieved.

The practical procedure is to set negative diagonal elements equal to their positive equivalents of the same value. This may be justified by observing that these elements represent variances and are therefore positive, so that the new positive-valued estimate must in fact be closer to the true variance. The matrix can now be made symmetric by pooling the off diagonal estimates, the new a_{ij} be equal to the mean of the old a_{ij} and a_{ij}. The principal minors are then examined in increasing dimension order. If they are found to be negative then the relevant off diagonal elements are set to zero. The covariance matrix estimate is now positive definite and provides a usable input into the model. It is noted once again that this procedure is only called upon when loss of positive definiteness occurs. In the majority of the systems examined the covariance matrix is significantly different from zero whereupon the normal estimation procedure converges satisfactory to the true value from the outset.

4.2.2. A BAYESIAN APPROACH TO COVARIANCE ESTIMATION

Before embarking on the presentation of the method employed in the model, an insight into Bayesian methodology is given. Bayesian statisticians use probabilities to describe degrees of belief in possible alternative parameter values or states of nature. In advance of experimentation it is assumed that a prior distribution of the parameter that it is desired to estimate can be obtained. This may be achieved as a last resort by introspection to arrive at odds for which one might bet for and against a parameter attaining. This prior distribution is then modified to a posterior distribution according to Bayes theorem. In essence a new observation changes ones belief in the set of admissable parameter values from the prior distribution to the posterior distribution. The exact prior distribution of the parameter is not however, imperative, for, if an adequate prior is given, a reasonably informative experiment will often yield a posterior distribution not much different from the true value. Further, it is not even necessary for the prior distribution to intergrate (sum) to unity over the whole sample space to achieve valid posterior distributions. In this case the prior is said to be improper. In this subsection the likelihood interpretation of Bayes theorem is adopted. A more comprehensive reference to the Bayesian philosophy is Lindley (1965).

DEFINITION: LIKELIHOOD FUNCTION

The likelihood function ($L(\theta/z_1,z_2,\ldots,z_n)$) of n random variables $Z_1,Z_2,\ldots,Z_n$ is their joint density;

$$f_{Z_1,Z_2,\ldots,Z_n}(z_1,z_2,\ldots,z_n/\theta)$$

which is to be regarded as a function of θ.

In particular, if $Z_1,Z_2,\ldots,Z_n$ is a random sample from $f(z/\theta)$

$$L(\theta/z_1,z_2,\ldots,z_n) = \prod_{i=1}^{n} f(z_1/\theta)$$

EXAMPLE 4.2.1.

If the Z_i have a Bernoulli distribution, i.e.;

$$f(z/\theta) = \theta^{z}(1 - \theta)^{1-z} \quad \text{where } 0 \leqslant \theta \leqslant 1$$

The likelihood function is;

$$L(\theta/z_1,z_2,\ldots,z_n) = \theta^{\Sigma z_i}(1-\theta)^{n-\Sigma z_i}$$

If $p'(\theta/z)$ denotes the posterior density of θ and $m(z)$ the marginal density of Z, the joint density of Z and θ is;

$$M(z)p'(\theta/z) = p(\theta)L(z/\theta)$$

so that

$$p'(\theta/z) = (M(z))^{-1}\,p(\theta)\,L(z/\theta)$$

Now if the integral (sum) of the posterior density is to be unity, it is clear that;

$$M(z) = \int_{\Omega} p(\theta)L(z/\theta)d\theta$$

Substituting this into the preceeding equation gives;

$$p'(\theta/z) = \frac{p(\theta)L(z/\theta)}{\int_{\Omega} p(\theta)L(z/\theta)d\theta}$$

The above result is known as Bayes theorem.

It is noted that Bayes estimators are C.A.N.E.(consistant asymptotically normal efficient). Even in those problems when the exact prior distribution is unknown this method can therefore be used. Thus, it provides a technique of determining estimators of functions that might not have otherwise have been considered estimiable.

Since the aim of this subsection is to estimate the covariance matrices R and S an admissable nth order parameter set $\{R^i,S^i\}$ is chosen. The elements of the parameter set is denoted by θ_i where $1 \leqslant i \leqslant n$. Recalling Bayes theorem it is easy to verify that;

$$p'(\theta_i/z_k,\{z_{k-1}\}) = \frac{L(z_k/\theta_i,\{z_{k-1}\})p(\theta_i/\{z_{k-1}\})}{\int_{\theta} L(z_k/\theta_i,\{z_{k-1}\})p(\theta_i/\{z_{k-1}\})d\theta}$$

The denominator of the left hand side of the above equation is a constant for all of the n models.

Therefore

$$p'(\theta_i/\{z_k\}) \propto L(z_k/\theta_i\{z_{k-1}\})p(\theta_i/\{z_{k-1}\}).$$

The notation can be further simplified by allowing $P_{i,k}$ to represent the probability density of model θ_i attaining, given observations up to and including z_k.

In this case;

$$P_{i,k} \propto L(z_k/\theta_i, \{z_{k-1}\})P_{i,k-1}$$

so that

$$P_{i,k} \propto \prod_{m=1}^{k} L(z_m/\theta_i, \{z_{m-1}\})P_{i,o}$$

As stated earlier there are two methods to be examined that relate to this Bayesian approach;

(i) The first method is to choose at time k the model that has highest posterior probability of attaining at time k as the model most likely to hold at time k + 1. That is, the model used to predict $\hat{x}_{k+1,k}$, for any decision making, is the model θ_i^* such that $i = \max. \left\{P_{i,k}\right\}$. Values of the other θ_i are also computed, so that the process may continue. In this case;

$$\hat{x}^*_{k+1,k} = \hat{x}^i_{k,k-1} + B^i_k(z_k - \hat{x}^i_{k,k-1})$$

(ii) The second method is to weight the parameters R^i and S^i according to the magnitude of the posterior probabilities of the models θ_i attaining. In this way;

$$R'_k = \frac{\Sigma P_{i,k}R^i}{\Sigma P_{i,k}} \quad \text{and} \quad S'_k = \frac{\Sigma P_{i,k}S^i}{\Sigma P_{i,k}}$$

and

$$\hat{x}'_{k+1,k} = \hat{x}'_{k,k-1} + B'_k(z_k - \hat{x}'_{k,k-1})$$

The first method is useful when cost is an inportant constraint - to reduce the computational burden the model chosen at time k can be deemed to hold for all succeeding time. The weighting method has intuitive appeal - if the true model θ is not an element of the parameter set θ_i , the estimate θ' can still be unbiased. Further, if θ is not time invariant, the second method will in fact follow the moments of θ_k through time. To employ either method, it is seen that an explicit form for the likelihood function is required for each θ_i. How to obtain these forms is the subject of the following part of this subsection.

4.2.3. SIMPLIFICATION OF THE MODEL BY A 'NORMALITY' ASSUMPTION

Under the hypothesis that model θ_i holds, the governing recurrence equations are;

1) $\hat{x}^i_{k,k-1} = \hat{x}^i_{k-1,k-1}$

2) $\hat{x}^i_{k,k} = \hat{x}^i_{k,k-1} + B^i_k(z_k - \hat{x}^i_{k,k-1})$

3) $C^i_{k,k-1} = C^i_{k-1,k-1} + R^i$

4) $C^i_{k,k} = (I - B^i_k)C^i_{k,k-1}$

where

5) $B^i_k = C^i_{k,k-1}(C^i_{k,k-1} + S^i)^{-1}$

At time k-1 the best estimate of z_k (see Chapter 3 Section 3.3.) under model θ_i will be;

$$\hat{z}^i_{k,k-1} = \hat{x}^i_{k,k-1} = \hat{x}^i_{k-1,k-1}$$

Its error covariance matrix may be calculated as below;

$$\rangle \tilde{z}^i_{k,k-1}, \tilde{z}^i_{k,k-1} \langle \; = \; \rangle \tilde{z}^i_{k,k-1}, z_k - \hat{z}^i_{k,k-1} \langle$$

$$= \rangle \tilde{z}^i_{k,k-1}, \; x_k + dz_k - \hat{x}^i_{k,k-1} \langle$$

$$= \rangle \tilde{x}^i_{k,k-1} + dz_k, \; \tilde{x}^i_{k,k-1} + dz_k \langle$$

$$= \quad C^i_{k,k-1} + S^i$$

$$= \quad C^i_{k-1,k-1} + S^i + R^i$$

Summarising, the best estimate for z_k at time $k - 1$ is $\hat{x}^i_{k-1,k-1}$ with an error covariance matrix $C^i_{k-1,k-1} + S^i + R^i$. The first two moments of the distribution of z_k are now known for each θ_i given $\{z_{k-1}\}$. These moments are enough to define a unique Normal distribution;

$$f(z_k/\theta_i, \{z_{k-1}\}) = N(\hat{x}^i_{k-1,k-1}, C^i_{k-1,k-1} + R^i + S^i)$$

The likelihood function of this distribution is;

$$L(z_k/\theta_i, \{z_{k-1}\})$$

$$= \frac{(\sqrt{2\pi})^{-n}}{|C_{k-1,k-1} + R^i + S^i|^{\frac{1}{2}}} \; e^{-\frac{1}{2}(z_k - \hat{x}^i_{k-1,k-1})^T (C^i_{k-1,k-1} + R^i + S^i)^{-1} (z_k - \hat{x}^i_{k-1,k-1})}$$

Recalling that;

$$P_{i,k} = L(z_k/\theta_i, \{z_{k-1}\}) P_{i,k-1}$$

the $P_{i,k}$ can now be recursively evaluated.

Each of the three methods described in this section for estimating R_k and S_k are shown in an operational mode in the Appendix (C). Comparisons of their efficiency are given, both on simulated data and on 'live' data from Dunlop U. K. Tyre Group.

CHAPTER FIVE

MANPOWER SUPPLY FORECASTING

5.1. INTRODUCTION

In this chapter the flow of employees within a graded organisation are considered. It is usual that an organisation will be hierarchical and that the only transitions will be one and two step promotions; demotions and accelerated promotion being rare. Dunlop U.K.T.D. is an example of such an organisation. However, a manpower sub-system might be such that one has to obtain experience in one or more areas of equal status, in this case transitions would occur between all 'grades'. The model developed in this chapter is equally capable of dealing with both situations.

After applying the length of service dependent wastage rates to the stock matrix at time T and summing over the length of service categories, the resultant vector is the estimated number of graded employees who will remain in the system at least until time T + 1. These employees are admissable for transition between grades. Recruits entering the manpower system in this period are not considered eligible for transition. It is apparent therefore that the manpower system under consideration is closed, as both inflow (recruits) and outflow (leavers) have been accounted for.

In a renewal model, once the wastage rates have been given, the transition matrix is exactly determined although it is not deterministic as its elements include the stochastic wastage rates. The optimal forecasting of these time dependent wastage rates has been dealt with in the previous two chapters. Their application to renewal models is obvious and it is not wished to elaborate further on this here.

The primary concern of this chapter is the estimation of the elements of the stochastic transition matrix and the one-step prediction of the stock flow matrix. The notation to be used throughout this chapter is presented in Section 1.1.. In Section 2 properties of the compound multinomial distribution are given. It should be remembered that, conditional upon the value of the transition matrix, the flows between grades are multinomially distributed. If the elements of transition matrix are considered to be random variables then the flow will be a realisation of a contagious distribution of the multinomial and transition distributions. General results are derived when the transition distribution is assumed to be arbitrary and a method of estimating its second order moments exhibited.

In Section 3, the rows of the transition matrix are assumed to be governed by Dirichlet distributions. The joint raw moments of this distribution are obtained and from these the generalised factorial moments of the stock vector can be calculated. The next problem examined is the estimation of the parameters of this distribution and a method based on Sampling Theory is developed. Concluding this section is an examination of the structure of the Dirichlet distribution. It is found that the distribution is not as general as it may first have been thought to be as it conceals a very specific assumption about flows within the organisation.

In the last section of this chapter, the prediction of manpower supply can be alligned with the General Problem of previous chapters. The particular representation for manpower supply applications is given.

5.1.1. <u>NOTATION</u>

The notation used in this chapter is now introduced. Occasionally it will be found convenient to drop either the T or i suffix, when transitions from the ith row at time T only are considered. It will be made clear in the text when this procedure is employed.

$n_{i,T}$ = number of employees in grade i at time T

$N_{ij,T}$ = random variable denoting the number of transitions between grade i and grade j within the time interval (T,T + 1)

$P_{ij,T}$ = random variable denoting the probability of a transition as above

$Q_{ij,T}$ = random variable denoting the proportion of transitions observed as above

$n_{ij,T}$, $p_{ij,T}$ and $q_{ij,T}$ are realisations of the above random variables

$n^{o}_{ij,T}$ = observed number of $n_{ij,T}$

d'( ) = density function of the Dirichlet distribution

m( ) = density function of the multinomial distribution

d'm( ) = density function of the Dirichlet-Multinomial distribution

$E_X(\ )$ = expectation function over an X-distribution

C( ) = covariance function

R( ) = correlation function

V( ) = variance function

$S_X(\ )$ = a matrix with (i,j) th element $C(X_i X_j)$

$\hat{X}$ = an estimate of X

$X^{(Z)} = X(X-1)\ldots(X-Z+1)$

$O(X^{-1})$ = terms of order at most X^{-1}

Γ = gamma function

a_{ij} = parameters of the Dirichlet distribution

5.2. PROPERTIES OF THE COMPOUNDED MULTINOMIAL DISTRIBUTION

A distribution is said to be multinomial if it has density function;

$$m(x_1,\dots x_{k-1}/n,p_1,\dots p_{k-1}) = \frac{n!}{\prod_{j=1}^{k} x_j!} \prod_{j=1}^{k} p_j^{\,x_j}$$

where for notational convenience the symbols x_k and p_k have been introduced such that;

$$x_k = n - \sum_{j=1}^{k-1} x_j \qquad \text{and} \qquad p_k = 1 - \sum_{j=1}^{k-1} p_j$$

As stated earlier, conditional upon the p_{ij} and the $n_{i,T}$ for all i, the flows $n_{ij,T}$ are multinomially distributed, that is;

$$m(n_{i1,T}\dots n_{ik-1,T}/n_{i,T}p_{i1}\dots p_{ik-1}) = \frac{n_{i,T}!}{\prod_{j=1}^{k} n_{ij,T}!} \prod_{j=1}^{k} p_{ij}^{\,n_{ij,T}} \qquad i=1\dots k$$

(5.2.1.)

The generalised factorial moment of this distribution is;

$$E(N_{i1,T}^{(Z_{i1})}\, N_{i2,T}^{(Z_{i2})} \dots N_{i\ k-1,T}^{(Z_{ik-1})}) \qquad i=1\dots k$$

$$= n_{i,T}^{\left(\sum_{j=1}^{k-1} Z_{ij}\right)} \prod_{j=1}^{k-1} p_{ij}^{\,Z_{ij}}$$

(5.2.2.)

So, clearly the generalised factorial moment of any compounded distribution is;

$$n_{i,T}^{\left(\sum_{j=1}^{k-1} Z_{ij}\right)} \quad E\left(\prod_{j=1}^{k-1} p_{ij}^{\,Z_{ij}}\right) \qquad i=1\dots k$$

where the expectation is now over the arbitrary distribution of the P's. We may therefore write;

$$E_N\left(\prod_{j=1}^{k-1} N_{ij,T}^{(Z_{ij})}\right) = n_{i,T}^{\left(\sum_{j=1}^{k-1} Z_{ij}\right)} E_p\left(\prod_{j=1}^{k-1} P_{ij}^{Z_{ij}}\right) \qquad i=1\ldots k$$

(5.2.3.)

5.2.1. MOMENTS OF THE COMPOUNDED MULTINOMIAL DISTRIBUTION FOR LARGE STOCK VECTORS

In this subsection it is shown that for large grade sizes, the moments of $N_{ij,T}/n_{i,T}$ approach those of P_{ij} and so the correlation of the flows tend to the correlations of the transition probabilities.

Dividing each side of the Equation (5.2.3.) by $n_{i,T}^{\sum_{j=1}^{k-1} Z_{ij}}$, the following identity is obtained;

$$E_N\left(\prod_{j=1}^{k-1}\left[\frac{N_{ij,T}}{n_{i,T}}\right]^{Z_{ij}}\right) + O(n_{i,T}^{-1}) = E_p\left(\prod_{j=1}^{k-1} P_{ij}^{Z_{ij}}\right) + O(n_{i,T}^{-1}) \qquad i=1\ldots k$$

where $O(n_{i,T}^{-1})$ represents all terms whose order is at most $n_{i,T}^{-1}$. In this way it is said that if the grade size is large, the moments of the N/n tend to those of the P. This may be written as;

$$E_N\left(\prod_{j=1}^{k-1}\left[\frac{N_{ij,T}}{n_{i,T}}\right]^{Z_{ij}}\right) \sim E_p\left(\prod_{j=1}^{k-1} P_{ij}^{Z_{ij}}\right)$$

In particular for all i and j;

$$E_N\left(\frac{N_{ij,T}}{n_{i,T}}\right) \sim E_p(P_{ij}) \qquad \text{for large } n_{i,T} \;,$$

and this is true at all times t.

Now $n_{i,T}$ is certainly known at time T, so conditioning on this it is found that for the compounded multinomial distribution;

$$C(N_{ij,T}N_{ik,T}) = -n_{i,T}E(p_{ij})E(p_{ik}) + n_{i,T}(n_{i,T} - 1)\ C(p_{ij}p_{ik}) \tag{5.2.4.}$$

and that;

$$V(N_{ij,T}) = n_{ij,T}\ E(p_{ij})(1 - E(p_{ij})) + n_{i,T}(n_{i,T}-1)V(p_{ij}) \tag{5.2.5.}$$

Considering the correlation of $N_{ij,T}$ and $N_{ik,T}$ and dropping the i and T suffices for ease of presentation;

$$R(N_jN_k) = \frac{-E(p_j)E(p_k) + (n - 1)C(p_jp_k)}{\left(\prod_{l=j,k} E(p_l)(1 - E(p_l)) + (n - 1)V(p_l)\right)^{\frac{1}{2}}}$$

so that for large grade sizes

$$R(N_{ij,T}N_{ik,T}) \sim R(p_{ij}\ p_{ik})$$

A method for estimating the elements of the covariance matrix of the P's, denoted $S_{i,p}$, is now given. From earlier work, Equation (5.2.2.) of this section, conditional upon $n_{i,T}$ and p_{ij};

$$C(N_{ij,T}N_{ik,T}) = -n_{i,T}p_{ij}p_{ik} \tag{5.2.6.}$$

and

$$V(N_{ij,T}) = n_{i,T}(p_{ij})(1 - p_{ij}) \tag{5.2.7.}$$

Replacing each p_{ij} in the above by $E(p_{ij})$, letting $S_{i,me}$ represent the covariance matrix so obtained and substituting into (5.2.4.) and (5.2.5.) we have;

$$S_{i,N} = S_{i,me} + n_{i,T}(n_{i,T} - 1)\ S_{i,p}$$

It is therefore clear that $S_{i,p}$ may be estimated by $\hat{S}_{i,p}$ where;

$$\hat{S}_{i,p} = \frac{\hat{S}_{i,N} - \hat{S}_{i,me}}{n_{i,T}(n_{i,T} - 1)}$$

It only remains to find estimates for S_N and S_{me} which is discussed in Section 5.3.2..

5.3. PROPERTIES OF THE DIRICHLET DISTRIBUTION AND OF THE COMPOUNDED DIRICHLET-MULTINOMIAL DISTRIBUTION

In the first part of this section, the rows of the stochastic transition matrix (P) are assumed to be sampled from Dirichlet distributions, as in Bartholomew (1975). The joint raw moments of this distribution are given in Theorem 5.3.1.. Following this the density function of the stock flow matrix (N) is derived and its general factorial moments computed. In subsection 5.3.2. a method of estimating the parameters of this distribution is presented and in the final part of the section the structure of the Dirichlet distribution is further examined.

5.3.1. THE MOMENTS AND DENSITY FUNCTION OF THE DIRICHLET-MULTINOMIAL DISTRIBUTION

Bartholomew (1975) gives the density function of the Dirichlet distribution, assumed to represent the density of the ith row of the stochastic transition matrix, as;

$$d'(p_{i1}\ldots p_{ik-1}/a_{i1},\ldots a_{ik}) = \frac{\Gamma\left(\sum_{j=1}^{k} a_{ij}\right)\prod_{j=1}^{k} p_{ij}^{\,a_{ij}-1}}{\prod_{j=1}^{k}\Gamma(a_{ij})} \qquad i=1\ldots k \qquad (5.3.1.)$$

where again p_{ik} symbolises;

$$1 - \sum_{j=1}^{k-1} p_{ij} \qquad i=1\ldots k$$

A formula which may be used to obtain the joint raw moments of this distribution is provided in the following theorem, where the i suffix is dropped in order to conserve space and make reading easier.

THEOREM 5.3.1.

$$E(p_1^{Z_1} p_2^{Z_2} \cdots p_{k-1}^{Z_{k-1}}) = \frac{\prod_{j}^{k-1} (a_j + Z_j - 1)\cdots(a_j)}{(\sum_{j=1}^{k} a_j + \sum_{j=1}^{k-1} Z_j - 1)\cdots(\sum_{j=1}^{k-1} a_j)}$$

▼

PROOF

$$E(p_i^{Z_1} p_2^{Z_2} \cdots p_{k-1}^{Z_{k-1}})$$

$$= \int_0^1 \cdots \int_0^{1-\sum_{j=1}^{k-2} p_j} \frac{\Gamma(\sum_{j=1}^{k} a_j)}{\prod_{j=1}^{k} \Gamma(a_j)} \, p_1^{Z_1+a_1-1} \cdots p_{k-1}^{Z_{k-1}+a_{k-1}-1} (1-\sum_{j=1}^{k-1} p_j)^{a_{k-1}} \, \underline{dp}$$

(5.3.2.)

where $\underline{dp}$ represents $dp_1 \; dp_2 \cdots dp_{k-1}$

Letting

$$G^* = \frac{\Gamma(\sum_{j=1}^{k} a_j)}{\prod_{j=1}^{k} \Gamma(a_j)}$$

the right hand side of Equation 5.3.2. reduces to;

$$G^* \int_0^1 p_1^{Z_1+a_1-1} dp_1 \cdots \int_0^{1-\sum_{j=1}^{k-2} p_j} p_{k-1}^{Z_{k-1}+a_{k-1}-1} (1-\sum_{j=1}^{k-1} p_j)^{a_{k-1}} dp_{k-1}$$

(5.3.3.)

Now it is well known that;

$$\int_0^C p^{r-1}(C - p)^{s-1}\,dp = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}\,C^{r+s-1}$$

(5.3.4.)

Noting this the first integration is easily performed, yielding;

$$\frac{G^*(a_k)\,(Z_{k-1}+a_{k-1})}{(a_k+Z_{k-1}+a_{k-1}}\int\int\cdots\int_0^{1-\sum_{j=1}^{k-3}p_j} p_{k-2}^{\,Z_{k-2}+a_{k-2}-1}\left(1-\sum_{j=1}^{k-2}p_j\right)^{Z_{k-1}+a_{k-1}+a_k-1}dp_{k-2}$$

The next integral is seen to be of the same form as (5.3.4.), so by continuing this process the expression is found to be;

$$\frac{G^*\Gamma(a_k)\Gamma(a_{k-1}+Z_{k-1})}{\cancel{\Gamma(a_k+Z_{k-1}+a_{k-1})}}\;\frac{\cancel{\Gamma(a_k+Z_{k-1}+a_{k-1})}\Gamma(a_{k-2}+Z_{k-2})\cdots}{\cancel{\Gamma(a_k+Z_{k-1}+a_{k-1}+a_{k-2}+Z_{k-2})}}\;\frac{\cancel{(\sum_{j=2}^{k}a_j+\sum_{j=2}^{k-1}Z_j)}\Gamma(a_1+Z_1)}{\Gamma(\sum_{j=1}^{k}a_j+\sum_{j=1}^{k-1}Z_j)}$$

$$=\frac{\Gamma(\sum_{j=1}^{k}a_j)}{\prod_{j=1}^{k}\Gamma(a_j)}\;\frac{\Gamma(a_k)\prod_{j=1}^{k-1}\Gamma(a_j+Z_j)}{\Gamma(\sum_{j=1}^{k}a_j+\sum_{j=1}^{k-1}Z_j)}$$

$$=\frac{\prod_{j=1}^{k-1}(a_j+Z_j-1)\cdots\cdots\cdots\cdots(a_j)}{\sum_{j=1}^{k}a_j+\sum_{j=1}^{k-1}Z_j-1\;\cdots\cdots\cdots(\sum_{j=1}^{k}a_j)}$$

■

It is remarked upon that an alternative representation of the above equation is;

$$E(P_1^{Z_1} P_2^{Z_2} \cdots P_{k-1}^{Z_{k-1}}) = \frac{\prod_{j=1}^{k-1} \prod_{h=1}^{Z_j} (a_j+h-1)}{\prod_{h=1}^{\sum_{j=1}^{k-1} Z_j} (\sum_{j=1}^{k} a_j+h-1)}$$

and as a corollary, the generalised factorial moment of the Dirichlet-Multinomial distribution is (see Equation 5.2.2.);

$$E_N(\prod_{j=1}^{k-1} N_{ij,T}^{(Z_{ij})}) = n_{i,T}^{(\sum_{j=1}^{k-1} Z_{ij})} \frac{\prod_{j=1}^{k-1} \prod_{h=1}^{Z_{ij}} (a_{ij}+h-1)}{\prod_{h=1}^{\sum_{j=1}^{k-1} Z_{ij}} (\sum_{j=1}^{k} a_{ij}+h-1)}$$

The first and second central moments of the stock flow vectors are now easily obtained;

$$E(N_{ij,T}) = n_{i,T} \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}} \qquad i=1\ldots k$$

$$C(N_{ij,T} N_{ih,T}) = -n_{i,T} \frac{a_{ij} \quad a_{ih} \quad (\sum_{j=1}^{k} a_{ij}+n_{i,T})}{(\sum_{j=1}^{k} a_{ij})^2 \ (\sum_{j=1}^{k} a_{ij}+1)}$$

$$V(N_{ij,T}) = n_{i,T} \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}} \left(1 - \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}}\right) \frac{(\sum_{j=1}^{k} a_{ij} + n_{i,T})}{(\sum_{j=1}^{k} a_{ij} + 1)}$$

(5.3.5.)

A relationship is now apparent between the covariance matrix of the stock flows under a Dirichlet distribution of the probability transition matrix $(S_{i,N})$ and the covariance matrix of the stock flows that would have been obtained if the p's were assumed known and equal to their expectation $(S_{i,me})$. So that we may write;

$$S_{i,N} = A_i S_{i,me}$$

(5.3.6.)

where

$$A_i = \frac{\sum_{j=1}^{k} a_{ij} + n_{i,T}}{\sum_{j=1}^{k} a_{ij} + 1}$$

(5.3.7.)

From the above it is easily seen that the correlation coefficients are unchanged by ascribing a probability distribution to the p's. So that;

$$R_{i,N} = R_{i,me}$$

Recalling the density functions of the Dirichlet and Multinomial distributions, Equations 5.3.1. and 5.2.1.;

$$d'(p_{i1}\ldots p_{ik-1}/a_{i1}\ldots a_{ik}) = \frac{\Gamma(\sum_{j=1}^{k} a_{ij}) \prod_{j=1}^{k} p_{ij}^{a_{ij}-1}}{\prod_{j=1}^{k} \Gamma(a_{ij})} \qquad i=1\ldots k$$

where p_{ik} symbolises $1 - \sum_{j=1}^{k-1} p_{ij}$, and;

$$m(n_{i1,T}\ldots n_{ik-1,T}/n_{i,T},p_{i1}\ldots p_{ik-1}) = \frac{n_{i,T}!}{\prod_{j=1}^{k} n_{ij,T}!} \prod_{j=1}^{k} p_{ij}^{\,n_{ij,T}} \qquad i=1\ldots k$$

where $n_{ik,T}$ symbolises $n_{i,T} - \sum_{j=1}^{k-1} n_{ij,T}$.

The density function of the Dirichlet-Multinomial distribution is given by;

$$d'm(n_{i1,T}\ldots n_{ik-1,T}/n_{i,T},a_{i1}\ldots a_{ik})$$

$$= \frac{n_{iT}!}{\prod_{j=1}^{k} n_{ij,T}!} \frac{\Gamma(\sum_{j=1}^{k} a_{ij})}{\prod_{j=1}^{k}\Gamma(a_{ij})} \int_0^1 \int_0^{1-p_{i1}} \cdots \int_0^{1-\sum_{j=1}^{k-2} p_{ij}} \prod_{j=1}^{k} p_{ij}^{\,n_{ij,T}+a_{ij}-1}\, dp_{i1}\ldots dp_{ik-1}$$

Equation 5.3.4. can now be recursively employed to yield the density function of the Dirichlet-Multinomial distribution as;

$$d'm(n_{i1,T}\ldots n_{ik-1,T}/n_{i,T},a_{i1}\ldots a_{ik})$$

$$= \frac{n_{iT}!}{\prod_{j=1}^{k} n_{ij,T}!} \frac{\Gamma(\sum_{j=1}^{k} a_{ij})}{\prod_{j=1}^{k}\Gamma(a_{ij})} \frac{\prod_{j=1}^{k}\Gamma(n_{ij,T}+a_{ij})}{\Gamma(n_{i,T}+\sum_{j=1}^{k} a_{ij})} \qquad i=1\ldots k$$

where

$$n_{ik,T} = n_{i,T} - \sum_{j=1}^{k-1} n_{ij,T}$$

This is the distribution of exactly $n_{ij,T}$ occurrences of a transition from grade i to grade j in the time interval (T,T + 1). At time T, $n_{i,T}$ is certainly known, but $a_{i1} \ldots a_{ik}$ must be estimated. A method of estimating these parameters is presented in the following subsection.

5.3.2. A METHOD FOR ESTIMATING THE PARAMETERS OF THE DIRICHLET-MULTINOMIAL DISTRIBUTION

Let the observed flow of employees between grade i and grade j, in the time interval (T, T + 1) be denoted $n^o_{ij,T}$. Over any time period there will be k^2 observed flows in a k-grade system, however only k(k - 1) of these will be independent.

Recalling Equation 5.3.5.;

$$E(N_{ij,T}) = \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}} n_{i,T} \qquad \begin{array}{l} i=1\ldots k \\ j=1\ldots k-1 \end{array}$$

and adopting the normal convention by denoting the mean of the $N_{ij,T}$ by $\bar{N}_{ij,t}$ where;

$$\bar{N}_{ij,t} = \sum_{T=1}^{t} N_{ij,T}/t$$

it is simply verified that;

$$E(\bar{N}_{ij,t}) = \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}} \bar{n}_{i,t}$$

Clearly $E(\bar{N}_{ij,t})$ may be estimated by $\bar{n}^o_{ij,t}$ and so it is possible to obtain k(k - 1) estimates of the a_{ij} as;

$$\hat{a}_{ij} = \frac{\bar{n}^o_{ij,t}}{\bar{n}_{i,t}} \sum_{j=1}^{k} \hat{a}_{ij} \qquad \begin{array}{l} i=1\ldots k \\ j=1\ldots k-1 \end{array}$$

A further k equations are required to complete the solution.

Recalling Equation 5.3.7.;

$$A_i = \frac{n_{i,T} + \sum_{j=1}^{k} a_{ij}}{1 + \sum_{j=1}^{k} a_{ij}} \qquad i=1...k$$

so all that is required is an estimate $\hat{A}_i$ for A_i.

Now Equation 5.3.6. states;

$$S_{i,N} = A_i S_{i,me} \qquad i=1...k$$

so on taking determinants;

$$|S_{i,N}| = A_i^{k-1} |S_{i,me}| \qquad i=1...k$$

From Section 5.2.;

$$S_{i,me\ jj} = n_{i,T} E(p_{ij})(1 - E(p_{ij}))$$

and

$$S_{i,me\ jh} = - n_{i,T} E(p_{ij}) E(p_{ih})$$

These may be unbiasedly estimated by;

$$S_{i,me\ jj} = \bar{n}^{o}_{ij,t} \frac{(\bar{n}_{i,t} - \bar{n}^{o}_{ij,t})}{\bar{n}_{i,t}}$$

and

$$S_{i,me\ jh} = \frac{- \bar{n}^{o}_{ij,t}\ \bar{n}^{o}_{ih,t}}{\bar{n}_{i,t}}$$

To achieve an unbiased estimator for $S_{i,T}$ it is necessary to standardise the data. The standardised flow from the ith grade to the jth grade in the time interval (T,T + 1) is denoted $n^{*}_{ij,T}$ and defined as;

$$n^{*}_{ij,T} = n^{o}_{ij,T} \times \frac{\bar{n}_{i,t}}{n_{i,T}}$$

Here

$$\overrightarrow{n}^{*}_{ij,t} = \bar{n}_{i,t}\left(\frac{n^{o}_{ij,t}}{n_{i,t}}\right) \neq \bar{n}^{o}_{ij,t}$$

In most organisations the grade sizes will not vary too greatly and;

$$\overrightarrow{n}^{*}_{ij,t} \simeq \bar{n}^{o}_{ij,t}$$

Estimators of the elements of $S_{i,N}$ are now easily calculated as;

$$\hat{S}_{i,N\ jj} = \sum_{T=1}^{t} (n^{*}_{ij,T} - \overrightarrow{n}^{*}_{ij,t})^2 \,/\, t-1$$

and

$$\hat{S}_{i,N\ jh} = \sum_{T=1}^{t} (n^{*}_{ij,T} - \overrightarrow{n}^{*}_{ij,t})(n^{*}_{ih,t} - \overrightarrow{n}^{*}_{ih,t}) \,/\, t-1$$

The estimation procedure is now completed by setting;

$$\sum_{j=1}^{k} a_{ij} = \frac{\bar{n}_{i,t} - \hat{A}_i}{\hat{A}_i - 1}$$

where

$$\hat{A}_i = (|\hat{S}_{i,N}| \,/\, |\hat{S}_{i,me}|)^{(k-1)^{-1}}$$

(5.3.8.)

Summarising, the estimates of the a_{ij} are given by;

$$\hat{a}_{ij} = \frac{\bar{n}^{o}_{ij,t}}{\bar{n}_{i,t}}\left(\frac{\bar{n}_{i,t} - \hat{A}_i}{\hat{A}_i - 1}\right) \qquad i,j=1\ldots k$$

where $\hat{A}_i$ is given in Equation 5.3.8..

EXAMPLE 5.3.1.

In order to generate observations, actual parameter values for the a_{ij} were assumed. Only transitions from the ith grade are given, where;

$a_1 = 4$

$a_2 = 8$ and so $\sum_{j=1}^{3} a_j = 20$

$a_3 = 8$

Given this $E(P_1) = 0.2$ $V(P_1) = 0.008$

$E(P_2) = 0.4$ $V(P_2) = 0.011$

These distributions were then simulated and three sampled values obtained from each.

TABLE 5.3.1.

t	1	2	3
P_1	.04	.23	.35
P_2	.60	.31	.39

The number of employees in this grade were assumed to vary 10, 11, 9 so that the flows between the grades could again be simulated. The resulting, and standardised flows are presented in the following table.

TABLE 5.3.2.

t	1	2	3	-
n_1^o	1	3	4	2.67
n_1^*	1	2.73	4.44	2.72
n_2^o	5	2	4	3.67
n_2^*	5	1.82	4.44	3.75

The estimated covariance matrices are;

$$\hat{S}_{i,N} = \begin{bmatrix} 2.96 & -0.49 \\ -0.49 & 2.88 \end{bmatrix} \qquad \hat{S}_{i,me} = \begin{bmatrix} 1.96 & -0.98 \\ -0.98 & 2.32 \end{bmatrix}$$

$$\left|\hat{S}_{i,N}\right| = 8.28 \qquad \left|\hat{S}_{i,me}\right| = 3.59$$

$$A_i = (8.28 \ / \ 3.59)^{\frac{1}{2}} = 1.52$$

and so we have;

$$\sum_{j=1}^{3} \hat{a}_{ij} = 16.31, \quad \hat{a}_1 = 4.35, \quad \hat{a}_2 = 5.99 \text{ and } \hat{a}_3 = 5.97$$

as estimates of;

$$\sum_{j=1}^{3} a_j = 20, \quad a_1 = 4, \quad a_2 = 8 \text{ and } a_3 = 8$$

5.3.3. THE STRUCTURE OF THE DIRICHLET DISTRIBUTION

In this subsection it is shown that the Dirichlet distribution represents a very particular form for the varying p_{ij}'s.

Recalling Equation 5.3.1.;

$$d'(p_{i1} \ldots p_{ik-1} \ / \ a_{i1} \ldots a_{ik})$$

$$= \frac{\Gamma\left(\sum_{i=1}^{k} a_{ij}\right) \prod_{j=1}^{k} p_{ij}^{a_{ij}-1}}{\prod_{j=1}^{k} \Gamma(a_{ij})} \qquad i=1\ldots k$$

Adopting the Stieltjes integral notation;

$$dD' = dD' \ (p_{i1} \ldots p_{ik-1} \ / \ a_{i1} \ldots a_{ik})$$

$$= \frac{\Gamma\left(\sum_{i=1}^{k} a_{ij}\right) \prod_{j=1}^{k} p_{ij}^{a_{ij}-1}}{\prod_{j=1}^{k} \Gamma(a_{ij})} \ dp_{i1} \ldots dp_{ik-1}$$

Now by definition;

$$\Gamma(a) = \int_0^{\infty} e^{-xb}\, x^{a-1}\, b^{a}\, dx$$

So we may form a new equation, if x_i is independent of $p_{i1}\dots p_{ik-1}$

$$dG(p_{i1}\dots p_{ik-1}, x_i \,/\, a_{i1}\dots a_{ik})$$

$$= \frac{\prod_{j=1}^{k} p_{ij}^{a_{ij}-1}}{\prod_{j=1}^{k} \Gamma(a_{ij})} \; e^{-x_i b_i} x_i^{\sum_{i=1}^{k} a_{ij}-1} \; b_i^{\sum_{i=1}^{k} a_{ij}} \; dp_{i1}\dots dp_{ik-1} dx_i$$

which on simplification becomes;

$$dG = \prod_{j=1}^{k} \frac{b_i^{a_{ij}} \; p_{ij}^{a_{ij}-1}}{\Gamma(a_{ij})} dp_{i1}\dots dp_{ik-1} \; e^{-b_i x_i} \; x_i^{\sum_{j=1}^{k} a_{ij}-1} \; dx_i$$

(5.3.9.)

Now making the transformation;

$$Y_{ij} = P_{ij} X_i \qquad i,j = 1\dots k$$

it is clear that;

$$X_i = \sum_{j=1}^{k} Y_{ij} \qquad i = 1\dots k$$

and so

$$P_{ij} = Y_{ij} \Big/ \sum_{j=1}^{k} Y_{ij} \qquad i = 1\dots k$$

The Jacobian is now given by;

$$\frac{\partial(p_{i1}\cdots p_{ik-1},x_i)}{\partial(y_{i1}\cdots y_{ik-1},y_{ik})} = \begin{vmatrix} \frac{\sum_{j=1}^{k} y_{ij}-y_{i1}}{(\sum_{j=1}^{k} y_{ij})^2} & \cdots\cdots & \frac{-y_{ik-1}}{(\sum_{j=1}^{k} y_{ij})^2} & 1 \\ \vdots & & \vdots & \vdots \\ \frac{-y_{i1}}{(\sum_{j=1}^{k} y_{ij})^2} & \cdots\cdots & \frac{\sum_{j=1}^{k} y_{ij}-y_{ik-1}}{(\sum_{j=1}^{k} y_{ij})^2} & 1 \\ \frac{-y_{i1}}{(\sum_{j=1}^{k} y_{ij})^2} & \cdots\cdots & \frac{-y_{ik-1}}{(\sum_{j=1}^{k} y_{ij})^2} & 1 \end{vmatrix}$$

$$= \frac{1}{(\sum_{j=1}^{k} y_{ij})^{2k-2}} \begin{vmatrix} \sum_{j=1}^{k} y_{ij}-y_{i1} & \cdots\cdots & -y_{ik-1} & 1 \\ \cdot & \cdots\cdots & & \cdot \\ \cdot & \cdots\cdots & & \cdot \\ \cdot & \cdots\cdots & & \cdot \\ -y_{i1} & \cdots\cdots & \sum_{j=1}^{k} y_{ij}-y_{ik-1} & 1 \\ -y_{i1} & \cdots\cdots & -y_{ik-1} & 1 \end{vmatrix}$$

which can be verified quickly to be;

$$= \frac{1}{(\sum_{j=1}^{k} y_{ij})^{k-1}}$$

This can now be substituted into Equation 5.3.9. to give;

$$dH(y_{i1}\ldots y_{ik})$$

$$= \prod_{j=1}^{k} \frac{b_i^{a_{ij}} \, y_{ij}^{a_{ij}-1} \, e^{-b_i \sum_{j=1}^{k} y_{ij}} \, \cancel{\left(\sum_{j=1}^{k} y_{ij}\right)^{\sum_{j=1}^{k} a_{ij}-1}}}{\Gamma(a_{ij}) \, \cancel{\sum_{j=1}^{k} y_{ij}^{a_{ij}-1}} \, \cancel{\left(\sum_{j=1}^{k} y_{ij}\right)^{k-1}}} \, dy_{i1}\ldots dy_{ik}$$

$$= \prod_{j=1}^{k} \frac{b_i^{a_{ij}} y_{ij}^{a_{ij}-1} \, e^{-b_i y_{ij}} \, dy_{ij}}{\Gamma(a_{ij})} \qquad i=1\ldots k$$

Therefore the marginal distributions of the Y_{ij} are gamma distributions, all with the same amplitude parameter b_i. Letting the dummy variable x_i be identified with $n_{i,T}$, it is clear that if the p_{ij}, for $i = 1\ldots k$ and $j = 1\ldots k - 1$, are to be governed by Dirichlet distributions and be independent of the $n_{i,T}$ then the N_{ij} must have independent gamma distributions with parameters b_i. The converse is also true, so that if the N_{ij} are not governed by gamma distributions with parameters b_i then the p_{ij} are not independent of $n_{i,T}$.

5.4. AN APPLICATION OF THE GENERAL PROBLEM TO SUPPLY FORECASTING

In Section 5.3. it was assumed that if the p_{ij} at time T were known, then the $n_{ij,T}$ would be multinomially distributed. Letting $q_{ij,T}$ represent the observed proportion of employees who move from grade i to grade j in the time interval (T,T + 1), clearly;

$$q_{ij,T} = n^{o}_{ij,T} / n_{i,T}$$

and

$$E(Q_{ij,T}) = E(P_{ij})$$

Previously the distribution of the probability transition matrix, from which the p_{ij} were sampled, has been assumed to be time invariant. Here this assumption is relaxed and a time dependent P-distribution is admissible. Letting $\bar{p}_{ij.T}$ represent the mean of the elements of the P-distribution and as before adopting matrix notation, a three stage process can be constructed;

$$q_T = p_T + dq_T \quad , \quad E(dq_T) = 0$$

$$p_T = \bar{p}_T + dp_T \quad , \quad E(dp_T) = 0$$

$$\bar{p}_T = \bar{p}_{T-1} + \bar{dp}_T \quad , \quad E(\bar{dp}_T) = 0$$

Relating the above equations to the General Problem we have;

$$\begin{bmatrix} p_T \\ \bar{p}_T \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{T-1} \\ \bar{p}_{T-1} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dp_T \\ \bar{dp}_T \end{bmatrix}$$

for the system equation and;

$$q_T = (1 \quad 0) \begin{pmatrix} p_T \\ \bar{p}_T \end{pmatrix} + dq_T$$

for the observation equation. This is now seen to be a special case of the General Problem presented in Appendix 2. Noting this the optimal filtered and predicted linear estimates of the grade sizes and their covariances are easily obtained.

CHAPTER SIX

MANPOWER DEMAND FORECASTING

6.1. Introduction

6.2. The forecasting of manpower demand given the future work loads of an organisation

6.2.1. The incorporation of technological change into the manpower demand forecasting model

6.3. The forecasting of workloads from series with growth and seasonal characteristics

6.3.1. A growth model

6.3.2. A cyclical model

6.3.3. Workload forecasting

6.1. INTRODUCTION

In this chapter the problem of forecasting an organisation's manpower requirements is considered. In the review of the present manpower demand models (Chapter 2 Section 4) it was suggested that the forecasting should be a subtle interplay between management foresight and methods of statistical analysis. The spirit of this statement pervades the models of this chapter.

In Section 2 the relationship between the firm's manpower requirements and the workloads serviced is explored. Here the workloads are assumed to be known or, more strictly, given for all future times. In some cases they will be taken directly from the corporate plan of the organisation and in others they may have been the results of some statistical forecasting method on past values. A suitable method that is useful in the latter case is the subject of the final section of this chapter.

Until now the technique that has been used to model the relationship of manpower demand to workloads has been time-independent multivariate linear regression. This means that the degree to which independent variables are deemed to contribute to the total manpower is assumed constant for all time. Cameron and Nash (1974) in their comprehensive analysis of this situation have shown, using a programme developed by the Central Statistical Office (Brown and Durbin (1968,1971)) that the regression coefficients estimated from their data were <u>not</u> stable over time. In fact it does not seem that this situation is a rare occurence, rather, that the time variance of the regression coefficients is a fundamental property of most real life systems. Beer (1966) when discussing industrial operations, and in particular systems with many variables, writes;

'...the importance of a particular variable in such a system

> is a question of degree, a question of judgement, a question of convention. Moreover, the importance it has by any of these criteria, change from moment to moment. This does not mean merely the numerical value assumed by the variable is changing - that is the nature of variables, one of the things about the system, that we know how to handle. No it means more: the STRUCTURAL RELEVANCE of the variable inside the system is changing with time.'

The emphasis given to the structural relevance of the variable is Beer's.

It is therefore concluded that any model used to represent the manpower requirements of an organisation, derived from the workloads it has to service, should reject the tenuous assumption that the regression coefficients (structure) of the model be time invariant. The fact that the regression coefficients vary over time is fully recognised and is incorporated into the model suggested for forecasting manpower demand, proposed in Section 2 of the chapter.

The future value of the workloads is an important input into the above model if it is to be used in its forecasting mode, as opposed to its pure filtering facility. In many real life planning exercises, including the above, it is necessary for managers to predict the future values of some variable from past data. Often this time series will be short, contain 'freak' observations and undergo major changes in time. Many excellent methods exist for dealing with long time series, notably 'Box and Jenkins' (1970) techniques, which are well behaved. However, it is a primary tenet that many real life series experienced in industrial situations are principally <u>not</u> well behaved.

In examining projected time series that had trend components over some period of time, Jantsch (1967) identified four main types of curves observed in technological forecasting. In each case these 'growth' curves exhibited either gradual or rapid change of shape after the initial observed time-stable situation which is indicative of a change in the external factors which effect the curves. What is important then, is the ability to identify these changes when they occur and to forecast in the light of this. The growth model expounded in the first part of Section 3 has this facility.

Another property that has been found to occur in many industrial time series is a cyclical nature. This may be monthly, seasonal or an x-year business cycle. Again, this type of attribute may in fact change over time. A model which copes with this type of data is presented in Section 3.2..

Of course, many series possess both a growth and cyclical nature. It is the property of linear models that their addition is also linear. The combination of the separate models is carried out in the final part of Section 3. Throughout this Section 3 the nomenclature applicable to the prediction of workloads from growth and seasonal data is employed. However, the models presented are quite generally applicable to this type of data.

6.2. THE FORECASTING OF MANPOWER DEMAND GIVEN THE FUTURE WORKLOADS OF AN ORGANISATION

Cameron and Nash (1974) proposed the following model to describe the total staff (S_t) of an organisation in time period t;

$$S_t = a_o + \sum_{i=1}^{n} a_i W_{it} + e_t$$

where

W_{it} = the number of units of workload w_i served in time period t

a_i = the regression parameters

and

e_t = the error term at time t

It is further assumed that the error term has zero mean and constant variance.

They discovered that the regression coefficients were, in fact, not time stable. In modern usage stability usually refers to a system governed by some dynamic equilibrium. Who would say that a governor used to control a mechanical instrument were unstable? But certainly it is not static. Cyberneticians have adopted the word homeostatic to describe a process that is inherently stable, although not necessarily static and the word homeorhetic for a process that is undergoing a stable change. It is therefore preferable to say that the regression coefficients of the above model are not static but that their instability has yet to be examined.

In keeping with the above, a model of the form;

$$S_{t+1} = \sum_{i=0}^{n} W_{it+1} a_{it+1} + e_{t+1} \quad (5.2.1.)$$

is proposed where

$a_{i,t+1}$ = the time dependent regression parameters and

$W_{o,t+1} = 1$, for all time

As before it is convenient to introduce matrix notation so that the Equation 6.2.1. is equivalent to;

$$S_{t+1} = W_{t+1} a_{t+1} + e_{t+1} \tag{6.2.2.}$$

where

a_{t+1} = n + 1 column vector of the a_{it+1}

and

W_{t+1} = 1 x n + 1 matrix of the W_{it+1}

Now as the regression coefficients are known to vary over time, this is accounted for by the admittance of a structural equation of the form;

$$a_{t+1} = a_t + da_t \tag{6.2.3.}$$

So that the regression parameters are assumed to be time dependent. The difference between sucessive time periods t + 1 and t is defined as da_t and it is further assumed that the expectation of this difference is zero, and that it's variance is finite.

By direct comparison of Equations 6.2.2. and 6.2.3. with Equations (1) and (2) of Chapter 3 Section 3.1., it is clear that the above model is yet another submodel contained in the General Problem. The relevant transformation is given by;

$$S_t = z_t$$

$$W_t = \psi_t$$

$$a_t = x_t$$

$$e_t = dz_t$$

The covariance matrices of the processes must also be specified, for which the methods of Chapter 4 Section 2 will be useful.

The recurence equations for the updating of the regression coefficients are then seen to be;

$$\hat{a}_{ot+1,t} = \hat{a}_{ot,t}$$

$$C_{ot+1,t} = C_{ot,t} + R_t$$

$$\hat{a}_{ot+1,t+1} = \hat{a}_{ot+1,t} + B_{t+1}(S_{t+1} - W_{t+1}\hat{a}_{ot+1,t})$$

$$C_{ot+1,t+1} = (I - B_{t+1}W_{t+1})C_{ot+1,t}$$

where

$B_{t+1} = C_{ot+1,t}W^T_{t+1}(W_{t+1}C_{ot+1,t}W^T_{t+1} + \mathfrak{Z}_{t+1})^{-1}$ and is a (n + 1 x 1) matrix.

$C_{ot+u,t}$ = the error covariance matrix of $\tilde{a}_{ot+u,t}$ given observations up to and including S_t and is an n x 1 symmetric positive definite matrix.

R_t = covariance matrix of the observed regression process at time t and is an n x 1 symmetric positive definite matrix.

$\mathfrak{Z}_{t+1}$ = covariance matrix of the structural process at time t+1 and is a positive scalar.

Now, <u>if</u> the regression coefficients were time independent, that is $R_t = 0$ for all time, and if the initial covariance matrix $C_{o,oo}$ is large in comparison with S_t, then the estimated regression coefficients are independent of $\hat{a}_{o,oo}$. Moreover, the model then reduces to the classical multiple regression analysis model and affords an elegant method for updating the regression coefficients as more observations are gathered. This is particularly useful with data containing a large number of observations or a large number of regression parameters, as the only data storage requirements are the optimal estimates of the regression coefficients and their error covariance matrix and also <u>no</u> matrix inversion is necessary. The objective of making the regression parameters time dependent has now been achieved.

6.2.1. <u>THE INCORPORATION OF TECHNOLOGICAL CHANGE INTO THE MANPOWER DEMAND FORECASTING MODEL</u>

In the new model the expectation of the change in these parameters was assumed to be zero. This assumption might be invalid under some interpretations of the model. In order to illustrate this point a model based on the corporate plan and on industrial engineering techniques is proposed and its solution in then given.

It might be that the workloads described above were in fact independent tasks that have to be carried out by any organisation to meet its corporate objective. It is usually the case that the industrial engineering department will have gathered information enabling the 'standard' man years required to produce a unit of each workload to be calculated. In the light of this a model of the form;

$$S_{t+1} = W_{t+1} a_{t+1} + e_{t+1} \qquad (6.2.4.)$$

is again appropriate. Here $a_{o,t+1}$ might be indicative of a basic staffing level needed, even though actual production may be near zero. This type of assumption is not uncommon in econometrics, where the staffing level may be related to the minimal capital investment.

As before the values of the 'manpower utilisation' coefficients (a_t) can be allowed to be time dependent and again a structural model of the form;

$$a_{t+1} = a_t + da_t$$

might be deemed appropriate. Due to the significance applied to the regression parameters, the assumption that the expectation of da_t is zero is in some doubt. This is due to changes in technology and efficiency, usually, and hopefully affecting organisations in such a way that productivity increases. This would clearly increase manpower utilisation and so lower the regression coefficients over time.

The question arises now of how this can be incorporated into the model and if possible also retain the simple recurrence equations of the General Problem. Two methods are proposed;

1) The introduction of an additive/subtractive parameter b'_t to represent the change in manpower utilisation due to increases in productivity and technology over time.

2) The introduction of a parameter b_t to represent the rate of change in manpower saving due to technical progress.

The first model has equations;

$$a_{t+1} = a_t + b'_{t+1} + da_t$$

$$b'_{t+1} = b'_t + db'_t$$

and are seen to be a special structural case of the General Problem by noting the equivalent formulation;

$$\begin{bmatrix} a_{t+1} \\ b'_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_t \\ b'_t \end{bmatrix} + \begin{bmatrix} da_t + db'_t \\ db'_t \end{bmatrix}$$

The second model is also a special case of the General Problem. This is so as its equations;

$$a_{t+1} = b''_{t+1} a_t + da_t$$

$$b''_{t+1} = b''_t + db''_t$$

are the same as;

$$\begin{bmatrix} a_{t+1} \\ b''_{t+1} \end{bmatrix} = \begin{bmatrix} b''_{t+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_t \\ b''_t \end{bmatrix} + \begin{bmatrix} da_t \\ db''_t \end{bmatrix}$$

In both cases Equation 6.2.4. has to be modified to;

$$S_{t+1} = (W_{t+1} \quad 0) \begin{pmatrix} a_{t+1} \\ b''_{t+1} \end{pmatrix} + e_{t+1} \tag{6.2.5.}$$

where $b'_{t+1} = b''_{t+1}$ when appropriate.

Both models are now simple to apply for one-step ahead forecasts and no problem exists in forecasting j-steps ahead with the first model. However, with the second model the forecasting equations are not simple as Theorem 3.3.1. does not apply. This is now shown below.

Firstly

$$\phi_{k+1,k} = \begin{pmatrix} \overset{*}{b}''_{t+1} & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{where } \overset{*}{b}_t = \text{diag}(b_t)$$

and therefore

$$\phi_{k+j,k} = \begin{pmatrix} \prod_{i=1}^{j} \overset{*}{b}''_{t+j} & 0 \\ 0 & 1 \end{pmatrix}$$

This is unknown for $j > 1$.

Now recalling the first term of Equation 9, Theorem 3.3.1.;

$$\langle \phi_{k+j,k} \tilde{x}_{ok,k}, \hat{x}_{k,k} \rangle$$

$$= \int_{\Omega} \left[\prod_{i=1}^{j} \overset{*}{b}''_{t+j} \tilde{a}_{ot,t} \right]^{T} \hat{a}_{t,t} + \tilde{b}''^{T}_{ot,t} \hat{b}''_{t,t} \, d\mu$$

$$= \int_{\Omega} \left[(\overset{*}{b}''_{t+1})^{j} \tilde{a}_{ot,t} \right]^{T} \hat{a}_{t,t} \, d\mu + \text{other terms in } (\overset{*}{b}''_{t+1})^{i} \quad i=j-1\ldots1 \tag{6.2.6.}$$

This expression is in general non-zero. It is therefore deduced that the j-step ahead prediction can be obtained from the first j moments of the distribution of $b_{t+1,t}$ and thus the simple recurrence equations are lost. However, if the rate of change in manpower saving due to technical change can be predicted independently of the above series or is assumed constant then Equation 6.2.6. can be factorised to terms involving;

$$\int_{\Omega} \tilde{a}^{T}_{ot,t} \hat{a}_{t,t} d\mu = 0$$

and once again Theorem 3.1.1. holds and the simple recurrence equations can be employed.

In summary; if the regression coefficients are static then an elegant method of updating their estimates which is computationally very efficient has been presented. If they are not static but their expected change in a unit time period is zero then the model described by Equations 6.2.2. and 6.2.3. is appropriate. Further, it is simple to incorporate an additive change, a constant rate of change or an independently estimated rate of change to represent the variation in the regression coefficients due to technical progress and other alterations in the efficiency of the organisation into the model and still retain the simple recurrence relations.

Throughout this section it has been assumed that the workloads have been known or given. In the next section methods of predicting workloads from series that have an inherent trend and seasonal characteristics are proposed.

6.3. THE FORECASTING OF WORKLOADS FROM SERIES WITH GROWTH AND SEASONAL CHARATERISTICS

In this section the forecasting of future workloads from past observations is considered. Firstly, a model that gives optimal estimates of the workloads when only a growth factor is present will be described. Secondly, a model that copes with the cyclical nature which is often present in industrial time series is given. Finally, these two models are combined to form a third model that is applicable to time series that possess both of these attributes.

Throughout this section the terms in which the models are described are particular to the estimation of workloads, however, the models have general applicability to other situations when the assumptions of the model are valid.

6.3.1. A GROWTH MODEL

Here it is assumed that the observed workloads serviced W_t^1, are related to an underlying workload parameters u_t exactly except for some small random error dW_t^1. This can be written as;

$$W_{t+1}^1 = u_{t+1} + dW_{t+1}^1 \qquad (6.3.1.)$$

Further, it is assumed that the underlying parameter has a growth charateristic, such that;

$$u_{t+1} = u_t + g_{t+1} + du_t \qquad (6.3.2.)$$

where g_{t+1} is the step increase in time interval (t, t+1) and du_t is a small disturbance vector. Now the incremental growth parameter is constant in time, barring a small error, so that;

$$g_{t+1} = g_t + dg_t \qquad (6.3.3.)$$

It is also assumed that the expectation of each of the error terms is zero and that they have finite variances.

Briefly expressed the model under consideration consists of observations from a process that exhibits a stochastic linear growth. The equivalence of the model to the General Problem of Chapter 3 is readily seen when it is written in the following form;

$$\begin{pmatrix} u_{t+1} \\ g_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_t \\ g_t \end{pmatrix} + \begin{pmatrix} du_t + dg_t \\ dg_t \end{pmatrix} \tag{6.3.4.}$$

$$W^1_{t+1} = (1 \quad 0) \begin{pmatrix} u_{t+1} \\ g_{t+1} \end{pmatrix} + dW^1_t \tag{6.3.5.}$$

Equation 6.3.4. can be further simplified to;

$$\begin{pmatrix} u_{t+1} \\ g_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_t \\ g_t \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} du_t \\ dg_t \end{pmatrix} \tag{6.3.6.}$$

In this case Equations 6.3.6. and 6.3.5. are a specific form of the extended model of Appendix 2. Once the respective error covariance matrices of dW^1_t, du_t and dg_t have been specified, estimation of future parameter values follows immediately by application of the recurrence equations of Theorem 3.3.6. to initial estimates.

6.3.2. A CYCLICAL MODEL

Consider a process, propagating through time, that exhibits only a cyclical nature of time period N. At each time point t an observation of the process, denoted W^2_t, is taken. It is assumed that the cyclical nature of the process at time t can be described by a (N x 1) factor vector denoted f_t. It is clear that a constraint of the factor vector must be;

$$(11\ldots1)f_t = 0 \tag{6.3.7.}$$

Now, t can be written as a multiple of the time cycle and a remainder such that;

$$t = Nj + k \qquad\qquad k = 1,\ldots,N$$

In this case the time point t relates to the kth part of the n-cycle.

A model of the form;

$$W^2_{t+1} = f^k_{t+1} + dW^2_{t+1} \tag{6.3.8.}$$

is therefore proposed, where f^k_{t+1} is the appropriate element of the cyclical factor vector f_{t+1} at time t + 1 and dW^2_{t+1} represents a small random observation error vector. It is also assumed that $E(dW^2_{t+1}) = 0$ and that $E(dW^2_{t+1}\ dW^{2T}_{\tau+1}) = S_t d_{t\tau}$

Further, the cyclical factor vector is allowed to change with time in such a way that the Equation;

$$f_{t+1} = f_t + df_t \tag{6.3.9.}$$

describes its transition. The term df_t takes on its obvious meaning of a small random disturbance vector where $E(df_t) = 0$ and $E(df_t df^T_\tau) = R_t d_{t\tau}$.

The correspondence of this process to the General Problem is easily seen by redefining Equation 6.3.8. as;

$$W^2_{t+1} = \psi^2_{t+1} f_{t+1} + dW^2_{t+1} \tag{6.3.10.}$$

where

$$\psi^2_t = \psi^2_{nj+k} = (\delta_{1k}\delta_{2k}\ldots\delta_{Nk})$$

The constraint of Equation 6.3.7., however introduces a restricted form for the covariance structure of the model. In order that the constraint be preserved by the recurrence relationships of Theorem 3.3.6., $C_{0,00}$ and R_t (the error covariance matrix of f_t) must be of the following form;

$$R_t = \begin{bmatrix} R_{11} & R_{12} & \cdot & R_{1N} \\ R_{21} & R_{22} & \cdot & R_{2N} \\ \cdot & \cdot & \cdot & \\ R_{N1} & R_{N2} & \cdot & R_{NN} \end{bmatrix}$$

such that

$$R_{ii} > 0 \qquad i = 1,\ldots N$$

$$R_{ij} = R_{ji} \qquad i,j = 1,\ldots,N$$

and

$$\sum_{i=1}^{n} R_{ij} = 0 \qquad j = 1,\ldots,N$$

This completes the cyclical model. In the next subsection the linear growth model described by Equation 6.3.2. is combined with a 4-cycle (seasonal) model. This type of combined series is frequently found in industrial time series.

6.3.3. WORKLOAD FORECASTING

Cameron and Nash (1974) use a multiplicative seasonal 'Box-Jenkins' $(0,1,1)(0,1,1)_4$ model to predict eight of the nine workloads under examination. The form of the model;

$$\nabla \nabla_4 W_{it} = (1 - \theta B)(1 - \Theta B^4) a_{it}, \qquad |\theta| < 1, \; |\Theta| < 1$$

carries the assumption that both the seasonal and growth variations are time invariant. After obtaining estimates of the parameters θ and Θ over 35 observations they then tested the model for time

periods 36-43. They found that one of the estimated series performed badly and attributed this to a probable change in the seasonal pattern, for which they had insufficient data to fit a new static model.
The seasonal linear growth stochastic model that follows is proposed as an improvement of the Box and Jenkins model. Not only does it allow for time variable parameters, little data and gives interpretation to the process but also encompasses Cameron and Nash's model for their remaining workload (W_5);

$$\nabla_4 W_{5t} = a_{5t}$$

Inspection of the graphs of their workload series would also suggest that the parameters for at least W_2, W_3 and W_8 might have changed over the observation period.

It is considered that the observation of workload i at time t + 1(W_{it+1}) is made up of an underlying value, a seasonal adjustment and a random observation error. The observation system is then;

$$W_{i\ t+1} = u_{i\ t+1} + \psi^2_{i\ t+1} f_{i\ t+1} + dW_{i\ t+1}$$

where the definition of the above terms follows through from the previous subsections. The structural system is taken straight from Equations 6.3.4. and 6.3.9. with the introduction of the suffix i to indicate the relationship to the ith workload.

$$u_{i\ t+1} = u_{it} + g_{i\ t+1} + du_{it}$$

$$g_{i\ t+1} = g_{it} + dg_{it}$$

$$f_{i\ t+1} = f_{it} + df_{it}$$

Putting this into the framework of the General Problem it is apparent that;

$$W_{i\,t+1} = (1 \quad 0 \quad \psi^2_{it+1}) \begin{bmatrix} u_{i\,t+1} \\ g_{i\,t+1} \\ f_{i\,t+1} \end{bmatrix} + dW_{i\,t+1}$$

and

$$\begin{bmatrix} u_{i\,t+1} \\ g_{i\,t+1} \\ f_{i\,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1_N \end{bmatrix} \begin{bmatrix} u_{it} \\ g_{it} \\ f_{it} \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1_N \end{bmatrix} \begin{bmatrix} du_{it} \\ dg_{it} \\ df_{it} \end{bmatrix}$$

for each workload W_i. Clearly, it is a simple matter to incorporate the prediction of the vector W_t into one model of the General Problem type but this would necessitate the storage of large matrices with many zero elements which is computationally inefficient.

Once again, this model could be used to update estimates from a static process; that is du_t, dg_t and df_t are zero for all time without matrix inversion or recalculating the parameters using all observations. In saying this however there is much to be lost in assuming stationality of time series parameters as we are forced to reject the model each time they change, whereas the above model handles this situation most elegantly and without the difficulties caused by the stationality assumptions.

CHAPTER SEVEN

STOCHASTIC CONTROL OF MANPOWER SYSTEMS

7.1. Introduction

7.1.1. A review of the preferability of a one-step strategy in the control of manpower systems

7.2. The choice of a cost function

7.3. Deterministic control theory

7.3.1. Solution of the one-period deterministic control problem

7.3.2. Solution of the two-period deterministic control problem

7.3.3. Solution of the N-period deterministic control problem

7.4. Stochastic control theory

7.4.1. Solution of the one-period stochastic control problem

7.4.2. Solution of the two-period stochastic control problem

7.4.3. Solution of the N-period stochastic control problem

7.5. An example in stochastic control theory

7.1. INTRODUCTION

In this chapter the problems encountered when controlling manpower systems are considered. In the review of manpower control theory, Chapter 2 Section 5, some prejudice against the one-step strategy advocated by Bartholomew (1975) was apparent. The general dislike for this type of strategy stems from its resemblance to the 'fire fighting' methods often used but even more frequently deplored by British management. In the second part of this section the justification of adopting such a strategy is questioned with specific reference to Bartholomew (1975).

In Section 2 a cost function is developed which is thought to portray the true cost of controlling a manpower system more accurately. Having generated such a function, the theory that enables the construction of sequential optimal control inputs is presented. Sections 3 and 4 comprise the solution of the deterministic and stochastic control problems respectively, both deal with general N time period planning horizon. Although there is no great difficulty in going straight to the stochastic situation this would, however, involve the introduction of a vast amount of notation in a short period in order to preserve the compactness of the solution. It is for this reason that the notation is built up slowly through the discussion of the deterministic problom, which is also included to facilitate the understanding of the structure of the control problem under consideration. Whilst working towards the solution of controlling manpower systems it is found quite generally that the method employed should be to start at the end of the manpower planning period and work, somewhat counter-intuitively, from the end of this period backwards in time to the present. This point is emphasised continually throughout the chapter. In this presentation the solution to an unconstrained control input is only considered, although the method of solution extends to systems having structural constraints.

This chapter is then concluded with a short example that exhibits the simplicity of the calculations which are involved in deriving control inputs. Moreover, it is found that the estimation problem discussed in Chapters 3 and 4 and the control problem examined here can be considered as completely independent processes. The re-evaluation of the one-step strategy now follows.

7.1.1. A REVIEW OF THE PREFERABILITY OF A ONE-STEP STRATEGY IN THE CONTROL OF MANPOWER SYSTEMS

Bartholomew (1973, 1975) has definite views in favour of the one-step strategy, in this section it is intended to re-examine the value of such a strategy. In Bartholomew (1975), he states;

> 'The one-step strategy will certainly make at least as much progress at the first step as any other strategy but, on the other hand, it cannot be better than a T-step strategy after any multiple of T steps. There can thus be no choice which is 'best' in an overall sense. The length of the planning horizon must, therefore, be decided by balancing short term and long term advantages. It appears difficult to give a satisfactory mathematical analysis of the question, but calculations are easily made and, fortunately, the matter seems to be so clear cut as to make a deeper analysis superfluous.'

The review of the one-step strategy that follows is extracted from work communicated to Professor Bartholomew early in 1976.

Firstly, it is shown that the one-step strategy may in fact be <u>better</u> than a T-step strategy after a multiple of T steps. Here the one and two step strategies are compared after 4 steps. As in Bartholomew (1975), Table 1, the following values are used;

$$P = \begin{bmatrix} .5 & .4 & 0 \\ 0 & .6 & .3 \\ 0 & 0 & .8 \end{bmatrix} \quad \text{Initial vector} \quad (0 \quad 1 \quad 0) \quad \text{and goal vector} \quad (.286 \quad .286 \quad .428)$$

The results of applying the one and two-step strategies to this problem are presented in Table 7.1.1. below;

TABLE 7.1.1.

Comparison of the one and two-step strategies over four time periods, details given in the text.

	ONE-STEP STRATEGY				TWO-STEP STRATEGY			
TIME	$N_i(1*)$			10D	$N_i(2*)$			10D
0	.000	1	.000	7.748	.000	1	.000	7.748
1	.100	.600	.300	1.496	.090	.600	.310	1.509
2	.180	.400	.420	0.243	.176	.396	.428	0.242
3	.232	.312	.456	0.044	.231	.308	.461	0.046
4	.262	.280	.458	0.016	.262	.277	.461	0.018

Clearly the two-step strategy is inferior to the one-step strategy after four time periods and this is sufficient to prove the first point.

Secondly it is thought that the preferability of the one-step strategy over the N-step strategy has not been adequately resolved. Bartholomew, with reference to his Table 2 (see Table 7.1.2.), argues that the values of the distance function at various time periods differ significantly in favour of the one-step strategy so as to make a deeper analysis superfluous. Never the less a deeper analysis was performed and it was found that the entries in his Table 2 were a little misleading. It is certainly true that the minimum distance of 0.61 after two time periods can be achieved by the two-step strategy from a distance of 2.16 after one time period, but it can also be reached from a distance of 1.16. Similarly the 3-step strategy can reach the minimum of 0.31, from a distance of 1.16 after one time period and 0.61 after two periods. Thus Table 2 of Bartholomew can be revised. This is presented in Table 7.1.3.

TABLE 7.1.2.

An extract of Table 2 of Bartholomew, (1975);

TIME \ T*	1	2	3
0	2.44	2.44	2.44
1	1.16	2.16	2.16
2	0.61	0.61	1.66
3	0.32	1.21	0.31

TABLE 7.1.3.

A revision of Table 7.1.2.;

TIME \ T*	1	2	3
0	2.44	2.44	2.44
1	1.16	1.16	1.16
2	0.61	0.61	0.61
3	0.32	0.32	0.31

Comparing the entries in Table 7.1.3., it would be difficult to conclude other than that the strategies have very similar properties in this specific example.

At this point a re-examination of the comparison process that has been employed above is undertaken. Various step strategies have been used and all time values of their distance from a goal have been calculated. These relative distances have then been compared, but the strategies have not been designed so as to minimise the distance from goal at all times, only for one specific time period T*. Thus, this comparison process is somewhat unsatisfactory. This naturally leads to the rejection of the simple distance function used above and therefore a new distance/cost function must be established. The choice of a suitable function is the topic under discussion in the ensuing section.

7.2. THE CHOICE OF A COST FUNCTION

In Chapter 2 it was stated that a simplified view of control theory was the matching of the supply and demand forecasts of the manpower system. By this it is assumed implicitly that there is something to be gained by 'controlling' the manpower system or equivalently a loss associated with not controlling it. So, already a subjective measure of 'cost' has arisen. It is the aim of the section to formulise an objective cost function that has the same properties as the actual loss associated in industry by failing to control manpower. The following two sections will then show optimal methods of minimising this cost function over time.

Dealing with a deterministic system first, it is clear that any cost function should include some measure that reflects the difference between the desired stock vector x* and the projected system vector x, that is;

$$£' = f((x^* - x),\ldots)$$

where £' represents some cost function. Now as the system propogates through time, cost will certainly accumulate through all stages of an N-period planning horizon. Hence;

$$£'_{N,0} = \sum_{i=1}^{N} f((x_i^* - x_i),\ldots)$$

since it is assumed that the system starts in state x_0 for which no control can be exerted.

One of the most widely used functions that expresses the cost of some discrepancy is the squared error function. This has the property that it attributes equal loss to both positive and negative errors of the same magnitude. It has the additional feature that the cost also increases with the amplitude and duration of the discrepancy. Both of these characteristics were thought to be valid first approximations to the costs incurred in industrial manpower systems. To add further realism to the cost function, a time variant symmetric positive definite weighting matrix is introduced. This may be used to model costs that are different for distinct grades which may vary over time. The cost function now takes the form;

$$£'_{N,0} = \sum_{i=1}^{N} (\overset{*}{x}_i - x_i)^T \, \mathbb{D}_i (\overset{*}{x}_i - x_i)$$

(7.2.1.)

where

$\mathbb{D}_i$ is an n x n symmetric positive definite weighting matrix.

The above cost function now reflects the loss associated with not obtaining a desired stock vector at each point of observation. However, this form assumes that no further cost is associated with controlling the manpower system. In industry costs are certainly incurred as when redundancy or recruitment occurs. This may be accounted for by the introduction of a second type of cost function. The deterministic cost function is now given as;

$$£_{N,0} = \sum_{i=1}^{N} (\overset{*}{x}_i - x_i)^T \mathbb{D}_i (\overset{*}{x}_i - x_i)^T + y_{i-1}^T \, \mathbb{B}_{i-1} y_{i-1}$$

(7.2.2.)

where

y_i is the control input at time i

and

$\mathbb{B}_i$ is symmetric positive definite weighting matrix.

The difference between the x and y suffices is due to control y_{i-1} affecting x_i.

It is clear that the objective of management must be to minimise this cost over the planning period. The function $L_{N,0}$ is defined to represent the minimum value attainable, that is;

$$L_{N,0} = \min_{y_0} \min_{y_1} \ldots \min_{y_{N-1}} \mathcal{L}_{N,0}$$

(7.2.3.)

So the aim is to minimise the 'system error' plus the 'control effect'.

Having now established a general deterministic cost function, it is possible to re-examine the cost function of Bartholomew (1975), discussed in the previous section. Bartholomew's cost criteria for the k-step strategy is equivalent to;

$$\min_{y_0} \min_{y_1} \ldots \min_{y_{k-1}} (\overset{*}{x} - x_k)^T 1 (\overset{*}{x} - x_k)$$

(7.2.4.)

with the additional constraint

$$y_{ij} \geqslant 0 \qquad \text{for all } i = 0, \ldots k-1$$
$$\text{and} \qquad j = 1, \ldots n$$

The one-step strategy involves minimising the cost function at each point in time adding further constraints to the minimising process and therefore cannot be better than a k-step strategy over k time periods, which is in agreement with Bartholomew. This is a rather unjust criticism of the one-step strategy as it is being compared with other strategies over <u>different</u> cost functions. What is more important is the fact that the one-step strategy cannot be better than the optimal control strategy obtained by finding the solution for $y_0 \ldots y_{N-1}$ of Equation 7.2.3., again subject to the constraint of positive values of y_{ij} only being admissable.

The control problem is now examined in more detail. It should be clear that since the only knowledge that is possessed of the manpower system is through observation, the control vectors must satisfy the the constants of the system and be a function of the observations. Such a sequence of control vectors will be called admissible. Suppose a sequence of admissible control vectors exist which are the solution to Equation 7.2.3. and let them be denoted by $\overset{*}{y}_0, \overset{*}{y}_1, \ldots \overset{*}{y}_{N-1}$. Using the same initial part of the optimal control sequence $\overset{*}{y}_0, \overset{*}{y}_1, \ldots \overset{*}{y}_{N'-1}$ we can assume that over the time interval (N',N) there is another admissible control sequence $y_{N'}, \ldots y_{N-1}$ that has a cost less than $\overset{*}{y}_{N'}, \ldots \overset{*}{y}_{N-1}$. Then it is clear that $\overset{*}{y}_0, \ldots \overset{*}{y}_{N'-1} y'_N \ldots y_{N-1}$ has a total cost less than $\overset{*}{y}_0, \ldots \overset{*}{y}_{N-1}$. However, this is a contradiction as the control sequence which is the solution of Equation 7.2.3. is the minimum. This means that an optimal control sequence must have the property that regardless of any earlier control sequence the remainder of the control sequence must be optimal. As will be seen in the following sections this property of optimal control sequences is paramount to their derivation.

The derivation of an optimal control sequence in one stage at first seems a mammoth task, however, the above indicates a much simpler approach. The method is not the intuitive way of trying to gain as much as possible in the first time period but the converse. Starting at the end of the planning horizon, the method is to obtain the minimal cost over the period (N-1,N) from an arbitrary x_{N-1} position and then work backwards in time to the present. This is the method employed in the next two sections.

So far the formulation has been purely deterministic, however the real problem is stochastic. It is necessary therefore that the cost function to be minimised is;

$$\pounds_{N,0} = E\left(\sum_{i=1}^{N} (\overset{*}{x}_i - x_i)^T \Phi_i (\overset{*}{x}_i - x_i) + y_{i-1}^T B_{i-1} y_{i-1}\right)$$

given the observations up to the time the control must be implemented. This representation of the cost is sufficient to cover both the deterministic problem and the stochastic problem and so further notation is unnecessary.

Before embarking upon the task of determining optimal control sequences, a convenient restructuring of the problem is established. Consider the system equation of the model as;

$$x_{k+1} = \phi_{k+1,k} x_k + \theta_{k+1,k} y_k + dx_k$$

where it is desired to find the solution of;

$$\min_{y_0} \min_{y_1} \cdots \min_{y_{N-1}} E\left(\sum_{i=1}^{N} (\overset{*}{x}_i - x_i)^T \Phi_i (\overset{*}{x}_i - x_i) + y_{i-1}^T \theta_{i-1} y_{i-1}\right)$$

An equivalent form of the problem is;

$$x'_{k+1} = \phi'_{k+1,k} x'_k + \theta'_{k+1,k} y'_k + dx'_k$$

where

$$x'_k = \begin{bmatrix} x_k \\ \overset{*}{x}_k \end{bmatrix} \qquad \phi'_{k+1,k} = \begin{bmatrix} \phi_{k+1,k} & 0 \\ 0 & \overset{*}{\phi}_{k+1,k} \end{bmatrix}$$

$$dx'_k = \begin{bmatrix} dx_k \\ 0 \end{bmatrix} \qquad \theta'_{k+1,k} = \begin{bmatrix} \theta_{k+1,k} \\ 0 \end{bmatrix}$$

$$y'_k = \begin{bmatrix} y_k \end{bmatrix}$$

and where $\overset{*}{\phi}_{k+1,k}$ is a matrix which satisfies;

$$\overset{*}{x}_{k+1} = \overset{*}{\phi}_{k+1,k} \overset{*}{x}_k$$

and it is desired to find the solution of;

$$\min_{y_0} \min_{y_1} \cdots \min_{y_{N-1}} E\left(\sum_{i=1}^{N} x_i'^T \Phi'_i x'_i + y_{i-1}^T \theta_{i-1} y_{i-1}\right)$$

where

$$\mathbb{Q}_i' = \begin{bmatrix} \mathbb{Q}_i & -\mathbb{Q}_i \\ -\mathbb{Q}_i & \mathbb{Q}_i \end{bmatrix}$$

The equivalence is simply established by multiplying out the terms. Noticing this reduction it is sufficient to find an optimal control sequence that minimises;

$$E\left(\sum_{i=1}^{N} x_i^T \mathbb{Q}_i x_i + y_{i-1}^T \mathbb{B}_{i-1} y_{i-1}\right)$$

given the observations up to the time of implementation of control.

7.3. DETERMINISTIC CONTROL THEORY

In this section the solution to the problem of controlling a deterministic manpower system is obtained. As noted in the previous section we will work from the end of the planning horizon backward in time to the present. The approach will be to solve the N period planning problem for time intervals (N - 1,N) and (N - 2,N). After this a form of the solution for time (N - J,N), that is compatable with the two previous solutions, will be assumed and the solution for time (N - J - 1,N) derived. The principle of mathematical induction can then be invoked in order to prove that the solution is valid for all J. In developing the solution both the optimal control sequence $\overset{*}{y}_{N-1}\cdots\overset{*}{y}_0$ and the costs $L_{N,N-1}\cdots L_{N,0}$ will be obtained. The solution given is for an unconstrained control input, however, the methodology for a constrained control input is analogous and is a development of Equation 15 of Bartholomew (1975).

Since, in the deterministic formulation of this problem the state vector x_k is directly observable there is no need to consider the redundant observation system. Here the task is to find the optimal control sequence $\overset{*}{y}_{N-1},\cdots\overset{*}{y}_0$ that minimises;

$$\pounds_{N,0} = \sum_{i=1}^{N} x_i^T \Phi_i x_i + y_{i-1}^T \theta_{i-1} y_{i-1}$$

where

$$x_{k+1} = \phi_{k+1,k} x_k + \theta_{k+1,k} y_k$$

As an aside, the usual formulation of the deterministic manpower problem can easily be placed within this structure;

$$n_{t+1} = P_{t+1,t} n_t + r_t$$

and the problem is to minimise;

$$\sum_{t=1}^{N} \sum_{j=1}^{n} w_j(\overset{*}{n}_{tj} - n_{tj})$$

where

n_{ti} = the ith element of the stock vector at time t

$P_{t+1,t}$ = the known transition matrix

r_t = the recruitment vector at time t

w_j = a weighting of the relative cost of errors between different grades

with the further constraints that;

$$r_{ij} \geqslant 0 \text{ and often } (1,1\ldots1)n_t = N \text{ for all time}$$

These latter constraints, although reducing the set of admissible control vectors, do not affect the method of solution that should be employed.

7.3.1. SOLUTION OF THE ONE-PERIOD DETERMINISTIC CONTROL PROBLEM

The optimal cost in this first problem is given by;

$$L_{N,N-1} = \min_{y_{N-1}} (x_N^T \Phi_N x_N + y_{N-1}^T B_{N-1} y_{N-1}) \tag{7.3.1.}$$

where

$$x_N = \phi_{N,N-1} x_{N-1} + \theta_{N,N-1} y_{N-1} \tag{7.3.2.}$$

On substituting Equation 7.3.2. into Equation 7.3.1., using the fact that Φ_N is symmetric and collecting terms, the following expression for $L_{N,N-1}$ is obtained;

$$L_{N,N-1} = \min_{y_{N-1}} (x_{N-1}^T \phi_{N,N-1}^T \Phi_N \phi_{N,N-1} x_{N-1} + 2x_{N-1}^T \phi_{N,N-1}^T \Phi_N \theta_{N,N-1} y_{N-1}$$

$$+ y_{N-1}^T (\theta_{N,N-1}^T \Phi_N \theta_{N,N-1} + B_{N-1})\, y_{N-1}) \tag{7.3.3.}$$

The turning point of this expression is found by taking the gradient with respect to y_{N-1} and equating this to zero, whereby;

$$2x^T_{N-1}\phi^T_{N,N-1}\Phi_N\theta_{N,N-1} + 2\overset{*}{y}{}^T_{N-1}(\theta^T_{N,N-1}\Phi_N\theta_{N,N-1} + B_{N-1}) = 0$$

and conditional on the inverse of the bracketed term existing, the optimal control for time n-1 is seen to be;

$$\overset{*}{y}_{N-1} = -(\theta^T_{N,N-1}\Phi_N\theta_{N,N-1} + B_{N-1})^{-1}\theta^T_{N,N-1}\Phi_N\phi_{N,N-1}x_{N-1}$$

(7.3.4.)

The fact that this is a minimum can be seen by observing that the second gradient of Equation 7.3.3. is positive definite if $\theta^{-1}_{N,N-1}$ exists and it can always be made to do so.

It is clear that notation and space difficulties are going to arise as the development continues towards the solution of the N-period control problem. For this reason two further symbols are introduced in order to ease the exposition;

$$U_N \triangleq \Phi_N$$

(7.3.5.)

and

$$\Upsilon_{N-1} \triangleq -(\theta^T_{N,N-1}U_N\theta_{N,N-1} + B_{N-1})^{-1}\theta^T_{N,N-1}U_N\phi_{N,N-1}$$

(7.3.6.)

by using these definitions Equation 7.3.4. reduces to;

$$\overset{*}{y}_{N-1} = \Upsilon_{N-1}x_{N-1}$$

(7.3.7.)

The evaluation of $L_{N,N-1}$ can now begin;

$$L_{N,N-1} = x^T_N U_N x_N + \overset{*}{y}{}^T_{N-1}B_{N-1}\overset{*}{y}_{N-1}$$

(7.3.8.)

which becomes on substitution of Equation 7.3.7. and Equation 7.3.2.;

$$L_{N,N-1} = x^T_{N-1}\phi^T_{N,N-1}U_N\phi_{N,N-1}x_{N-1} \cdots$$

$$+ 2x^T_{N-1}\phi^T_{N,N-1}U_N B_{N,N-1}\Upsilon_{N-1}x_{N-1}$$

$$+ x^T_{N-1}(\Upsilon^T_{N-1}(B^T_{N,N-1}U_N B_{N,N-1} + B_{N-1})\Upsilon_{N-1})x_{N-1}$$

$$= x^T_{N-1}\left\{ \phi^T_{N,N-1}U_N\phi_{N,N-1} + \phi^T_{N,N-1}U_N B_{N,N-1}\Upsilon_{N-1} \right\}x_{N-1}$$

$$= x^T_{N-1}V_{N-1}x_{N-1} \qquad (7.3.9.)$$

where

$$V_{N-1} \triangleq \phi^T_{N,N-1}U_N\phi_{N,N-1} + \phi^T_{N,N-1}U_N B_{N,N-1}\Upsilon_{N-1} \qquad (7.3.10)$$

The one-period solution can now be summarised as;

$$\overset{*}{y}_{N-1} = \Upsilon_{N-1}x_{N-1} \qquad (7.3.11)$$

$$L_{N,N-1} = x^T_{N-1}V_{N-1}x_{N-1} \qquad (7.3.12.)$$

where

$$\Upsilon_{N-1} = -(B^T_{N,N-1}U_N B_{N,N-1} + B_{N-1})^{-1}B^T_{N,N-1}U_N\phi_{N,N-1} \qquad (7.3.13.)$$

$$U_N = \mathbb{O}_N \qquad (7.3.14.)$$

$$V_{N-1} = \phi^T_{N,N-1}U_N(\phi_{N,N-1} + B_{N,N-1}\Upsilon_{N-1}) \qquad (7.3.15.)$$

7.3.2. SOLUTION OF THE TWO-PERIOD DETERMINISTIC CONTROL PROBLEM

The optimal cost is now given by;

$$L_{N,N-2} = \min_{y_{N-2}} \min_{y_{N-1}} \sum_{i=N-1}^{N} (x_i^T \mathbb{O}_i x_i + y^T_{i-1}B_{i-1}y_{i-1})$$

As noted previously the value of y_{N-1}, does not affect the stock vector at time N - 1, hence $L_{N,N-2}$ reduces to;

$$L_{N,N-2} = \min_{y_{N-2}} (x^T_{N-1}(\Phi_{N-1} + V_{N-1})x_{N-1} + y^T_{N-2}B_{N-2}y_{N-2}) \quad (7.3.16.)$$

Letting

$$U_{N-1} = \Phi_{N-1} + V_{N-1} \quad (7.3.17.)$$

Equation 7.3.16. further reduces to;

$$L_{N,N-1} = \min_{y_{N-2}} (x^T_{N-1}U_{N-1}x_{N-1} + y^T_{N-2}B_{N-2}y_{N-2}) \quad (7.3.18.)$$

which is the same form as Equation 7.3.1. of the previous subsection, the solution of which is already known. Noting this it is apparent that;

$$\overset{*}{y}_{N-2} = \Upsilon_{N-2}x_{N-2}$$

$$L_{N,N-2} = x^T_{N-2}V_{N-2}x_{N-2}$$

where

$$\Upsilon_{N-2} = -(\theta^T_{N-1,N-2}U_{N-1}\theta_{N-1,N-2} + B_{N-2})\theta^T_{N-1,N-2}U_{N-1}\phi_{N-1,N-2}$$

$$U_{N-1} = \Phi_{N-1} + V_{N-1}$$

$$V_{N-2} = \phi^T_{N-1,N-2}U_{N-1}(\phi_{N-1,N-2} + \theta_{N-1,N-2}\Upsilon_{N-2})$$

It is now a simple matter to postulate the solution of a J-period problem.

7.3.3. SOLUTION OF THE N-PERIOD DETERMINISTIC CONTROL PROBLEM

It is first assumed that the equations for a J-period problem are;

$$\overset{*}{y}_{N-J} = \Upsilon_{N-J}x_{N-J} \quad (7.3.19.)$$

$$L_{N,N-J} = x^T_{N-J}V_{N-J}x_{N-J} \quad (7.3.20.)$$

$$\Upsilon_{N-J} = -(\theta^T_{N-J+1,N-J}U_{N-J+1}\theta_{N-J+1,N-J} + B_{N-J})\theta^T_{N-J+1,N-J}U_{N-J+1}\phi_{N-J+1,N-J} \quad (7.3.21.)$$

$$U_{N-J+1} = \mathbb{D}_{N-J+1} + V_{N-J+1}$$

(7.3.22.)

$$V_{N-J} = \phi^T_{N-J+1,N-J} U_{N-J+1} (\phi_{N-J+1,N-J} + \theta_{N-J+1,N-J} \Upsilon_{N-J})$$

(7.3.23.)

The J + 1-period problem consists of finding the control sequence;

$$\overset{*}{y}_{N-J-1}, \overset{*}{y}_{N-J} \ldots \overset{*}{y}_{N-1}$$

that minimises;

$$\pounds_{N,N-J-1} = \sum_{i=N-J}^{N} x_i^T \mathbb{D}_i x_i + y_{i-1}^T \theta_{i-1} y_{i-1}$$

Now the control vectors $\overset{*}{y}_{N-J}, \ldots, \overset{*}{y}_{N-1}$ do not affect the stock vector x_{N-J}, therefore the task reduces to finding $\overset{*}{y}_{N-J-1}$ that minimises;

$$\pounds_{N,N-J-1} = x^T_{N-J} \mathbb{D}_{N-J} x_{N-J} + y^T_{N-J-1} \theta_{N-J-1} y_{N-J-1} + x^T_{N-J} V_{N-J} x_{N-J}$$

since we assumed that the solution of the J-period problem is as above.

By writing;

$$U_{N-J} = \mathbb{D}_{N-J} + V_{N-J}$$

the task is to find $\overset{*}{y}_{N-J-1}$ that satisfies;

$$\min_{y_{N-J-1}} x^T_{N-J} U_{N-J} x_{N-J} + y^T_{N-J-1} \theta_{N-J-1} y_{N-J-1}$$

which is the same form as Equation 7.3.1., whose solution is already known. Thus we obtain;

$$\overset{*}{y}_{N-J-1} = \Upsilon_{N-J-1} x_{N-J-1}$$

$$L_{N,N-J-1} = x^T_{N-J-1} V_{N-J-1} x_{N-J-1}$$

$$\Upsilon_{N-J-1} = -(\theta^T_{N-J,N-J-1} U_{N-J} \theta_{N-J,N-J-1} + \theta_{N-J-1}) \theta^T_{N-J,N-J-1} U_{N-J} \phi_{N-J,N-J-1}$$

$$U_{N-J} = \mathbb{D}_{N-J} + V_{N-J}$$

$$V_{N-J-1} = \phi^T_{N-J,N-J-1} U_{N-J} (\phi_{N-J,N-J-1} + \theta_{N-J,N-J-1} \Upsilon_{N-J-1})$$

The principle of mathematical induction can now be employed. Since the solution is true for $J = 1$ and for $J + 1$ given it is true for J, it is true for all $J = 1,2,...N$.

The solution of the deterministic control problem is again a set of recursive equations. The computation procedure would be as follows;

1. Set $U_N = \Phi_N$ and a counter equal to zero
2. Using U_N in Equation 7.3.21., obtain the value of Υ_{N-1}.
3. Store the value of Υ_{N-1}
4. Substitute Υ_{N-1} and U_N into Equation 7.3.23., this yields V_{N-1}
5. Store the value of V_{N-1}
6. Obtaining Φ_{N-1} and V_{N-1} from above, Equation 7.3.22. now gives U_{N-1}
7. Set counter = counter + 1
8. If counter = N go to Instruction 10
9. Set N - 1 = N, return to Instruction 1
10. The series $\Upsilon_0,...,\Upsilon_{N-1}$ and $V_0,...,V_{N-1}$ have been stored.

Given x_0 and the updating relationship $x_{k+1} = \phi_{k+1,k}x_k + \theta_{k+1,k}y_k$,

$$y_k = \Upsilon_k x_k \qquad k = 0,1...N-1$$

and

$$L_{k,0} = L_{N,0} - L_{N,k}$$

where

$$L_{N,k} = x_k^T V_k x_k$$

can be evaluated.

Having determined the optimal control sequence for a deterministic manpower system, in the next section the stochastic problem is tackled.

7.4. STOCHASTIC CONTROL THEORY

The development of the optimal control sequence for the stochastic problem can cow be carried out in an analogous way to the deterministic solution obtained in the previous section. It is known that any control sequence must be a function of the available observations, this is because direct knowledge of the system is clouded by random error. As in the deterministic solution, derivation of the optimal control sequence will commence at the end of the planning horizon and work towards the present.

7.4.1. SOLUTION OF THE ONE-PERIOD STOCHASTIC CONTROL PROBLEM

The optimal cost is now given by;

$$L_{N,N-1} = \min_{y_{N-1}} E(x_N^T \mathbb{D}_N x_N + y_{N-1}^T \mathbb{B}_{N-1} y_{N-1}) \tag{7.4.1.}$$

where

$$x_N = \Phi_{N,N-1} x_{N-1} + \mathbb{B}_{N,N-1} y_{N-1} + dx_{N-1} \tag{7.4.2.}$$

Substituting Equation 7.4.2. into Equation 7.4.1., using the symmetry of $\mathbb{D}_N$ and collecting terms, the following equation holds;

$$\begin{aligned} L_{N,N-1} = \min_{y_{N-1}} \; & E((\Phi_{N,N-1} x_{N-1} + \mathbb{B}_{N,N-1} y_{N-1} + dx_{N-1})^T \mathbb{D}_N \\ & (\Phi_{N,N-1} x_{N-1} + \mathbb{B}_{N,N-1} y_{N-1} + dx_{N-1}) \\ & + y_{N-1}^T \mathbb{B}_{N-1} y_{N-1}) \end{aligned} \tag{7.4.3.}$$

Dropping the time suffices, expanding and rearranging the terms;

$$\begin{aligned} L_{N,N-1} = \min \; & E(x^T \Phi^T \mathbb{D} \, \Phi \, x + 2x^T \Phi^T \, \mathbb{D} \, \mathbb{B} \, y + y^T (\mathbb{B} \, \mathbb{D} \, \mathbb{B} + \mathbb{B}) y \\ & + dx^T \mathbb{D} dx + 2x^T \Phi^T \mathbb{D} dx + 2dx^T \mathbb{D} \, \mathbb{B} \, y) \end{aligned}$$

Now the second to last term on the right hand side is equivalent to;

$$2\langle \mathbb{D}\,\phi\, x, dx\rangle = 0$$

the last term also vanishes since y_{N-1} is a function of $z_{N-1},\ldots,z_1$

and

$$\langle z_i dx_{N-1}\rangle = 0 \qquad \forall\ i = 1,\ldots,N-1$$

The task therefore reduces to finding $\overset{*}{y}_{N-1}$ that minimises;

$$E(x^T\phi^T\mathbb{D}\,\phi x + 2x^T\phi^T\mathbb{D}\,\theta y + y^T(\theta\,\mathbb{D}\,\theta + \mathbb{B})y + dx^T\mathbb{D}dx)$$

(7.4.4.)

The fact that

$$E_X = E_Z(E(x/z))$$

is now employed, so that Equation 7.4.5. may be written as;

$$E_{Z_{N-1},\ldots,Z_1}(E(x^T\phi^T\mathbb{D}\phi x + 2x^T\phi^T\mathbb{D}\theta y + y^T(\theta\mathbb{D}\theta + \mathbb{B})y + dx^T\mathbb{D}dx/z_{N-1},\ldots,z_1)$$

(7.4.5.)

Now obviously it is enough to minimise the inner expectation in order to minimise $L_{N,N-1}$.

Taking the gradient of Equation 7.4.5. and equating this to zero and reinstating the time suffices the following expression obtains;

$$2E(x_N^T/z_{N-1},\ldots,z_1)\phi^T_{N,N-1}\mathbb{D}_N\theta_{N,N-1} + 2\overset{*}{y}{}^T_{N-1}(\theta^T_{N,N-1}\mathbb{D}_N\theta_{N,N-1} + \mathbb{B}) = 0$$

But as before $E(x^T_{N-1}/z_{N-1},\ldots,z_1) = \hat{x}_{0\ N-1,N-1}$ and so

$$\overset{*}{y}_{N-1} = \Upsilon_{N-1}\hat{x}_{0\ N-1,N-1}$$

(7.4.6.)

where Υ_{N-1} takes on the same value as it did in the deterministic problem, which is shown in Equation 7.3.6..

$L_{N,N-1}$ can now be evaluated, defining once again;

$$U_N \triangleq \mathbb{O}_N \quad ,$$

(7.4.7.)

$$L_{N,N-1} = E(x_N^T U_N x_N + y_{N-1}^T \mathbb{B}_{N-1} y_{N-1}$$

(7.4.8.)

which equals on dropping the time suffices;

$$E(x^T \phi^T \mathbb{O} \phi x + 2x^T \phi^T \mathbb{O} \theta \Upsilon \hat{x} + \hat{x}^T \Upsilon^T (\theta \mathbb{O} \theta + \mathbb{B}) \Upsilon \hat{x} + dx^T \mathbb{O} dx)$$

(7.4.9.)

$$= E(x^T Vx) + E(\tilde{x}^T \phi^T U\theta (\theta^T U\theta + \mathbb{B})^{-1} \theta^T U \phi \tilde{x}) + E(dx^T \mathbb{O} dx)$$

$$= E(x^T Vx) + E(\tilde{x}^T \phi^T U\theta \Upsilon \tilde{x}) + E(dx^T \mathbb{O} dx)$$

$$= E(x_{N-1}^T V_{N-1} x_{N-1}) + \mathrm{TRACE}(\phi^T_{N,N-1} U_N \theta_{N,N-1} \Upsilon_{N-1} C_{0\ N-1,N-1}) + \mathrm{TRACE}(U_N R_{N-1})$$

(7.4.10.)

where R_{N-1} and $C_{0\ N-1,N-1}$ were first defined in Chapter 3 Section 2.1. and Section 3.3. respectively and where V_{N-1} is defined by Equation 7.3.15.. Equation 7.4.10. is now further simplified to give;

$$L_{N,N-1} = E(x_{N-1}^T V_{N-1} x_{N-1}) + \xi_{N-1}$$

(7.4.11.)

where

$$\xi_{N-1} = \mathrm{TRACE}(\phi^T_{N,N-1} U_N \theta_{N,N-1} \Upsilon_{N-1} C_{0\ N-1,N-1} + U_N R_{N-1})$$

(7.4.12.)

In examining Equation 7.4.11. it is found that the first term is the expected value of the deterministic solution and the second term represents the additional cost due to stochastic dynamics of the system and the estimation error.

7.4.2. SOLUTION OF THE TWO-PERIOD STOCHASTIC CONTROL PROBLEM

The optimal cost for the two-period problem is given by;

$$L_{N,N-2} = \min_{y_{N-2}} \min_{y_{N-1}} E \sum_{i=N-1}^{N} (x_i^T \mathbb{O}_i x_i + y_{i-1}^T \mathbb{B}_{i-1} y_{i-1})$$

(7.4.13.)

As noted in the previous section, the value of y_{N-1} has no effect on x_{N-1}, the stock vector at time N-1, and so $L_{N,N-2}$ can be rewritten as;

$$L_{N,N-2} = \min_{y_{N-2}} E(x^T_{N-1}(\Phi_{N-1} + V_{N-1})x_{N-1} + y^T_{N-2}B_{N-2}y_{N-2} + \xi_{N-1})$$

(7.4.14.)

which as in Equation 7.3.17., letting;

$$U_{N-1} = \Phi_{N-1} + V_{N-1}$$

(7.4.15.)

further reduces to;

$$L_{N,N-2} = \min_{y_{N-2}} E(x^T_{N-1}(U_{N-1})x_{N-1} + y^T_{N-2}B_{N-2}y_{N-2}) + \xi_{N-1}$$

(7.4.16.)

Now this is the same form as Equation 7.4.1., except for the addition of the term ξ_{N-1} which is independent of y_{N-2}. Thus the solution to this problem is immediate;

$$\overset{*}{y}_{N-2} = \Upsilon_{N-1}\hat{x}_{0\ N-2,N-2}$$

(7.4.17.)

$$L_{N,N-2} = E(x^T_{N-2}V_{N-2}x_{N-2}) + \xi_{N-2}$$

(7.4.18.)

where

$$\Upsilon_{N-2} = -(\theta^T_{N-1,N-2}U_{N-1}\theta_{N-1,N-2} + B_{N-2})^{-1}\theta^T_{N-1,N-2}U_{N-1}\phi_{N-1,N-2}$$

(7.4.19.)

$$V_{N-2} = \phi^T_{N-1,N-2}U_{N-1}(\phi_{N-1,N-2} + \theta_{N-1,N-2}\Upsilon_{N-2})$$

(7.4.20.)

and

$$\xi_{N-2} = \xi_{N-1} + \mathrm{TRACE}(\phi^T_{N-1,N-2}U_{N-1}\theta_{N-1,N-2}\Upsilon_{N-2}C_{0\ N-2,N-2}+U_{N-1}R_{N-2})$$

(7.4.21.)

7.4.3. SOLUTION OF THE N-PERIOD STOCHASTIC CONTROL PROBLEM

Noting the form of the solutions to the one and two period problem it is easy to postulate the solution to a j-period problem. Thus, it is assumed that the solution of the J-period problem is;

$$\overset{*}{y}_{N-J} = \Psi_{N-J}\hat{x}_{0\ N-J,N-J}$$

(7.4.22.)

$$L_{N,N-J} = E(x^T_{N-J}V_{N-J}x_{N-J}) + \xi_{N-J}$$

(7.4.23.)

where

$$\Psi_{N-J} = -(\Theta^T_{N-J+1,N-J}U_{N-J+1}\Theta_{N-J+1,N-J} + \Theta_{N-J})^{-1}\Theta^T_{N-J+1,N-J}U_{N-J+1}\phi_{N-J+1,N-J}$$

(7.4.24.)

$$U_{N-J+1} = \Omega_{N-J+1} + V_{N-J+1}$$

(7.4.25.)

$$V_{N-J} = \phi^T_{N-J+1,N-J}U_{N-J+1}(\phi_{N-J+1,N-J} + \Theta_{N-J+1,N-J}\Psi_{N-J})$$

(7.4.26.)

$$\xi_{N-J} = \xi_{N-J+1} + \mathrm{TRACE}(\phi^T_{N-J+1,N-J}U_{N-J+1}\Theta_{N-J+1,N-J}\Psi_{N-J}C_{0N-J,N-J}$$

$$+ U_{N-J+1}R_{N-J})$$

(7.4.27.)

The J+1-period problem consists of finding the control sequence $\overset{*}{y}_{N-J-1}, \overset{*}{y}_{N-J}, \ldots, \overset{*}{y}_{N-1}$ that minimises;

$$\pounds_{N,N-J-1} = E\sum_{i=N-J}^{N}(x^T_i\Omega_i x_i + y^T_{i-1}\Theta_{i-1}y_{i-1})$$

Now the control vectors $\overset{*}{y}_{N-J}, \ldots, \overset{*}{y}_{N-1}$ do not affect the stock vector at time N-J, and since it has been <u>assumed</u> that the solution to the J-period problem is true, then the task is simply to find $\overset{*}{y}_{N-J-1}$ that minimises;

$$\pounds_{N,N-J-1} = E(x^T_{N-J}\Omega_{N-J}x_{N-J} + y^T_{N-J-1}\Theta_{N-J-1}y_{N-J-1} + x^T_{N-J}V_{N-J}x_{N-J} + \xi_{N-J})$$

(7.4.28.)

Compatable with Equation 7.4.25.;

$$U_{N-J} = \Phi_{N-J} + V_{N-J} ,$$

(7.4.29.)

Equation 7.4.28. reduces to;

$$£_{N,N-J-1} = E(x^T_{N-J}U_{N-J}x_{N-J} + y^T_{N-J-1}B_{N-J-1}y_{N-J-1}) + \xi_{N-J}$$

(7.4.30.)

Strictly, we should now go through the minimising process as in Section 7.4.1., but as this is so similar, it is considered enough to note the similarity between Equation 7.4.1. and Equation 7.4.30. and then state the solution to the J+1-period problem as;

$$\overset{*}{y}_{N-J-1} = \Upsilon_{N-J-1}\hat{x}_{0\ N-J-1,N-J-1}$$

$$L_{N,N-J-1} = E(x^T_{N-J-1}V_{N-J-1}x_{N-J-1}) + \xi_{N-J-1}$$

where

$$\Upsilon_{N-J-1} = -(\theta^T_{N-J,N-J-1}U_{N-J}\theta_{N-J,N-J-1} + B_{N-J-1})\theta^T_{N-J,N-J-1}U_{N-J}\phi_{N-J,N-J-1}$$

$$U_{N-J} = \Phi_{N-J} + V_{N-J}$$

$$V_{N-J-1} = \phi^T_{N-J,N-J-1}U_{N-J}(\phi^T_{N-J,N-J-1} + \theta_{N-J,N-J-1}\Upsilon_{N-J-1})$$

$$\xi_{N-J-1} = \xi_{N-J} + \text{TRACE}(\phi^T_{N-J,N-J-1}U_{N-J}\theta_{N-J,N-J-1}\Upsilon_{N-J-1}C_{0N-J-1,N-J-1}$$

$$+ U_{N-J}R_{N-J-1})$$

So, given that the solution is true for a J-period problem, it is true for a J+1-period problem and noting that it is also true for J=1, by the principle of mathematical induction, the solution holds for J=1,2,...N. This completes the stochastic control problem.

In the next section a simple example is given so that the theory of the last two sections can be related to a real situation.

7.5. AN EXAMPLE IN STOCHASTIC CONTROL THEORY

Consider a three-period stochastic control problem where there are no costs associated with control, that is;

$$£_3 = E\sum_{i=1}^{3}(\overset{*}{x}_i - x_i)^T \mathbb{D}_i(\overset{*}{x}_i - x_i)$$

where

$$x_{k+1} = \phi_{k+1,k}x_k + \theta_{k+1,k}y_k + dx_k$$

$$z_{k+1} = \psi_{k+1}x_{k+1} + dz_{k+1}$$

and where the objective is to find controls $\overset{*}{y}_0, \overset{*}{y}_1, \overset{*}{y}_2$ that minimise $£_3$. Using the work of Section 2 of this chapter, the above problem may be reduced to;

$$£_3' = E\sum_{i=1}^{3} x_i'^T \mathbb{D}_i x_i$$

where

$$x'_{k+1} = \phi'_{k+1,k}x'_k + \theta'_{k+1,k}y_k + dx'_k$$

$$z_{k+1} = \psi'_{k+1}x'_{k+1} + dz_{k+1}$$

The definitions of the terms with an apostrophe carry over from Section 2. The solution can be written down immediately as;

$$\overset{*}{y}_k = \Upsilon_k \hat{x}'_{0\ k,k} \qquad k = 0,1,2.$$

where Υ_k simplifies to;

$$\Upsilon_k = \begin{bmatrix} -\theta^{-1}_{k+1,k}\phi_{k+1,k} & \theta^{-1}_{k+1,k}\overset{*}{\phi}_{k+1,k} \end{bmatrix}$$

Also

$$\hat{x}'_{0k+1,k} = \begin{bmatrix} \overset{*}{x}_{k+1} \\ \overset{*}{x}_{k+1} \end{bmatrix} \qquad \hat{x}'_{0k,k} = \begin{bmatrix} \overset{*}{x}_k + B_k(z_k - \overset{*}{x}_k) \\ \overset{*}{x}_k \end{bmatrix}$$

and therefore

$$x'_{k+1} = \begin{bmatrix} \overset{*}{x}_{k+1} + dx_k \\ \overset{*}{x}_{k+1} \end{bmatrix} \qquad \tilde{x}_{0\ k+1,k} = \begin{bmatrix} dx_k \\ 0 \end{bmatrix}$$

It is simple to verify that $V_k = 0, \forall k$ and so the total cost is;

$$L_3 = \sum_{i=0}^{2} \mathrm{TRACE}(\Phi_{i+1,i} U_{i+1} B_{i+1,i} \Upsilon_i C_{0\ i,i} + U_{i+1} R_i)$$

which after some algebraic manipulation becomes;

$$L_3 = \sum_{i=0}^{2} \mathrm{TRACE}\ \mathbb{D}_{i+1}\ C_{0\ i+1,i}$$

A noteworthy characteristic of this type of problem is that the optimal control filter Υ_k and the optimal filtered estimate $\hat{x}_{0,k,k}$ are determined independently. Thus, the control filter can, and in fact must, be determined for all times of the planning period. This is because in most problems its form is given by applying the recursive Equations; 7.4.24.,25.,26. backward in time.

In the above simplistic example the optimal control filter can be determined without reference to the planning period. This occurs because the expected error is zero. In cases where the B_k are non-zero or if other constraints were present, as in Bartholomew (1975), the optimal control filter can only be obtained from the recursive equations starting at the end of the planning period and working backwards in time.

We have now developed models for the measurement and prediction of wastage, supply and demand forecasting, and the control of manpower systems. In the next chapter we bring these together and discuss their application to an industrial manpower system.

CHAPTER EIGHT

THE DUNLOP MANPOWER SYSTEM AT WORK

8.1. INTRODUCTION

It is the aim of this chapter to review the current state of manpower planning in Dunlop U.K.T.G.. The following section begins with a presentation of some of the results obtained by using the models developed in the previous chapters. Firstly, wastage estimation is considered; here the problem is to estimate the time varying elements of the survivor curve so that the estimated function may be applied to the current grade dependent length of service distribution in order to ascertain the probable stock remaining at the end of the time period. The actual data exhibited refers to operators and their propensity to leave during the first three months of service. The three methods proposed in Chapter 4 are compared with what might be called the classical, sample mean - sample variance, approach. On this data the results of the new approach are very much better as measured by two different techniques.

After the wastage rates and numbers have been estimated by the above methods, one is left with a stock that is eligible for promotion. In Chapter 5 the Dirichlet-Multinomial distribution was proposed as a suitable distribution to govern the grade transitions. In Section 3 this method is employed to estimate the promotional transition over Dunlop data for the years 1972-1977 inclusive.

Naturally any industry based thesis must be judged not only on its academic success but on the benefits it offers to the sponsoring company. With this in mind Section 4 deals with current and potential uses of the proposed manpower models to Dunlop U.K.T.G.. It would be misleading to suggest that the complex models expounded in the main chapters of the thesis have been understood by all management and that they use them unreservedly, that is not the case, however the simpler aspects, such as an accessible up-to-date data base, extracted programmes for tabulation and graphical representation of data and methods of wastage estimation have gained considerable approval. The reasons for this situation and the resistance to accepting the complete 'package' are further examined in Section 4.

The manpower models, although specifically designed for the Dunlop contextual environment, will be applicable to other organisations for manpower planning and because of the generality of the formulated stochastic dynamic models, will have applications outside the manpower planning field. A variety of possible applications will be discussed in Section 5.

In this way it is hoped to trace the progress of the project, the results and implications as related to Dunlop U.K.T.G. and also to discuss the relevance of the models which have been developed in the more general context of manpower and industrial planning.

8.2. RESULTS OF THE WASTAGE MODELS ON DUNLOP ENGINEERING DATA 1971-77

Every manpower supply model must consider wastage. It is therefore of great importance that the best estimation procedure available should be used to estimate wastage, for the accuracy of any supply forecast will be ultimately dependent upon it. It is now widely accepted that one of the prominent, if not most prominent, correlate of wastage is length of service. It is clear then that wastage can only be discussed sensibly with reference to length of service. Forbes (1971a) has given a comprehensive account of methods of estimating the survivor function; amongst other methods he deals with census and cohort analysis. Both of these approaches have their advantages and disadvantages. In order to overcome the disadvantages of these static models, the problem of wastage estimation was reformulated into a time dependent stochastic context. The new dynamic estimation procedure was established in Chapters 3 and 4, where three methods of estimating the covariance structure of the optimal predictor were proposed.

These three methods and the time invariant, sample mean - sample variance, method have been compared on data from Dunlop Engineering Group. The data comprised 28 observations of the quarterly wastage rate over the period 1971-1977. The comparison was made using two different statistics; The mean sum of squared errors and the mean likelihood of obtaining the observation from its estimated distribution. The former is the more usual test statistic whereas the latter affords the inclusion of the variance of the estimate which is most useful for decision making. Graphical results of the relative efficiencies under the test statistics are presented in Figures 8.2.1. and 8.2.2., and a summary of the results are given in Table 8.2.1., where the weighted posterior probability estimator is scaled to 100. Further details of the calculations and results can be found in Appendix C. Section 2.

FIGURE 8.2.1. GRAPHICAL REPRESENTATION OF THE RELATIVE EFFICIENCIES OF THE FOUR ESTIMATORS ON DUNLOP DATA USING THE MEAN SUM OF SQUARED ERRORS STATISTIC

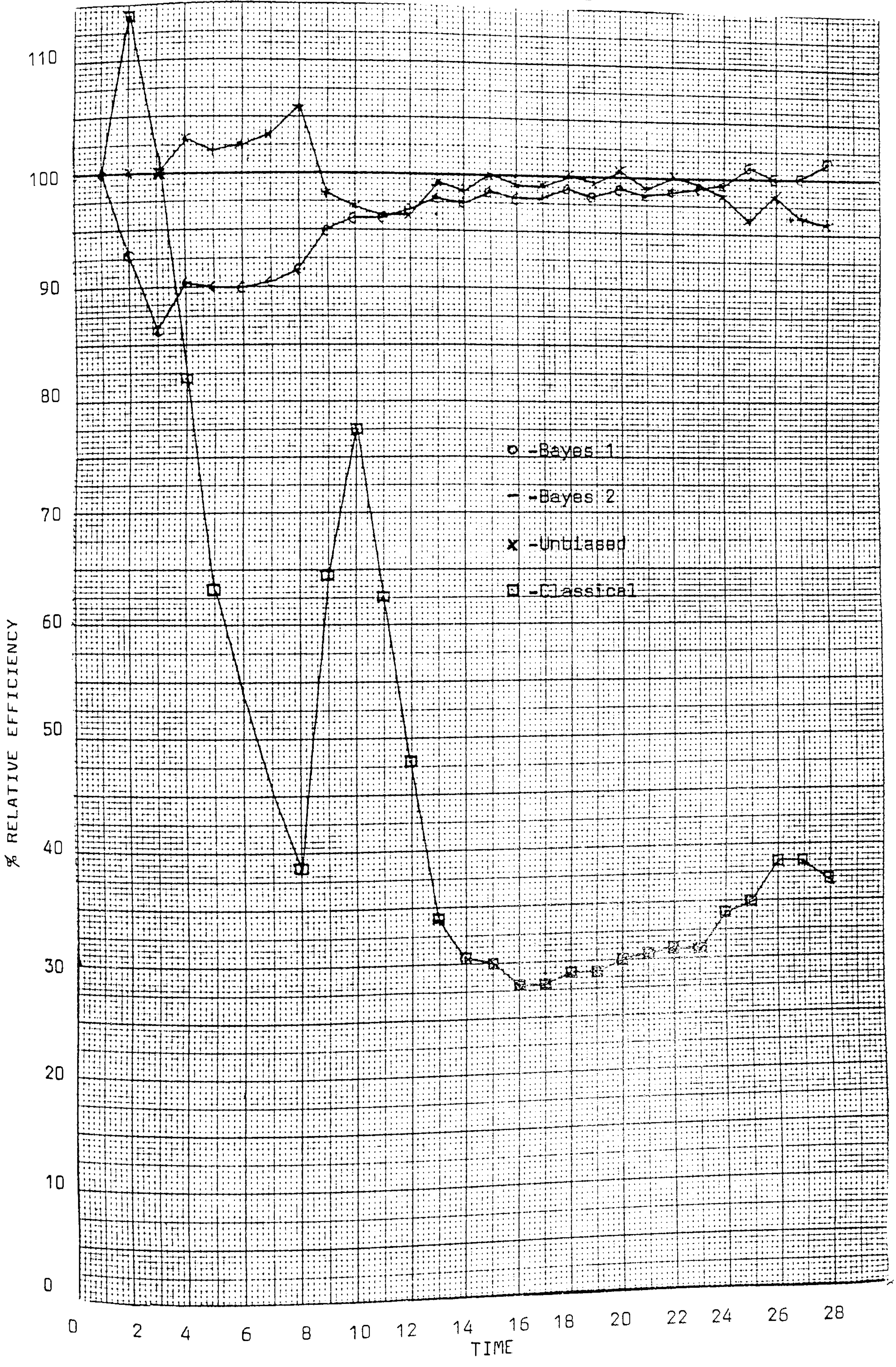


FIGURE 8.2.2. GRAPHICAL REPRESENTATION OF THE RELATIVE EFFICIENCIES OF THE FOUR ESTIMATORS ON DUNLOP DATA USING THE MEAN LIKELIHOOD STATISTIC

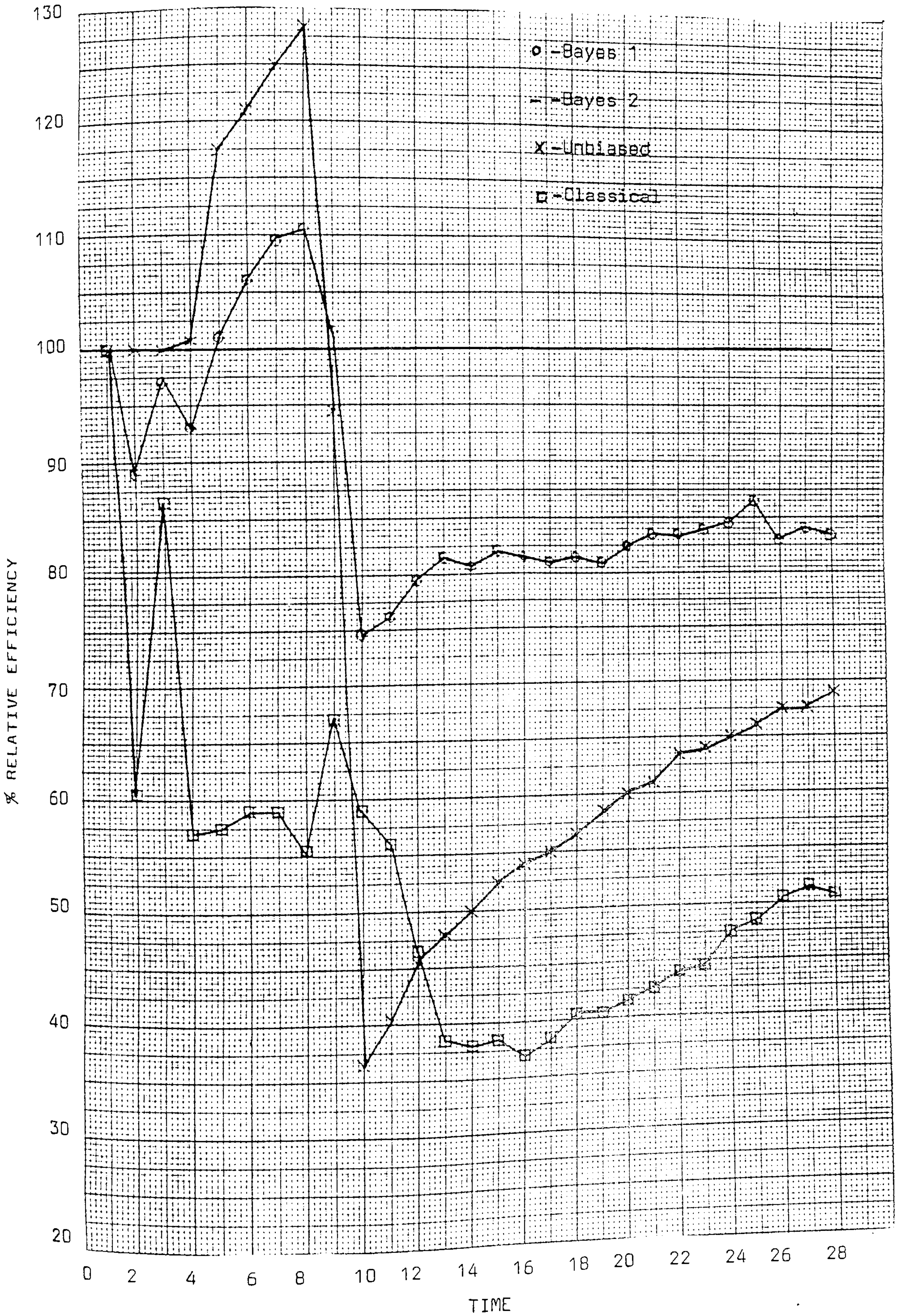


TABLE 8.2.1.

Comparison of the relative efficiences of the four estimators over the whole Dunlop data series:

	UNBIASED	BAYES 1	BAYES 2	CLASSICAL
Mean sum of squared errors	96.1	101.4	100.0	37.1
Mean likelihood	68.8	83.3	100.0	51.0

The graphs give a clear indication of the necessity to use the time variant estimation methods on data which is thought might be time dependent. Comparison by the mean sum of squared errors shows little difference between the proposed estimators, however the improvement on the classical estimator is enormous. On this data, the second test statistic separates the estimators still further and reveals the Bayesian-type estimators to have the greater power in resolving the observation distribution. Further, it is noted that little loss is incurred if the data is time invariant by using the time variant method and much is to be gained computationally through the smaller data storage requirements of the dynamic method which will obviously affect maintenance costs of the system.

8.3. RESULTS OF USING THE DIRICHLET-MULTINOMIAL DISTRIBUTION TO PREDICT PROMOTIONS ON DUNLOP U.K.T.G. DATA 1972-1977

Commencing with the stock vector on January 1st, 1973 knowledge of the observed length of service dependent wastage rates and the promotions occuring in the preceeding year, estimation of future transitions could begin. Initially, the grade-length of service estimated wastage rates were applied to the stock vector yielding an estimate of the number of employees who would remain in the system at least until January 1st, 1974. The surviving graded stock were then considered eligible for transition. The probability transition rate for 1973-1974 was now estimated by the methods discussed in Chapter 5. The observation of what actually occured during this period was then recorded. After the addition of all surviving recruits for that year, the stock vector for January 1st, 1974 was obtained and the recursive estimation process continued. Details of the results are presented in Appendix E. whilst a summary of the results follows in Tables 8.3.1. and 8.3.2..

TABLE 8.3.1.

The actual and estimated leavers (staff) from Dunlop U.K.T.G. 1972-1977:

GRADE	1	2	3	4	5	6	TOTAL
ACTUAL 72-73	57	143	52	16	7	5	280
ESTIMATE 73-74	90	293	71	12	3	5	474
ACTUAL 73-74	92	286	72	13	3	5	471
ESTIMATE 74-75	117	462	133	40	13	12	777
ACTUAL 74-75	108	458	144	37	14	13	774
ESTIMATE 75-76	77	392	96	36	11	9	621
ACTUAL 75-76	83	400	92	35	12	10	632
ESTIMATE 76-77	76	357	150	39	10	2	634
ACTUAL 76-77	72	386	148	35	11	2	654

Due to the intermittent nature of leaving with those groups having more than eight years service, predictions here were extremely varied, but as the numbers leaving were small, the results when aggregated to the grade level are most satisfactory.

The second table relates to the estimated stock distribution after account has been taken of the estimated wastage and promotional transitions but before the addition of the surviving recruits during the relevant period.

TABLE 8.3.2.

Actual and estimated stock distribution for the year end on Dunlop U.K.T.G. data 1972-1977.

GRADE	1	2	3	4	5	6	TOTAL
ACTUAL 72-73	299	1882	834	139	51	23	3228
ESTIMATE 73-74	315	1890	874	150	59	28	3316
ACTUAL 73-74	253	1894	936	151	56	29	3319
ESTIMATE 74-75	204	1709	911	129	51	24	3028
ACTUAL 74-75	234	1691	902	125	55	24	3031
ESTIMATE 75-76	187	1484	891	106	46	20	2734
ACTUAL 75-76	186	1494	874	109	42	18	2723
ESTIMATE 76-77	164	1516	823	99	45	25	2672
ACTUAL 76-77	156	1565	759	105	47	24	2652

Once again the total stock size is predicted with great accuracy ($\pm 1\%$) and the grade sizes are of usable precision. The actual and estimated transitions are given in Appendix E. Here, as may be expected, the promotional flows are more varied but when coupled with the variance estimation given by the Dirichlet-Multinomial method a sensible interpretation of the flows can be made. The method for estimating the parameters and moments of the Dirichlet-Multinomial distribution, which is described in Chapter 5, has proved a most useful tool in the analysis of promotional flows.

8.4. PROJECT BENEFITS TO DUNLOP U.K.T.G.

At the beginning of the project U.K.T.G. had a manual based record system, access to which was effective only on an employee name basis. The collection of aggregate data by any other attribute was almost impossible due to the excessive workload that would be involved. The same was true for any form of cross tabulation between work locations, job type or factory. In order to obtain usable data for the input into the models proposed in this thesis it was necessary to construct an accurate, up-to-date data base. This was achieved at minimal cost to the centre and, perhaps more importantly, minimal disruption to the smaller operating units during both the initial stages and the necessary procedures for updating the system. This computerised information system was established as an interlaced element of the salary suite and not merely an appendage of it. The value of the total system was thus enhanced; now information on both salary, job evaluation schemes and manpower data could be cross referenced. The elementary programmes written for data extraction have been extensively used by Dunlop U.K.T.G. giving immediate results.

Another feature of this system is the small amount of effort that is required for the maintenance of the file after the initial implementation stage. This is invaluable to the company as the usable data base is now expanding and no further information is being lost through either the absence of data or its ineffective storage.

Not all of the models are being used to their full advantage and it is necessary to examine some of the reasons why this is the case. Firstly, in industry information is not neutral, in the sense that it carries socio-political implications. The proposed manpower planning system may have far reaching affects on the organisation and lead to changes being made. These changes may influence many areas; recruitment and redundancy policies, manning levels, skill distribution, career development patterns etc.. All of which are politically and emotionally charged. Any proposed change may meet with resistance especially when its implications are so extensive.

Increased information on the manpower within an organisation could lead to industrial conflict. The disclosure of information was of particular concern to the executive management of Dunlop U.K.T.G.; should knowledge be given of plans to reduce the labour force, the probability of promotion between various grades or the closure of a factory? Present leglislation on the disclosure of information increases the difficulty of this problem. It should be realised, therefore, that the information given by the manpower models, however accurate, will have socio-political implications for the organisation and its work groups.

The above situation is true of many organisations. Dunlop, in common with other tyre companies, has one more important feature that it is thought to have influenced the acceptance of the models. This is the desperate over-capacity in the tyre industry and the accumulated over manning in certain areas. This results in a situation of 'crisis' management such as the drowning man who thinks that he sees land on the horizon and starts swimming for it although there <u>might</u> be an island just behind him to the left or right. Time to thoroughly analyse a situation in order to determine the best long-term strategies seldom seems available, even though the short-term policies adopted are rarely optimal. Nevertheless 'fire-fighting' seems to be the usual approach and it is difficult to change management attitudes quickly on decision making approaches.

The initial terms of reference for the project provided for three different planning horizons to the manpower methodology. Within these boundaries, it could be said that over the short-term (1-2 years) considerable success has been achieved with the one year ahead forecasts of the new models e.g. wastage estimation and promotional transition. Consequently, the confidence in these techniques is growing due to the realised benefits of their adoption. The attempt to justify a long term (10 years plus) strategy over a short time period is a problem of a much higher magnitude and it is unlikely that acceptance of the long term strategy will increase significantly beyond the 'scenario' level for some time to come. With the medium-term planning horizon, the situation is a little more optimistic. The Manpower Services Commission and the Industrial Training Boards are becoming more aware of the rewards of medium-term manpower planning. Their educational influence on industry through the provision of courses and discussions at all levels of management, coupled with the inclusion of manpower planning as a pre-requisite for levy exemption for some of the Boards, has noticeably raised the level of consciousness as to the benefits of using manpower planning techniques.

As management becomes more informed it is believed that the techniques developed in this work, or similar techniques, will become widely used as the potential gains are realised. However, manpower modelling must be seen as a useful <u>tool</u> for effective manpower management and decision making, not a substitute for it.

8.5. FURTHER APPLICATIONS OF THE STOCHASTIC PLANNING MODELS

The models described in the thesis have been developed specifically to satisfy the manpower problems of Dunlop U.K.T.G.. The U.K.T.G. acts as a administrative centre of the U.K. tyre operation and has direct control over the four main new-tyre factories, the retreading factories, distribution depots and the sales organisation. In this respect the tyre organisation is highly centralised but has many operating locations. The Group is easily seen to be a complex system having a high degree of hetrogeneity in job types. The manpower system therefore needed to be flexible enough to operate at the local level, with the minimum of intervention, whilst providing the centre with meaningful data both in quantity and quality, for the development of manpower strategies. In order to achieve this objective the extra data requirements were restricted to eight items which were specified on one form; the 'Staff Changes Form'. One of the strengths of the system lies in the timely and accurate completion of this document, used for both salary and manpower information. For manpower applications the model has been seen to work well on a small amount of readily available data.

Being centralised is not, of course, a constraint on the use of the models, but mentioned only as a consideration when information is scarce and resentment exists against the provision of data to the centre. The advantages are not felt solely by the central unit, as now valuable manpower statistics may be disseminated to the satellite units providing useful background to discussions on a variety of topics including;

1) Leaving rates and reasons for leaving of salesmen in the South East as opposed to other areas.
2) Foreman to operative ratio in the manufacturing units.
3) Compilation of possible candidates in other locations for succession planning.

This type of information and the benefits to be derived from them should not be limited to use in one organisation. The manpower system could be adapted for use in a number of organisations where similar problems already exist, such as;

1) Insufficient data on manpower already in existence where the use of classical statistical methods is excluded.
2) Concern for the amount of time and money to be expended on the provision of a manpower planning system.
3) A central unit with smaller units distributed throughout a region/country.

The formulation of the stochastic estimation model, as stated in Chapter 3, was very general and its association with the control strategies of Chapter 7 means that the model will be applicable to many situations where control of knowledge within a dynamic system is required and where observation of the system is coloured by random disturbances or the lack of precision is some measuring process. Applications will include such diverse things as; quality control on manufacturing lines, prediction of mortality rates, market forecasting or more obscure uses such estimating the destruction occuring at a distance from the epicentre of an earthquake in which case the process diffuses through both space and time.

The Dirichlet-Multinomial distribution also has many applications outside organisational manpower planning; wherever unknown flows occur between identifiable units. Some obvious uses would be in the areas of social and labour mobility, consumer brand switching, international export and import analysis and the less obvious area of nuclear particle collision theory.

It is apparent from the few possible applications mentioned in this section that stochastic planning methods have great flexibility and potential uses within many diverse situations which is a clear strength of the model.

8.6. CONCLUDING REMARKS

The content, development and results of the project within Dunlop U.K.T.G. have been fully discussed now and it is worthwhile, in conclusion, to review the underlying philosophy of an industry based I.H.D. scheme.

The foundation of the work lies in the synergy of industrial practicality and academic analysis. As was mentioned in Chapter 1, classical theory may be suitable only where 'well behaved' data is available. It has frequently been a problem, therefore to make the transition between the classical methods and the industrial application. The industrial environment varies greatly from the ideal situation in which academic work is often based. Despite the difference in environment, priorities and attitude the two worlds, academic and industrial, may benefit greatly by constructive communication. Industrial managers are stimulated into thought and action which the pressures of industry may often inhibit. The academic gains a clearer insight of the situations in which proposed models must function and the effect of socio-political influences on the acceptance of methodology.

In some way the project which took place within Dunlop U.K.T.G. made some progress by bringing about this type of communication which, when successful, brings mutual benefits and the foundation upon which to build a closer co-operation.

APPENDIX A

THE HILBERT SPACE

Here the concept of the Lebesgue and Hilbert Spaces are formally defined and some of their properties, which are usefully employed in the main chapters, derived. Both of these mathematical spaces are not new and can be found in basic texts on algebra, (Halmos 1951), however, the mode of presentation and proofs as applied to the solution of the General Problem are original to the present work.

In Appendix A.1.1. the structure of the Hilbert Space is built up from first principles. Starting with basic algebra, the definition of a ring is given and by adding further constraints, the structure of an σ-algebra is obtained. The concept of a measure is then defined and, in particular, the probability measure is introduced. From these definitions a formal definition of a random variable can be established.

By the introduction of addition and multiplication as mathematical operators on a vector space, random vectors are seen to be elements of the space of all mappings from $\Omega \mapsto \mathbb{R}^N$ that are measurable with respect to a σ-field, $\mathbb{R}$ of Ω. The concepts of length and distance between vectors of Euclidean space are formulaised in the topographical structure of the norm and metric. Following this is given the definition of the Cauchy sequence and we are able to define a Banach Space. This leads to the proof that the Lebesgue Space is a Banach Space. One last structural concept that of an 'angle' between vectors in Euclidean Space, and its formal equivalent of the inner product enables the definition of the Hilbert Space. It is then easily shown that the Lebesgue Space is indeed a Hilbert Space.

In Section 2. of this Appendix, the important properties of the Hilbert Space that accounts for its natural choice as a space in which to solve the General Problem are explained. Two important theorems are proved;

i) Within a closed convex subspace of an Hilbert Space there is an element with smallest norm

ii) Every vector in an Hilbert Space may be written as an unique element of a closed subspace and one in its orthogonal complement

Section A.1.1. now follows beginning with a set of elementary definitions.

A.1.1. DEFINITION OF THE HILBERT SPACE

We begin by defining a probability space and work towards the definition of a random vector. This work is standard to any text on Probability Theory.

Definition: RING

A non-empty class $\mathbb{R}$ of subsets of a set Ω is called a ring if;

R(1) $A_i, A_j \in \mathbb{R} \Rightarrow A_i \cap A_j \in \mathbb{R}$

R(2) $A_i, A_j \in \mathbb{R} \Rightarrow A_i \setminus A_j \in \mathbb{R}$

R(3) $A_i, A_j \in \mathbb{R} \Rightarrow A_i \cup A_j \in \mathbb{R}$

Definition: σ -ALGEBRA

$\mathbb{R}$ is a σ-algebra if it is a ring and if;

R(4) $\Omega \in R$

R(5) $A_n \in \mathbb{R}, \forall_n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathbb{R}$

It is noted that R(2,4,5,) are sufficient for $\mathbb{R}$ to be a σ-algebra. The double $(\Omega, \mathbb{R})$, where $\mathbb{R}$ is a σ-algebra over Ω is called a measureable space. The real numbers $\mathbb{R}$ extended by the symbols $+\infty, -\infty$ (not real numbers) will be denoted $\mathbb{R}^*$.

Definition: MEASURE

A measure on a measurable space $(\Omega, \mathbb{R})$ is a function μ that maps $\mathbb{R}$ into $\mathbb{R}^*$ such that;

M(1) $\mu(\emptyset) = 0$

M(2) $\mu(A) \geq 0 \qquad \forall A \in \mathbb{R}$

M(3) $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} (A_n)$ whenever $\{A_n\}$ is a sequence of mutually disjoint sets in $\mathbb{R}$.

Condition M(3) is called 'countably additive'.

Definition: PROBABILITY MEASURE

A measure P is a probability measure if;

P(1) $P(\Omega) = 1$

The triple $(\Omega, \mathbb{R}, P)$, where $\mathbb{R}$ is a σ-algebra over the non-empty set Ω and P is a probability measure is called a Probability Space, where no confusion arises we may talk of a probability space Ω.

Definition: RANDOM VARIABLE

Let $(\Omega, \mathbb{R}, P)$ be a probability space and let $X : \Omega \longmapsto \mathbb{R}$. Then if $\forall x_0 \in \mathbb{R}$, the set;

$$\{\omega \in \Omega : X(\omega) < x_0\} \in \mathbb{R}$$

is measurable then $X(\omega)$ is a random variable or measurable function.

A.1.2. VECTOR SPACE

A vector space consists of a set V, whose elements are called vectors, a scalar field F and some mathematical structure. We will only be interested in the scalar field, the real numbers $\mathbb{R}$, and thus confine ourselves to Real Vector Spaces.

Definition: VECTOR SPACE

The 4-tuple $(V,F,+,.)$, less strictly denoted by its underlying set V, is a vector space if the following hold:

1) A mapping + of $V \times V \longmapsto V$, called addition.

$\forall$ vectors $v_1, v_2 \in V$, $\exists$ a unique $v_3 \in V$ such that;

$$+(v_1,v_2) = v_3 \quad \text{or more usually} \quad v_1+v_2 = v_3$$

V(1) $v_1 + v_2 = v_2 + v_1 \qquad \forall\, v_1,v_2 \in V$

V(2) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3 \qquad \forall\, v_1,v_2,v_3 \in V$

V(3) $\exists$ a unique element in V denoted 0_v, such that;

$$0_v + v = v \qquad \forall v \in V$$

V(4) Associated with each $v \in V$, $\exists$ a unique element denoted $(-v)$, such that;

$$v + (-v) = 0$$

2) A mapping $\cdot$ of $F \times V \longmapsto V$, called scalar multiplication.

$\forall$ scalars $f \in F$ and $v_1 \in V$ $\exists$ a unique element $v_2 \in V$ such that;

$$\cdot(f, v_1) = v_2 \quad \text{or more usually} \quad f \cdot v_1 = v_2$$

If no ambiguity exists then the operation . can be dropped and we write;

. (f, v) as fv.

V(5) $f_1(f_2 v) = (f_1 f_2)v$ $\qquad \forall\, f_1, f_2 \in F,\ v \in V$

V(6) $1_f v = v$ $\qquad \forall\ v \in V$

V(7) $(f_1 + f_2)v = f_1 v + f_2 v$ $\qquad \forall\, f_1 f_2 \in F,\ v \in V$

V(8) $f(v_1 + v_2) = fv_1 + fv_2$ $\qquad \forall\, f \in F,\ v_1, v_2 \in V$

An immediate consequence of the definition is;

$$0_f v = 0_v \qquad \forall v \in V$$

The vector space we will use later is the space of all mappings from $\Omega \longmapsto \mathbb{R}^n$, that are measurable with respect to a σ-field $\mathcal{R}$ of Ω. Thus the elements of this space are to be called <u>random vectors</u>.

A.1.3. NORMED VECTOR SPACES

Given a Vector Space V we will find it useful to be able to compare vectors. In order to do this we introduce three types of mapping from the Vector Space V, (or its product space $V \times V$) to the real numbers $\mathbb{R}$.

We consider firstly the concept of the metric space.

Definition: <u>METRIC SPACE</u>

Let X be a set and $\rho : X \times X \longmapsto \mathbb{R}^+ = [0, +\infty)$ a map such that;

$\forall x_1, x_2, x_3$ in X

MS(1) $\rho(x_1, x_2) = 0 \iff x_1 = x_2$

MS(2) $\rho(x_1, x_2) = \rho(x_2, x_1)$

MS(3) $\rho(x_1, x_2) \leq \rho(x_1, x_3) + \rho(x_3, x_2)$ $\qquad$ Triangle inequality

Then ρ is called a <u>metric</u> of X and the double (X, ρ) a <u>metric space</u>.

The structure of the metric space is related to our concept of distance in Euclidean geometry. For this reason we called the metric on $\mathbb{R}^n$

defined by $\rho(x_1,x_2) = \left[\sum_{i=1}^{n} (x_{1i} - x_{2i})^2\right]^{\frac{1}{2}}$ the Euclidean or Usual Metric.

The class of metrics is vast so we will not discuss them here, but will examine a special sub-class later in the chapter.

We now progress further in our development of the mathematical structure by defining another real valued function which is not only similar to the idea of Euclidean length but also generates a suitable distance function (metric).

Definition: NORMED VECTOR SPACE

Let V be a vector space and $\| \; \| : V \longmapsto \mathbb{R}^+$ a map such that;

$$\forall \lambda \in \mathbb{R}, v, v_1, v_2 \in V$$

N(1) $\|v\| = 0 \iff v = 0_V$

N(2) $\|\lambda v\| = |\lambda| \; \|v\|$

N(3) $\|v_1+v_2\| \qquad \|v_1\| + \|v_2\|$

The number $\|v\|$ is called the norm or sometimes the length of v.

The double $(V, \| \; \|)$ is called a normed vector space (n.v.s.).

THEOREM A.1.1.

If $(V, \| \; \|)$ is a normed vector space then;

$$(v_1,v_2) = \|v_1-v_2\|$$

defines a metric on V called the induced metric. (See footnote) ▼

PROOF

MS(1) $\rho(x_1x_2) = 0 \overset{\Delta\rho}{\iff} \|x_1-x_2\| = 0 \overset{N1}{\iff} x_1-x_2 = 0 \iff x=y$

MS(2) $\rho(x_1,x_2) \overset{\Delta\rho}{=} \|x_1-x_2\| \overset{N2}{=} \|x_2-x_1\| \overset{\Delta\rho}{=} \rho(x_2,x_1)$

MS(3) $\rho(x_1,x_2) = \|x_1-x_2\| = \|(x_1-x_3) + (x_3 - x_2)\|$

$\overset{N3}{\leq} \|x_1-x_3\| + \|x_3-x_2\|$

$\overset{\Delta\rho}{=} \rho(x_1,x_3) + \rho(x_3,x_2)$ ▪

Normed Vector Space: This means that every n.v.s. can be thought of as a metric space.

We now state, without proof, that all norms on $\mathbb{R}^n$ are equivalent - this, for our purposes, means they have similar topographical properties.

Some familar norms on $\mathbb{R}^n$ are;

$$\sqrt{\sum_{i=1}^{n} x_1^2} \;,\quad \sum_{i=1}^{n} |x_i| \;,\quad \max_{1\leq i\leq n} |x_i| \quad \text{and} \quad \left[\int_0^1 |x(t)|^p \, dt\right]^{1/p}$$

We shall show the following map is a well defined norm and use it extensively later;

$$\|x\|_p = \left[\int_{\Omega} |x(w)|^p \mathcal{P}(dw)\right]^{1/p} < \infty \qquad p > 1$$

where $(\Omega, R, \mathcal{P})$ is a probability space and $x \in L_p$ the space of all measurable mappings from Ω to $\mathbb{R}$.

A.1.4. <u>BANACH SPACES</u>

Another property that is deemed essential to our mathematical space is that sequences $\{x_n\}$ that converge to some element x, do so within the vector space, i.e. $x \in V$, more precisely we define a <u>Cauchy</u> sequence.

A sequence $\{x_n\}$ in a metric space (X, ρ) is Cauchy, if given $\varepsilon > 0$, there exists an $n_0(\varepsilon)$ such that;

$$\rho(x_m, x_n) < \varepsilon \quad \text{whenever } m,n \geq n_0(\varepsilon)$$

The metric space (X, ρ) is <u>complete</u> if every Cauchy sequence converges to a point within X.

A <u>Banach Space</u> is a normed vector space that when regarded as a metric space (though induced metric) is complete.

We now consider a special class of spaces that are most important to the development of a forecasting system. These are the Lebesgue Spaces and are closely related to the set of random variables that have finite moments. We will define an appropriate norm and go on to prove that the Lebesgue Spaces are complete and thus are Banach Spaces.

Definition: <u>LEBESGUE SPACE $L_p(\Omega, \mathcal{R}, \mu)$</u>

The set $\mathcal{L}p\ (\Omega, \mathcal{R}, \mu)$ is defined as the set of measurable functions X on the measure space $(\Omega, \mathcal{R}, \mu)$ such that $|X|^p$ for $p \geq 1$ is integrable and finite, i.e.:

$$\int_\Omega |X(\omega)|^p d\mu(\omega) < \infty \qquad 1 \leq p < \infty$$

We will denote the set $\mathcal{L}p\ (\Omega, \mathcal{R}, \mu)$ by $\mathcal{L}p$, where no confusion arises to which measure space is under consideration.

We define for $p \geq 1$, the p norm on $\mathcal{L}p$ given by;

$$\|X\|_p = \left(\int_\Omega |X(\omega)|^p d\mu(\omega) \right)^{1/p}$$

As $\|X-Y\|_p = 0 \iff x=y$ except possibly on a set of measure zero it is necessary to define a null set $\mathcal{N}$;

$$\mathcal{N} = \{X : X = 0 \ \text{ a.e.}\}$$

where a.e. means 'almost everywhere'.

We now give the Lebesgue space $L_p\{\Omega, \mathcal{R}, \mu\}$ as the quotient $\mathcal{L}_p/\mathcal{N}$, where there is no ambiguity the space is represented by L_p only.

The measurable functions X of L_p are thus equivalent classes of those of $\mathcal{L}_p$. The distinction is unimportant for our purposes and we will speak of functions of L_p when we formally mean the representation of that function's equivalence class in $\mathcal{L}_p$.

The space L_p is easily shown to be linear. The properties of the norm follow immediately, by noticing that N(3) (the triangle inequality for norms) is given by directly applying the famous Minkowski Inequality. L_p is thus quickly verified as being a normed linear space. To show L_p is a Banach Space we must show it is complete. This is more difficult, and is given in the following theorem:

THEOREM A.1.2.

The space L_p is complete and, hence, is a Banach Space. ▼

PROOF

Let $\{x_n\}$ be a Cauchy sequence in L_p; then for each k there exists an N_k, such that; if n, m, $\geq N_k$, we have

$$\|x_n - x_m\|_p < 2^{-k}$$

There is no loss in generality if it is assumed that $N_k < N_{k+1}$ $\forall k$.

Thus

$$\| x_{N_k} - x_{N_{k+1}} \|_p < 2^{-k} \qquad *$$

We define

$$y_n = \sum_{k=1}^{n} |x_{N_k} - x_{N_{k-1}}| \qquad (X_{N_0} = 0)$$

Clearly y_n is positive and increasing with n so by applying Levi's Monotone Convergence Theorem;

$$\left(\int_\Omega |y_n(\omega)|^p d\mu(\omega)\right)^{1/p} \mapsto \left(\int_\Omega |y(\omega)|^p d\mu(\omega)\right)^{1/p}$$

where y denotes;

$$y = \sum_{k=1}^{\infty} |x_{N_k} - x_{N_{k-1}}|$$

Here

$$\|y\|_p = \lim_{n\mapsto\infty} \|y_n\|_p \leq \lim_{n\mapsto\infty} \sum_{k=1}^{\infty} \|x_{N_k} - x_{N_{k-1}}\|_p$$

Using * we see that $\|y\|_p$ is less than a finite constant, so that;

$$\int_\Omega |y(\omega)|^p d\mu(\omega) < \infty$$

This gives $y = \sum_{k=1}^{\infty} |x_{N_k} - x_{N_{k-1}}| < \infty$ except on the null set $\mathcal{N}$.

It is immediately clear that $\sum_{k=1}^{\infty} (x_{N_k} - x_{N_{k-1}})$ converges almost everywhere.

Letting

$$x = \sum_{k=1}^{\infty} (x_{N_k} - x_{N_{k-1}}) \text{ we show } x \in L_p$$

Clearly

$$|x| = \left|\sum_{k=1}^{\infty} (x_{N_k} - x_{N_{k-1}})\right| \leqslant \sum_{k=1}^{\infty} \left|x_{N_k} - x_{N_{k-1}}\right| = y \quad \text{a.e.}$$

Therefore

$$(\|x\|_p)^p = \int_\Omega |x(w)|^p d\mu(w) \leqslant \int_\Omega y^p d\mu = (\|y\|_p)^p < \infty$$

wich implies

$$\|x\|_p < \infty \qquad \text{and } x \in L_p$$

To show that $\|x_n - x\|_p \mapsto 0$ we first prove $\|x_{N_k} - x\|_p \mapsto 0$ as $k \mapsto \infty$

Now

$$\left|x_{N_k} - x_{N_j}\right| \leqslant \sum_{i=k+1}^{j} \left|x_{N_i} - x_{N_{i-1}}\right| < y$$

Therefore $\left|x_{N_k} - x_{N_j}\right|^p$ is dominated by y^p and we may apply the Lebesgue's Dominance Convergence Theorem.

Letting $j \mapsto \infty$ we have,

$$\left|x_{N_k} - x\right|^p \leqslant y^p$$

and

$$\int \lim_{k \mapsto \infty} \left|x_{N_k} - x\right|^p d\mu = \lim_{k \mapsto \infty} \int \left|x_{N_k} - x\right|^p d\mu = \lim_{k \mapsto \infty} (\|x_{N_k} - x\|_p)^p = 0$$

Now for fixed $\varepsilon > 0$, $\exists$ N such that if $m, n \geqslant N$;

$$\|x_n - x_m\|_p < \varepsilon/2$$

So for N_k, $n > N$

$$\|x_n - x_{N_k}\|_p < \varepsilon/2 \qquad (1)$$

and since

$$\|x_{N_k} - x\|_p \longmapsto 0$$

We can, by choosing a K sufficiently large obtain an N_k such that;

$$\|x_{N_k} - x\|_p < \varepsilon/2 \qquad (2)$$

Substituting into the triangle inequality;

$$\|x - x_n\|p \leq \|x - x_{N_k}\| + \|x_{N_k} - x_n\| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

for some $n \geq N$.

Thus, x_n converges to x in L_p, L_p is complete and hence a Banach Space. ∎

We note here and use later that the particular space L_2,

$$\int_{\Omega} |x(\omega)|^2 d\mu(\omega) < \infty$$

with the 2-norm is a Banach Space.

A.1.5. HILBERT SPACES

We have already given the concepts of the metric and the norm noting their similarities to the ideas of distance and length in Euclidean Space. In this, the final part of Appendix A we introduce the concept of the inner product - this is the formulisation of the Euclidean idea of an 'angle' between elements of a Vector Space. More importantly, we will be able to talk about the perpendicularity or orthogonality of these vectors. The induced metric was discussed earlier, pursuing this notion, the norm associated with an inner product will be introduced. It is shown that this norm is well defined. A Banach space with its norm so described is termed a Hilbert Space.

We begin with the definition of the inner product.

Definition: INNER PRODUCT

An inner product on a real vector space V is a mapping $\langle\ \rangle : V \times V \longmapsto \mathbb{R}$, and satisfies the following rules;

For v, v_1, $v_2 \in V$ and r_1, $r_2 \in \mathbb{R}$

I(1) $\langle r_1 v_1 + r_2 v_2, v\rangle = r_1 \langle v_1, v\rangle + r_2 \langle v_2, v\rangle$

I(2) $\langle v_1, v_2\rangle = \langle v_2, v_1\rangle$

I(3) $\langle v, v\rangle > 0 \iff v \neq 0$

The double $(V,\langle\ \rangle)$ is called an inner product space (i,p.s.). The 'dot' product defined by;

$$\langle v_1,v_2\rangle = v_1\cdot v_2 = \sum_{i=1}^{3} v_{1i}v_{2i}$$

may already be familiar from elementary vector analysis.

If $\langle v_1,v_2\rangle = 0$ we say that v_1 is orthogonal to v_2 and denoted this by $v_1 \perp v_2$. We will assume that our inner product spaces always carry an associated norm given by;

$$\|v\| = \langle v,v\rangle^{\frac{1}{2}},$$

that it is truely a norm, we leave until Theorem 4.2.4.

It is easily shown that $\langle\ \rangle$ is linear in both vectors of the ordered pair, (I(1) + I(2)), we therefore say $\langle\ \rangle$ is a bilinear functional.

THEOREM A.1.3.

The Cauchy-Schwarz Inequality:

Let x_1,x_2 be elements of the i.p.s. X then;

$$|\langle x_1,x_2\rangle| \leq \|x_1\|\,\|x_2\|$$

where

$$\|x\| = \langle x,x\rangle^{\frac{1}{2}}$$

If x_1 or x_2 is the origin,

$$\langle x_1,0\rangle = \langle 0,x_2\rangle = \langle 0,0\rangle = 0$$

and the inequality holds. So let us assume $x_1,x_2 \neq 0$

Clearly

$$\langle x_1-rx_2,x_1-rx_2\rangle \geq 0 \qquad \forall r\in \mathbb{R}$$

In particular let;

$$r = \frac{\langle x_1,x_2\rangle}{\langle x_2,x_2\rangle}$$

Then

$$\langle x_1-rx_2,\ x_1-rx_2\rangle = \langle x_1,x_1\rangle - \frac{|\langle x_1,x_2\rangle|^2}{\langle x_2,x_2\rangle} \geq 0$$

Taking roots we obtain the desired inequality. ■

THEOREM A.1.4.

The inner product norm given by;

$$\|x\| = \langle x,x\rangle^{\frac{1}{2}}$$

is a well defined norm. ▼

PROOF

The only property that does not follow trivially is N(3), the triangle inequality;

$$\|x_1+x_2\|^2 = \langle x_1+x_2, x_1+x_2\rangle = \|x_1\|^2+2\langle x_1,x_2\rangle+\|x_2\|^2$$

$$\leqslant \|x_1\|^2+2|\langle x_1,x_2\rangle|+\|x_2\|^2$$

We now employ the Cauchy-Schwarz inequality;

$$\|x_1\|^2+2\|x_1\|\|x_2\|+\|x_2\|^2$$

$$(\|x_1\|+\|x_2\|)^2$$

Taking square roots we satisfy N(3) and the theorem is proved. ■

Definition: HILBERT SPACE

A Hilbert Space is an inner product space whose associated norm (i.p.norm) is complete.

We now show that the function;

$$\langle x_1,x_2\rangle = \int_\Omega x_1(w)^T x_2(w)\,d\mu(w) \qquad x_1,x_2 \in L_2$$

is an inner product and hence $(L_2,\langle\ ,\ \rangle)$ is an inner product space.

I(1) $\langle r_1x_1+r_2x_2,x\rangle = \int_\Omega (r_1x_1+r_2x_2)^T x\,d\mu = r_1\langle x_1,x\rangle+r_2\langle x_2,x\rangle$

I(2) $\langle x_1,x_2\rangle = \int_\Omega x_1^T x_2\,d\mu = \int_\Omega x_2^T x_1\,d\mu = \langle x_2,x_1\rangle$

I(3) $\langle x,x\rangle = \int_\Omega x^2 d\mu > 0 \iff x \neq 0$

We have shown already (THEOREM A.1.2.) that the associated norm

$$\|x\| = \langle x,x\rangle^{\frac{1}{2}} = \left(\int_\Omega x^2 d\mu\right)^{\frac{1}{2}}$$

is complete, hence the Lebesgue Space L_2 is a Hilbert Space.

A.2. PROPERTIES OF THE HILBERT SPACE

The object of this section is to prove two very important theorems viz;

1) Given a closed convex subset of a Hilbert Space there exists a unique element in the subset with smallest norm.

2) Each vector of the Hilbert Space may be represented by the combination of a unique element in a closed subspace and one in the orthogonal complement of that subspace.

Loosely, these two theorems may be taken to mean;

If we wish to estimate a vector in a Hilbert Space, given a subspace, there is a unique estimate of this vector in this subspace that minimises our estimation error. In fact, we will see that the estimate is the orthogonal projection of the vector into the subspace.

In order to put this into the context of the General Problem (Section 3.2.) let x be a vector of the State Space X. The estimate of this vector will be denoted $\hat{x}$ and the error $(x-\hat{x})$ by $\tilde{x}$. The object is to minimise the estimation error in some way. If we can define a function that represents the cost of estimation error in a meaningful way as an inner product, then there is a unique $\hat{x}_o$ that minimises this cost. In this way we say $\hat{x}_o$ is the <u>optimal</u> estimate of x. Further, if the subspace is the space generated by all linear combinations of the observations then $\hat{x}_o$ is the orthogonal projection of x into this subspace.

We begin by giving two elementary but extremely useful identities.

(1) GENERALISED PYTHAGOREAN IDENTITY

Let $x_1 \dots x_n$ be pairwise orthogonal i.e.; $\langle x_i, x_j \rangle = 0 \ \forall i \neq j$, then,

$$\|\sum_{i=1}^{n} x_i\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

$$\|\sum_{i=1}^{n} x_i\|^2 = \langle x_1+x_2+\dots x_n, x_1+x_2+\dots x_n \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x_i x_j \rangle$$

$$= \sum_{i=1}^{n} \langle x_i x_i \rangle = \sum_{i=1}^{n} \|x_i\|^2$$

(2) PARALLELOGRAM IDENTITY

For all $x_1, x_2 \in X$

$$\|x_1+x_2\|^2 + \|x_1-x_2\|^2 = 2(\|x_1\|^2 + \|x_2\|^2)$$

$$\|x_1+x_2\|^2 + \|x_1-x_2\|^2 = \langle x_1+x_2, x_1+x_2 \rangle + \langle x_1-x_2, x_1-x_2 \rangle$$

$$= \langle x_1, x_1 \rangle + 2\langle x_1, x_2 \rangle + \langle x_2, x_2 \rangle$$

$$+ \langle x_1 x_1 \rangle - 2\langle x_1, x_2 \rangle + \langle x_2, x_2 \rangle$$

$$= 2\|x_1\|^2 + 2\|x_2\|^2$$

We now give two definitions;

If A is a subset of a ' ' space X and is itself a ' ' space, we say that A is a (' ') subspace of X.

If $\forall x_1, x_2 \in X$ and $0 \leq \lambda \leq 1$, $\lambda x_1 + (1-\lambda)x_2 \in X$ we say that X is convex.

THEOREM A.2.1.

Let $\tilde{X}$ be a convex subspace of a Hilbert Space X, then there is a unique element $\tilde{x}_0 \in \tilde{X}$ that has the smallest norm. Moreover for this vector $\tilde{x}_0$;

$$\langle \tilde{x}_0, \tilde{x} - \tilde{x}_0 \rangle \geq 0 \qquad \forall \tilde{x} \in \tilde{X}$$

PROOF

There are three parts to the proof:-

(1) The existence of $\tilde{x}_o$

(2) Its uniqueness

(3) $\langle \tilde{x}_o, \tilde{x}-\tilde{x}_o \rangle \geqslant 0 \qquad \forall \tilde{x} \in \tilde{X}$

Let $\varepsilon = \inf\{\|\tilde{x}\| : \tilde{x} \in \tilde{X}\}$, then there is a sequence $\{\tilde{x}_n\} \in \tilde{X}$ such that

$$\|\tilde{x}_n\| \longmapsto \varepsilon \quad \text{as } n \longmapsto \infty$$

So let

$$\tilde{x}_n, \tilde{x}_m \in \tilde{X}, \text{ since } \tilde{X} \text{ is convex } \tfrac{1}{2}(\tilde{x}_n, \tilde{x}_m) \in \tilde{X}$$

Clearly

$$\|\tfrac{1}{2}(\tilde{x}_n+\tilde{x}_m)\| \geqslant \varepsilon^2 \qquad \therefore \qquad \|\tilde{x}_n+\tilde{x}_m\| \geqslant 4\varepsilon^2$$

By applying the Parellogram Identity to $\tilde{x}_m, \tilde{x}_n$;

$$\begin{aligned}\|\tilde{x}_n-\tilde{x}_m\|^2 &= 2\|\tilde{x}_n\|^2+2\|\tilde{x}_m\|^2-\|\tilde{x}_n+\tilde{x}_m\|^2 \\ &\leqslant 2\|\tilde{x}_n\|^2+2\|\tilde{x}_m\|^2-4\varepsilon^2\end{aligned}$$

On taking limits;

$$\|\tilde{x}_n-\tilde{x}_m\|^2 \mapsto 0$$

and is therefore Cauchy in $\tilde{X}$.

Since $\tilde{X}$ is closed (being complete CX), $\tilde{x}_n$ converges to a point in $\tilde{X}$, $\tilde{x}_o$ say.

■1.

To show uniqueness we again employ the Parallelogram Identity. For $\tilde{x}_1, \tilde{x}_2 \in \tilde{X}$ we state from the first part of the proof;

$$\|\tilde{x}_1-\tilde{x}_2\|^2 \leqslant 2\|\tilde{x}_1\|^2+2\|\tilde{x}_2\|^2-4\varepsilon^2$$

So, if $\tilde{x}_1$ and $\tilde{x}_2$ both give the smallest norm (ε);

$$\|\tilde{x}_1-\tilde{x}_2\|^2 \leqslant 0$$

which implies $\tilde{x}_1=\tilde{x}_2$ and uniqueness is shown.

■2.

We now show $\langle \tilde{x}_0, \tilde{x}-\tilde{x}_0 \rangle \geqslant 0 \quad \forall \tilde{x} \in \tilde{X}$

Since $\tilde{X}$ is convex we have for $0 < \lambda \leqslant 1$

$$\lambda\tilde{x}+(1-\lambda)\tilde{x}_0 = \tilde{x}_0+\lambda(\tilde{x}-\tilde{x}_0) \in \tilde{X}$$

Now

$$\epsilon^2 = \|\tilde{x}_0\|^2 \leqslant \|\tilde{x}_0+\lambda(\tilde{x}-\tilde{x}_0)\|^2$$
$$= \|\tilde{x}_0\|^2+2\lambda\langle \tilde{x}_0,\tilde{x}-\tilde{x}_0\rangle+\lambda^2\|\tilde{x}-\tilde{x}_0\|^2$$

$$0 \leqslant \langle \tilde{x}_0,\tilde{x}-\tilde{x}_0\rangle+\tfrac{1}{2}\lambda\|\tilde{x}-\tilde{x}_0\|^2$$

and letting $\lambda \mapsto 0$, we have;

$$0 \leqslant \langle \tilde{x}_0,\tilde{x}-\tilde{x}_0\rangle$$

■ 3.

Definition: <u>ORTHOGONAL COMPLEMENT</u>

Let A be a non-empty subset of X. If for any $x \in X$ and for all $a \in A$ $\langle x,a\rangle = 0$, we say x is orthogonal to the set A, and further define $A^{\perp}$, the orthogonal complement of A, as the set,

$$A^{\perp} = \{x \in X: \ \langle x,a\rangle = 0, \ \forall a \in A\}$$

This may be shortened by writing;

$$A^{\perp} = \{x \in X, \ \langle x,A\rangle = 0\}$$

In the next theorem we prove that any vector $x \in X$ may be uniquely decomposed into two vectors; $\hat{x}_0$ belonging to the estimation space $\hat{X}$ and $\tilde{x}_0$ belonging to its orthogonal complement $\tilde{X}$, the error space.

<u>THEOREM A.2.2 (DECOMPOSITION THEOREM)</u>

Let $\hat{X}$ be a closed vector space of X and let $x \in X$, then there exists a unique $\hat{x}_0 \in \hat{X}$ such that $x-\hat{x}_0 \in \tilde{X} = (\hat{X}^{\perp})$. ▼

PROOF

We define the coset of $\hat{X}$ containing x as the set;

$$x+\hat{X} = \{x+\hat{x};\ \hat{x} \in \hat{X}\}$$

This is both closed and convex.

Let $\tilde{x}_0$ be the point in $x+\hat{X}$ with smallest norm, so that $\tilde{x}_0$ may be written as $\tilde{x}_0 = x-\hat{x}_0$ for some $\hat{x}_0 \in \hat{X}$.

Now if $\tilde{x}_0$ were not in $\tilde{X}$ we could find an $\hat{x} \in \hat{X}$ such that $\langle \tilde{x}_0, \hat{x} \rangle \neq 0$ therefore, there is an $\hat{x}^* = \lambda \hat{x} \in \hat{X}$ with $\langle \tilde{x}_0, \hat{x}^* \rangle < 0$

But

$$\tilde{x}_0+\hat{x}^* = x-\hat{x}_0+\hat{x}^* = x + (\hat{x}^* - \hat{x}_0) \in x + \hat{X}$$

We now apply the third part of the previous theorem and obtain;

$$0 \leq \langle \tilde{x}_0, (\tilde{x}_0+\hat{x}^*) - \tilde{x}_0 \rangle = \langle \tilde{x}_0, \hat{x}^* \rangle < 0$$

which is a contradiction and therefore $x-\hat{x}_0 = \tilde{x}_0 \in \tilde{X}$

■ 1.

To prove uniqueness:

If $\hat{x}_1, \hat{x}_2$ are two vectors in $\hat{X}$, clearly $\hat{x}_1-\hat{x}_2 \in \hat{X}$, since $\hat{X}$ is closed, also $x-\hat{x}_1$, $x-\hat{x}_2 \in \tilde{X}$

$$\begin{aligned} 0 &= \langle x-\hat{x}_2, \hat{x}_1-\hat{x}_2 \rangle - \langle x-\hat{x}_1, \hat{x}_1-\hat{x}_2 \rangle \\ &= \langle \hat{x}_1-\hat{x}_2, \hat{x}_1-\hat{x}_2 \rangle \\ &\Rightarrow \hat{x}_1 = \hat{x}_2 \end{aligned}$$

■ 2.

APPENDIX B

SOLUTION OF AN EXTENDED STOCHASTIC ESTIMATION PROBLEM

In Chapter 3, a general stochastic problem was solved. It was remarked there that the problem could be easily extended to include; a deterministic control input, colouration of the system and observation disturbances and a system error transition matrix. In this Appendix the extended General Problem is solved. We begin then with the statement of the problem;

Let the equation representing the propagation of the system through time be;

$$x_{k+1} = \phi_{k+1,k}x_k + \theta_{k+1,k}y_k + \chi_{k+1,k}dx_k$$

where

x_k is the state of the system at time k

$\phi_{k+1,k}$ is the state transition matrix

y_k is the control input at time k

$\theta_{k+1,k}$ is the control transition matrix

dx_k is the random error vector at time k

$\chi_{k+1,k}$ is the system error transition matrix

It is further assumed that;

$$E(dx_k) = 0 \qquad \forall k$$

and that

$$E(dx_k dx_j^T) = \langle dx_k, dx_j \rangle = R_k \delta_{kj}$$

where δ_{kj} is the 'kronecker delta' and R_k is a real symmetric positive definite matrix. The observation system is the same as in Chapter 3, so that;

$$z_{k+1} = \psi_{k+1}x_{k+1} + dz_{k+1}$$

where

z_k is the observation vector at time k

ψ_k is the observation transition matrix

dz_k is the observation error vector at time k

It is further assumed that;

$$E(dz_k) = 0 \qquad \forall k$$

and that

$$E(dz_k dz_j^T) = \langle dz_k, dz_j \rangle = S_k \delta_{kj}$$

Now, to allow for correlation between the observation and system disturbances;

$$E(dx_k, dz_j^T) = \langle dx_k, dz_j \rangle = T_k \delta_{kj}$$

but the independence of the system and observation errors to any initial state vector x_0 is retained.

The filtered estimates and their error covariance matrices are the same as in Chapter 3 so the task is to find the optimal estimates of the one-step prediction of the state vector and its error covariance matrix.

<u>THEOREM B.1.</u>

$$\hat{x}_{0\ k+1,k} = \phi_{k+1,k}\hat{x}_{0\ k,k} + B_k'(z_k - U_k\hat{x}_{k,k-1}) + \theta_{k+1,k}y_k$$

where

$$B_k' = \chi_{k+1,k}T_k(\psi_k C_{0\ k,k-1}\psi_k^T + S_k)^{-1}$$

<u>PROOF</u>

Now

$$x_{k+1} = \hat{x}_{0\ k+1,k} + \tilde{x}_{0\ k+1,k}$$

So it is sufficient to prove that;

$$x_{k+1} - (\phi_{k+1,k}\hat{x}_{0\ k,k} + B_k'(z_k - \psi_k\hat{x}_{0\ k,k-1}) + \theta_{k+1,k}y_k) \in \tilde{X}_k$$

Let $\hat{x}_{k,k} \in \hat{X}_k$ and consider;

$$\langle x_{k+1} - (\phi_{k+1,k}\hat{x}_{0\ k,k} + B'_k(z_k - \psi_k\hat{x}_{0\ k,k-1}) + \theta_{k+1,k}y_k) \ , \ \hat{x}_{k,k}\rangle$$

$$= \langle \phi_{k+1,k}\tilde{x}_{0\ k,k} - B'_k\psi_k\tilde{x}_{0\ k,k-1} - B'_k dz_k + \chi_{k+1,k}dz_k \ , \ \hat{x}_{k,k}\rangle$$

$$= \langle - B'_k\psi_k\tilde{x}_{0\ k,k-1} - B'_k dz_k + \chi_{k+1,k}dz_k, \hat{x}_{k,k}\rangle \qquad \text{B.1.1.}$$

since $\langle \tilde{x}_{0\ k,k}, \hat{x}_{k,k}\rangle = 0$

Now as $\hat{x}_{k,k} \in \hat{X}_k$;

$$\hat{x}_{k,k} = \sum_{i=1}^{k} A_i z_i \qquad \text{for some } A_i$$

This can now be expanded because;

$$z_i = \psi_i x_i + dz_i$$

thus

$$\hat{x}_{k,k} = \sum_{i=1}^{k} A_i(\psi_i(\tilde{x}_{0\ i,i-1} + \hat{x}_{0\ i,i-1}) + dz_i)$$

Substituting into Equation B.1.1. and ignoring terms identically zero, yields;

$$\langle - B'_k U_k \tilde{x}_{0\ k,k-1} - B'_k dz_k + \chi_{k+1,k}dx_k, \ A_k\psi_k\tilde{x}_{0\ k,k-1} + A_k dz_k\rangle$$

$$= \quad - B'_k\psi_k C_{0\ k,k-1}\psi_k^T A_k^T - B'_k S_k A_k^T + \chi_{k+1,k}T_k A_k^T$$

$$= \quad (-B'_k(\psi_k C_{0\ k,k-1}\psi_k^T + S_k) + \chi_{k+1,k}T_k)A_k^T$$

and substituting for B'_k, the above term vanishes and the theorem is proven.

THEOREM B.2.

$$C_{0\ k+1,k} = \phi_{k+1,k}C_{0\ k,k}\phi_{k+1,k}^T + \chi_{k+1,k}R_k\chi_{k+1,k}^T - \phi_{k+1,k}B_k T_k^T\chi_{k+1,k}^T$$
$$- \chi_{k+1,k}T_k B_k^T\phi_{k+1,k}^T - B'_k T_k^T\chi_{k+1,k}^T$$

where as in Chapter 3;

$$B_k = C_{0\;k,k}\psi_k^T S_k^{-1}$$

$$\hat{x}_{0\;k,k} = \hat{x}_{0\;k,k-1} + B_k(z_k - \psi_k\hat{x}_{0\;k,k-1})$$

and

$$C_{0\;k,k} = (I - B_k\psi_k)C_{0\;k+1,k}$$

<u>PROOF</u>

From the previous theorem;

$$\hat{x}_{0\;k+1,k} = \phi_{k+1,k}\hat{x}_{0\;k,k} + B'_k(z_k - \psi_k\hat{x}_{0\;k,k-1}) + \theta_{k+1,k}y_k$$

so that

$$\tilde{x}_{0\;k+1,k} = \phi_{k+1,k}\tilde{x}_{0\;k,k} - B'_k\psi_k\tilde{x}_{0\;k,k-1} - B'_k dz_k + \chi_{k+1,k}dx_k$$

Therefore

$$C_{0\;k+1,k} = \rangle\, \tilde{x}_{0\;k+1,k}, \tilde{x}_{0\;k+1,k} \langle$$

$$= \phi_{k+1,k}C_{0\;k,k}\phi^T_{k+1,k} + B'_k(\psi_k C_{0\;k,k-1}\psi^T_k + S_k)B^T_k$$

$$+ \chi_{k+1,k}R_k\chi^T_{k+1,k} - B'_k T^T_k\chi^T_{k+1,k} - \chi_{k+1,k}T_k B'^T_k$$

$$+ \rangle\, \phi_{k+1,k}\tilde{x}_{0\;k,k}, - B'_k\psi_k\tilde{x}_{0\;k,k-1} - B'_k dz_k + \chi_{k+1,k}dx_k \langle$$

$$+ \rangle - B'_k\psi_k\tilde{x}_{0\;k,k-1} - B'_k dz_k + \chi_{k+1,k}dx_k, \phi_{k+1,k}\tilde{x}_{0\;k,k} \langle$$

+ other terms equal to zero.

B.1.2.

Now

$$\hat{x}_{0\;k,k} = \hat{x}_{0\;k,k-1} + B_k(z_k - \psi_k\hat{x}_{0\;k,k-1})$$

Hence

$$\tilde{x}_{0\;k,k} = \tilde{x}_{0\;k,k-1} - B_k\psi_k\tilde{x}_{0\;k,k-1} - B_k dz_k$$

$$= (1 - B_k\psi_k)\tilde{x}_{o\;k,k-1} - B_k dz_k$$

This expression may be substituted into the sixth term of Equation B.1.2. to give;

$$\rangle\, \phi_{k+1,k}(1 - B_k\psi_k)\tilde{x}_{0\ k,k-1} - \phi_{k+1,k}B_k dz_k, - B'_k\psi_k\tilde{x}_{0\ k,k-1} - B'_k dz_k + \chi_{k+1,k}dx_k \,\langle$$

$$= - \phi_{k+1,k}(1 - B_k\psi_k)C_{0\ k,k-1}\psi_k^T\, {B'_k}^T + \phi_{k+1,k}B_kS_k{B'_k}^T - \phi_{k+1,k}B_kT_k^T\chi_{k+1,k}^T$$

$$= - \phi_{k+1,k}C_{0\ k,k}\psi^T{B'_k}^T + \phi_{k+1}B_kS_k{B'_k}^T - \phi_{k+1,k}B_kT_k^T\chi_{k+1,k}^T$$

But

$$C_{0\ k,k}\psi_k^T = B_kS_k$$

and the above equation reduces to;

$$- \phi_{k+1,k}B_kT_k^T\chi_{k+1,k}^T$$

Substituting this back into Equation B.1.2. the following is obtained;

$$C_{0\ k+1,k} = \phi_{k+1,k}C_{0\ k,k}\phi_{k+1,k}^T + B'_k(\psi_kC_{0\ k+1,k}\psi_k^T + S_k){B'_k}^T$$

$$+ \chi_{k+1,k}R_k\chi_{k+1,k}^T - B'_kT_k^T\chi_{k+1,k}^T - \chi_{k+1,k}T_k{B'_k}^T$$

$$- \phi_{k+1,k}B_kT_k^T\chi_{k+1,k}^T - X_{k+1,k}T_kB_k^T\phi_{k+1,k}^T$$

Noting that the second and fifth terms are complements the Theorem is proven.

The solution of the extended solution is now summarised in the following Theorem.

<u>THEOREM B.3.</u>

The optimal filtered and one-step prediction estimates and their respective covariance matrices for the extended General Problem may be obtained recursively given initial estimates $\hat{x}_{00,0}$ and $C_{00,0}$ from the following equations;

$$\hat{x}_{0\ k+1,k} = \phi_{k+1,k}\hat{x}_{0\ k,k} + B_k(z_k - \psi_k\hat{x}_{0\ k,k-1}) + \theta_{k+1,k}y_k$$

B.3.1.

$$C_{0\ k+1,k} = \phi_{k+1,k}C_{0\ k,k}\phi^T_{k+1,k} + \chi_{k+1,k}R_k\chi^T_{k+1,k} - B'_kT^T_k\chi^T_{k+1,k}$$
$$- \phi_{k+1,k}B_kT^T_k\chi^T_{k+1,k} - \chi_{k+1,k}T_kB^T_k\phi^T_{k+1,k}$$

B.3.2.

$$\hat{x}_{0\ k,k} = \hat{x}_{0\ k,k-1} + B_k(z_k - \psi_k\hat{x}_{0\ k,k-1})$$

B.3.3.

$$C_{0\ k,k} = (I - B_k\psi_k)C_{0\ k,k-1}$$

B.3.4.

where

$$B'_k = \chi_{k+1,k}T_k(\psi_kC_{0\ k,k-1}\psi^T_k + S_k)^{-1}$$

B.3.5.

$$B_k = C_{0\ k,k-1}U^T_k(\psi_kC_{0\ k,k-1}U^T_k + S_k)^{-1}$$

B.3.6.

▼

PROOF

Equations B.3.1. and B.3.5. are given in Theorem B.1., Equation B.3.2. in Theorem B.2., Equation B.3.3. and B.3.6. in Theorem 3.3.4. and Equation B.3.4. in Theorem 3.3.5..

■

APPENDIX C

THE EVALUATION OF THE THREE MODELS ON SIMULATED AND DUNLOP DATA

When the covariance matrices of the system and observation disturbances are known for all time, they may be substituted into the recursive relationships of the General Problem to derive the optimal solution. In manpower planning it is rare that the error covariance matrices will be known exactly for all future time periods and in this case they have to be estimated. Three methods of estimating their values were detailed in Chapter 4 and are recalled here for ease of reference;

(1) Unbiased estimation by the evaluation of certain differences in the observations

(2) Bayesian estimation by the selection of values from a class of alternative models so that the model employed for the (t, t + 1) time period is the one with the highest posterior probability of attaining at time t

(3) Bayesian estimation by constructing the model for time period (t, t + 1) as a posterior probability weighting of the models at time t

In order that the proposed models may be compared with the more usual (classical) estimator (4), that of sample mean and sample variance, results using this model are also included. All four methods are compared on two data series; one simulated and the other taken from Dunlop. The comparison is made by two techniques; the first and probably the more familiar is the often used relative mean sum of squared error and the second is the relative mean likelihood. On the simulated data relative refers to the optimal estimator and on Dunlop data, where the optimal estimator is 'a priori' unknown, to the third estimator.

C.1. SIMULATED DATA

The data was created on computer by assuming a starting point for the process as 100 and generating a time series for the system by adding at each time period a realisation from a random process with standard deviation 2 and zero mean to the previous system value. An observation series of these system values was generated by assuming random error with zero mean and variance 20. These observations are presented in the following table.

TABLE C.1. SIMULATED OBSERVATIONS

TIME	OBS	TIME	OBS	TIME	OBS	TIME	OBS	TIME	OBS
1	97.66	11	96.11	21	91.11	31	92.58	41	102.46
2	101.42	12	86.52	22	99.11	32	95.48	42	99.90
3	101.30	13	102.11	23	97.56	33	97.41	43	95.29
4	96.98	14	86.13	24	92.85	34	90.54	44	93.15
5	97.46	15	86.29	25	97.02	35	90.83	45	94.05
6	100.19	16	92.16	26	102.00	36	96.64	46	80.77
7	90.19	17	91.11	27	89.64	37	94.55	47	86.07
8	97.40	18	94.09	28	92.47	38	100.06	48	80.74
9	93.42	19	94.70	29	98.46	39	92.25	49	90.05
10	95.37	20	94.45	30	93.70	40	96.23	50	83.15

To form the Bayesian type (2,3) estimators of the process, only three alternatives, assumed to cover the normal range of variance, were proposed;

$$\theta_1 = \quad R = 0, \quad S = 10, \quad C_o = 10$$

$$\theta_2 = \quad R = 10, \quad S = 10, \quad C_o = 30$$

$$\theta_3 = \quad R = 10, \quad S = 100, \quad C_o = 210$$

In this case the optimal estimator is known and can be used as a standard by which the other estimators may be compared.

TABLE C.2. ESTIMATES USING THE OPTIMAL ESTIMATOR

TIME	EST	TIME	EST	TIME	EST	TIME	EST	TIME	EST
1	100.00	11	95.37	21	93.53	31	94.94	41	95.36
2	98.64	12	95.63	22	92.67	32	94.10	42	97.90
3	99.86	13	92.37	23	94.97	33	94.59	43	98.62
4	100.42	14	95.86	24	95.90	34	95.60	44	97.43
5	99.14	15	92.37	25	94.81	35	93.79	45	95.89
6	98.53	16	90.19	26	95.60	36	92.73	46	95.23
7	99.13	17	90.90	27	97.89	37	94.13	47	90.05
8	95.92	18	90.97	28	94.94	38	94.28	48	88.63
9	96.45	19	92.09	29	94.05	39	96.35	49	85.80
10	95.36	20	93.03	30	95.63	40	94.88	50	87.32

TABLE C.3. ESTIMATES USING ESTIMATOR 1(UNBIASED)

TIME	EST	TIME	EST	TIME	EST	TIME	EST	TIME	EST
1	100.00	11	96.27	21	93.50	31	95.01	41	95.05
2	98.29	12	96.19	22	92.50	32	94.28	42	96.85
3	99.39	13	90.88	23	95.28	33	94.63	43	97.49
4	101.24	14	94.16	24	96.17	34	95.42	44	97.07
5	97.29	15	91.40	25	94.81	35	94.07	45	96.34
6	97.45	16	89.40	26	95.68	36	93.25	46	95.88
7	100.12	17	90.53	27	98.19	37	94.06	47	93.01
8	99.10	18	90.17	28	95.00	38	94.16	48	91.34
9	99.02	19	91.78	29	94.15	39	95.37	49	88.05
10	97.16	20	92.85	30	95.60	40	94.70	50	88.67

TABLE C.4. ESTIMATES USING ESTIMATOR 2 (BAYES 1)

TIME	EST	TIME	EST	TIME	EST	TIME	EST	TIME	EST
1	100.00	11	97.40	21	94.28	31	94.70	41	95.08
2	98.83	12	97.29	22	92.32	32	93.39	42	99.82
3	99.69	13	90.02	23	96.52	33	94.68	43	99.87
4	100.09	14	97.49	24	97.16	34	96.37	44	97.04
5	99.47	15	90.47	25	94.50	35	92.77	45	94.64
6	99.14	16	87.89	26	96.06	36	91.57	46	94.27
7	99.29	17	90.53	27	99.73	37	94.70	47	85.93
8	98.15	18	90.89	28	93.49	38	94.61	48	86.02
9	98.07	19	92.87	29	92.86	39	97.98	49	81.90
10	97.60	20	94.00	30	96.32	40	94.44	50	86.12

TABLE C.5. ESTIMATES USING ESTIMATOR 3 (BAYES 2)

TIME	EST	TIME	EST	TIME	EST	TIME	EST	TIME	EST
1	100.00	11	95.63	21	93.70	31	94.88	41	95.39
2	98.29	12	95.81	22	92.83	32	94.04	42	97.33
3	99.39	13	93.69	23	94.81	33	94.58	43	98.37
4	100.10	14	96.09	24	95.88	34	95.58	44	97.15
5	99.04	15	93.77	25	94.81	35	94.03	45	95.72
6	98.45	16	90.67	26	95.62	36	92.83	46	95.11
7	99.07	17	91.23	27	97.50	37	94.23	47	92.59
8	97.05	18	91.18	28	95.35	38	94.35	48	89.86
9	97.20	19	92.37	29	94.11	39	96.26	49	86.46
10	95.81	20	93.26	30	95.62	40	94.87	50	87.70

TABLE C.6. ESTIMATES USING THE 'CLASSICAL' ESTIMATOR

TIME	EST	TIME	EST	TIME	EST	TIME	EST	TIME	EST
1	100.00	11	97.14	21	94.75	31	94.97	41	94.89
2	97.66	12	97.05	22	94.58	32	94.89	42	95.07
3	99.54	13	96.17	23	94.79	33	94.91	43	95.19
4	100.13	14	96.63	24	94.91	34	94.98	44	95.14
5	99.34	15	95.88	25	94.82	35	94.85	45	95.10
6	98.96	16	95.24	26	94.91	36	94.74	46	95.07
7	99.17	17	95.04	27	95.18	37	94.79	47	94.76
8	97.89	18	94.81	28	94.98	38	94.92	48	94.58
9	97.83	19	94.77	29	94.89	39	94.92	49	94.29
10	97.43	20	94.77	30	95.01	40	94.85	50	94.20

The mean sum of squared error is calculated as;

$$(N_1 - N_2)^{-1} \sum_{i=N_2}^{N_1} (z_i - \hat{z}_i)^2$$

where $\hat{z}_i$ is the relevant estimate at time i. Relative efficiences are obtained by taking the inverse ratio of this value to that of the optimal estimator and scaling to 100. The results are shown for the whole series ($N_1 = 50$, $N_2 = 1$) and since the estimators are all 'learning', for the last 10 observations ($N_1 = 50$, $N_2 = 41$).

TABLE C.7.

Comparison of the relative efficiency of the various estimators over the whole simulated data series by the mean squared error statistic:

OPTIMAL	EST. 1	EST. 2	EST.3	CLASSICAL
100	87.9	88.3	97.2	81.9

TABLE C.8.

Comparison of the relative efficiency of the various estimators over the last 10 simulated observations by the mean squared error statistic:

OPTIMAL	EST.1	EST. 2	EST. 3	CLASSICAL
100	79.9	108.9	91.1	58.8

The second method of comparison, that of relative likelihood, is thought to be more valuable as it emcompasses not only the estimation error but also the estimated observation variance. Loosely speaking for estimators with the same magnitude of error, the smaller the variance the greater the value of the estimator. The likelihood of an observation is calculated as;

$$\left(\prod_{i=N_2}^{N_1} \hat{Var}(z_i)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(z_i - \hat{z}_i)^2 / \hat{Var}(z_i))\right)^{(N_1 - N_2)^{-1}}$$

Again values for the whole series and the last ten observations are presented.

TABLE C.9.

Comparison of the relative efficiency of the various estimators over the whole simulated data series by the mean likelihood statistic:

OPTIMAL	EST.1	EST.2	EST.3	CLASSICAL
100	84.8	88.6	91.3	75.0

TABLE C.10.

Comparison of the relative efficiency of the various estimators over the last ten simulated observations by the mean likelihood statistic:

OPTIMAL	EST.1	EST.2	EST.3	CLASSICAL
100	76.4	91.39	90.31	44.5

In each case the three proposed estimators show an improvement over the 'classical' estimator and in the latter part of the series the improvement is immense.

C.2. DUNLOP DATA

The data, on operatives of Dunlop Engineering Group, was obtained from the Central Personnel Division. It consists of 28 values of the wastage rate for the first 13 weeks of employment observed over the period 1971-1977.

TABLE C.11. OBSERVED QUARTERLY % WASTAGE RATES IN DUNLOP ENGINEERING GROUP

1971	1972	1973	1974	1975	1976	1977
28.4	39.5	34.9	10.1	21.3	21.4	22.7
33.7	39.5	22.6	15.9	25.8	23.6	10.1
33.5	42.0	25.2	19.8	22.2	20.2	17.8
39.3	43.9	15.7	16.9	14.1	29.8	12.3

The three estimators of Chapter 4 and the classical estimator were used to estimate this series and the following results were obtained:

TABLE C.12. ESTIMATES USING ESTIMATOR 1 (UNBIASED)

1971	1972	1973	1974	1975	1976	1977
25.0	37.8	43.8	16.4	17.1	15.0	29.7
27.9	39.5	37.0	11.0	21.0	21.1	23.2
32.5	39.9	26.6	15.5	25.3	23.5	10.6
33.5	42.0	25.3	19.3	22.4	20.4	17.3

TABLE C.13. ESTIMATES USING ESTIMATOR 2 (BAYES 1)

1971	1972	1973	1974	1975	1976	1977
25.0	37.8	43.3	17.1	17.2	15.3	26.4
27.6	39.1	36.9	11.1	20.7	20.5	24.1
31.3	39.7	24.6	15.2	25.1	22.2	12.1
33.2	41.5	25.1	19.1	22.6	21.0	13.8

TABLE C.14. ESTIMATES USING ESTIMATOR 3 (BAYES 2)

1971	1972	1973	1974	1975	1976	1977
25.0	38.0	43.3	17.4	17.3	15.9	28.0
27.9	39.2	36.7	11.4	20.4	20.1	23.9
32.5	39.7	24.8	14.9	24.7	22.8	12.4
33.3	41.4	25.1	18.7	22.8	20.8	16.6

TABLE C.15. ESTIAMTES USING THE 'CLASSICAL' ESTIMATOR

1971	1972	1973	1974	1975	1976	1977
25.0.	33.7	37.5	33.2	28.8	27.2	26.7
28.4	34.9	37.2	31.4	28.4	27.0	26.5
31.1	35.7	35.8	30.3	27.9	26.8	25.9
31.9	36.6	34.8	29.6	27.9	26.5	25.6

These four estimators were them compared using the test statistics detailed in Section C.1.. Since the observations come from real data, the optimal estimator was of course unknown, and for this data the third estimator was scaled to 100.

TABLE C.16. RELATIVE EFFICIENCIES OF THE FOUR ESTIMATORS ON DUNLOP DATA USING THE MEAN SUM OF SQUARED ERROR STATISTIC

TIME	EST.1.	EST.2.	EST.3.	CLASSICAL
1	100.0	100.0	100.0	100.0
2	100.0	92.7	100.0	114.0
3	100.0	86.2	100.0	101.7
4	103.0	90.5	100.0	82.1
5	102.0	90.1	100.0	63.1
6	102.5	90.0	100.0	53.5
7	103.4	90.6	100.0	45.5
8	106.1	91.5	100.0	38.3
9	98.2	94.9	100.0	64.6
10	96.9	96.2	100.0	77.6
11	96.5	96.1	100.0	62.7
12	96.3	96.8	100.0	47.9
13	99.3	98.0	100.0	34.2
14	98.6	97.5	100.0	30.6
15	99.7	98.1	100.0	30.1
16	99.2	97.9	100.0	27.8
17	99.0	97.8	100.0	27.9
18	100.0	98.4	100.0	29.2
19	99.5	98.1	100.0	29.0
20	100.5	98.8	100.0	29.9
21	99.0	97.9	100.0	30.7
22	99.8	98.3	100.0	31.1
23	99.3	98.7	100.0	30.8
24	98.5	99.2	100.0	34.1
25	96.2	101.0	100.0	35.0
26	98.6	100.2	100.0	38.6
27	96.6	99.9	100.0	38.8
28	96.1	101.4	100.0	37.1

TABLE C.17. RELATIVE EFFICIENCIES OF THE FOUR ESTIMATORS ON DUNLOP DATA USING THE MEAN LIKELIHOOD STATISTIC

TIME	EST.1.	EST.2.	EST.3.	CLASSICAL
1	100.0	100.0	100.0	100.0
2	100.0	89.2	100.0	60.4
3	100.0	96.8	100.0	86.4
4	100.8	93.2	100.0	56.8
5	117.6	100.8	100.0	57.6
6	121.2	106.1	100.0	59.1
7	125.0	109.6	100.0	58.8
8	128.6	110.7	100.0	55.7
9	94.5	101.6	100.0	66.9
10	36.3	74.5	100.0	58.8
11	40.4	76.0	100.0	55.8
12	45.4	79.4	100.0	46.4
13	47.9	81.3	100.0	38.5
14	50.0	80.6	100.0	37.8
15	52.5	81.8	100.0	38.4
16	53.9	81.4	100.0	37.3
17	54.8	80.8	100.0	38.5
18	56.7	81.7	100.0	40.4
19	58.5	81.1	100.0	40.6
20	60.2	82.5	100.0	41.7
21	61.2	83.5	100.0	42.7
22	63.5	83.7	100.0	44.2
23	64.2	84.0	100.0	44.3
24	65.0	84.5	100.0	47.6
25	66.0	86.4	100.0	48.5
26	67.4	83.2	100.0	50.5
27	67.4	84.2	100.0	51.6
28	68.8	83.3	100.0	51.0

It is clear that under both test statistics the three proposed estimators are far superior to the classical estimator on the live Dunlop data. Graphical representations of the efficiency of the estimators are given in Chapter 8 Section 2 of the main text.

APPENDIX D

PRELIMINARY WORK IN DUNLOP U.K.T.G.

Before any meaningful manpower analysis could be undertaken in Dunlop U.K.T.G. a reliable and up-to-date data base needed to be established. One characteristic was paramount to the success of such a data base, that of the interlacing of salary and manpower information. This link assured the quality of the manpower data as salary information is obviously a priority and therefore its accuracy is ensured. The basic input to both the salary and manpower suites was the pink Staff Changes Form GSA 17 T/O reproduced on the following two pages.

One of the key units for data analysis at the occupational level was the Job Activity Code and the Job Appointment Date (Datacode 781). A full list of these codes and their I.M.S. Survey equivalent are presented in Table D.1..

Once the data base had been constructed a preliminary manpower analysis of the 'camera' type was undertaken on 2880 staff employed in Dunlop U.K.T.G. factories in England. The results of this analysis are shown in Section 1 of this appendix.

DUNLOP STAFF CHANGES FORM (OBVERSE)

FORM GSA 17 T/D

DUNLOP LIMITED
(and Subsidiary Companies)

CONFIDENTIAL

STAFF CHANGES FORM

i) THIS FORM WILL NOT BE ACCEPTED UNLESS COMPLETED WHERE APPLICABLE IN EVERY DETAIL.
ii) PLEASE READ " NOTES FOR GUIDANCE " OVERLEAF BEFORE COMPLETING THIS FORM.

THE FOLLOWING CHANGES HAVE BEEN AUTHORISED AFFECTING:-

SURNAME:	INITIALS:	MR/MRS/MISS	DATE OF BIRTH	DATE JOINED DUNLOP GROUP (or Subsidiary Co)	DATE JOINED PRESENT DIVISION

No		NEW DETAILS	OLD DETAILS
1	i) Company or Division		
	ii) Location		
2	i) Departmental Title		
	ii) Departmental Cost Code		
3	i) Job Title 553		
	ii) Job Appointment Date 781		
4	i) Job Evaluation Scheme 550		
	ii) Points-Grade-Agreed Salary 551		

No		BASIC	SUPPLEMENTARY/ADDITIONAL
5	SALARY		
	i) Present Salary	£	
	ii) Increase Recommended	£	
	iii) New Salary	£	

No			
6	Date Change Operates		For Salaries Dept. Use Only.
7	Number of Days Pay in Lieu of Holiday		
8	Standard Weekly Hours		
9	Job Activity Code 781		
10	Keydate 552		Date Rec'd ....

No		CODE	REASON(S)
11	Reason(s) for Change(s) (Leaving) 554 (Salary) 555		
12	Manager's Signature & Date		

No		Complete for All Types of Change	Transfer Only - Receiving Div,unit,Dept
13	Approved by: i) Signature & Date	................................/.........../......	................................/............./......
	ii) Job Title	....	....
14	Authorised by: i) Signature & Date	................................/.........../......	................................/............./......
	ii) Job Title	....	....

DUNLOP STAFF CHANGES FORM (REVERSE)

NOTES FOR GUIDANCE

Alterations must be signed by the Authorising Official

MONTHLY SALARIES	The form must reach the Salaries Section London not later than 13 working days before payment is due	for change to operate in same month or week respectively
WEEKLY SALARIES	The form must reach the Salaries Section not later than Monday morning	

If paid monthly show SALARY PER ANNUM; if weekly SALARY PER WEEK

(1) SALARY CHANGES.

Use one of the following codes to describe the 'Reason(s) for Change(s)' Section (11)

P = Promotion	G = General Increase (Group - Wide)	Q = Qualifications Gained
M = Merit	N = Negotiated Award (Local Agreement)	J = In line with J.E.
R = Rate for Age	A = Market Rate Adjustment /Anomaly Correction	T = Completion of Training/Probation
I = Increased Responsibility		V = Annual Increment
H = Increased Hours		X = Other Reason
B = Back Dated Increase		

(2) INTER-DEPARTMENTAL TRANSFERS WITHIN TYRE DIVISION GROUP

The form should be completed for old details (Items 1 to 5) by the department from which the employee is being transferred and should be signed by the Manager of that department. It should then be sent to the Manager of the department to which the employee is being transferred and he will complete the 'new' part and countersign the form. It should then be authorised before being sent to Chief Salary Administrator F.D.

(3) TRANSFERS TO OTHER DUNLOP DIVISIONS

This form should be completed by the department from which the employee is being transferred and sent to Chief Salary Administrator, Tyre Division, Fort Dunlop.

In addition, the Company form GSA 17 should be completed by the same department and sent to the establishment to which the employee is being transferred.

(4) STAFF LEAVING THE COMPANY

(a) In cases where staff are taking a paid holiday immediately prior to leaving the Company, the leaving date on the form should be the date on which the holiday is completed and not the date the employee actually ceases work.

(b) Where payment in lieu of holiday is being made, the actual date of leaving should be given against item 6 and the number of day's pay due stated against item 7.

(c) Payment in lieu of notice will not be made unless the form is supported by a separate written request from the Division.

(d) Use one of the following codes to describe the 'Reason for Change' (Item 11).

Discharge	07 = Dissatisfaction with Job/Company	14 = Marriage
01 = Unsuitable	08 = Work relationships with other staff	15 = Pregnancy
02 = Disciplinary	09 = Personal Betterment	16 = Moving House
03 = Redundancy or Re - organisation		17 = Retirement
Resignation	10 = Transport difficulties	18 = Death
04 = Remuneration	11 = Housing difficulties	19 = Emigration
05 = Hours of Work	12 = Domestic Responsibility	20 = Other Reason
06 = Working Conditions	13 = Illness/Accident	21 = Cause Unknown

(5) JOB EVALUATION

(a) Use one of the following codes to describe the type of Job Evaluation Scheme (item 4i)

A = TD/FD	H = TASS	P = Pregnancy Absence
B = F.D. J.J.E.	HA= TASS - TD/FD	R = Research - ASTMS Sch.
C = Union Agreed Minimum	I = ASTMS - Exec.	S = Senior Secretary
D = Training Scale	IA = ASTMS - Exec - TD/FD	T = Off Site
E = Graduate Scale	J = Clerical Grade	TA = Off Site TD/FD
F = Rate for Age	K = Staff Mgt Grade	TB = Off Site F.D. J.J.E.
FA = R.F.A. - TD/FD	L = Machine Tool	V = Depot Scales - Marketing
FB = R.F.A. - F.D. J.J.E.	M = Retread Bilateral	W = Washington - Bilateral ASTMS
G = Accountancy Scale	N = Inchinnan - J.J.E.	

(b) Item 4ii) refers to the Points/Grade/Union Agreed Salary awarded under the above scheme Item 4i).

TABLE D.1. JOB ACTIVITY CODES

JOB ACTIVITY	CODE	I.M.S. CODE
PROFESSIONAL/MANAGERIAL		
Engineering		
Mechanical	111	PEM
Electrical	112	PEE
Civil	113	PEC
Chemical	114	PEH
Electronics	115	PEE
Instrument	116	PEO
Maintenance	117	PEO
Other	119	PEO
Scientific		
Physical	121	PSH
Chemical	122	PSC
Mathematical/Statistical	123	POM
Other	129	PSO
Computer		
Systems Analyst	131	SCU
Programmer	132	SCU
Operator	133	CUC
Other	139	CUM
Manager	130	POO
Managerial		
Director/General Works Manager	141	POO
Sales/Marketing	142	POO
Technical	143	POO
Planning/Distribution	144	POO
Office/Section/Department	145	POO
Personnel	146	POO
Buying	147	POO
Production	148	POO
Other	149	POO
Selected		
Accountant	161	POY
Architect	162	POA
Estate Surveyor	163	POS
Quantity Surveyor	164	POO
Legal	165	POL
Draughtsman	166	SEU
Designer	167	POO
Buyer	168	POO

JOB ACTIVITY	CODE	I.M.S. CODE
Selected Continued...		
Salesman/Rep.	171	-
Industrial Engineer	172	-
Planner	173	-
O & M	174	-
Auditor	175	POO
Other	179	-
CLERICAL		
Typing & Secretarial		
Secretary	211	CUS
Copy Typist	213	CUS
Shorthand Typist	214	CUS
Clerk/Typist	215	CUO
Audio Typist	216	CUS
Other Secretary or Typist	219	CUS
Clerical		
Accounts Clerk	221	CUO
Other Clerical	229	CUO
Machine Operator		
Telephonist	231	-
Telex	232	-
Photo/Reprographic	233	-
Other Machine Op.	239	-
Selected		
Postal & Mailing	281	-
Other Clerical	299	CUO
SUPERVISORY		
Senior Foreman	311	-
Foreman	319	-
Section Leader	329	-
Inspector/Checker	339	-
Shift Leader	399	-
Other Supervisory	389	-

JOB ACTIVITY	CODE	I.M.S. CODE
TECHNICAL		
Technologist	419	-
Technician	429	-
Experimental Operator	489	-
Other Technical	499	-
OTHER STAFF		
Commissionaire	811	-
Fire/Security Patrol	815	-
Fireman	821	-
Tyre Fitter	831	-
Storeman/Storekeeper	841	-
Driver	851	-
Nurse	861	-
Trainer	871	-
Trainee	881	-
Sales Executive	891	-
Other Staff	899	-

D.1. A PRELIMINARY ANALYSIS OF DUNLOP U.K.T.G. STAFF

The tables that follow can be extracted in similar format through interpretive programmes written for the Dunlop Manpower Suite.

Table D.2. is a simple age-grade matrix of all Dunlop U.K.T.G. monthly paid employees (excluding India Tyres). Totals of the number of employees in each grade band appear in the extreme right hand column, whereas totals of employees in each grade are in the bottom row. Any age/length of service cross referenced with grade/salary band/job evaluation band, can be constructed in an analogous way.

Table D.3. is a graphical representation of Table D.2. and in this case each character represents five employees. In general the computer problem chooses the character/employee ratio to be the interger so that the graph fills most of the computer output.

Table D.4. is another representation of Table D.2. where the proportion of the employees in each grade is given instead of the actual number and is used as the input to a simple 'camera' analysis.

Table D.5. is a graphical representation of Table D.4.(Unsmoothed) and may be thought of as a snapshot of the career progression of the employees in the current analysis.

Table D.6. is a version of Table D.5. using the smoothing function;

$$10x^* = x_{-2} + 2x_{-1} + 4x + 2x_1 + x_2$$

Table D.7. is the result of a programme which averages the smoothed readings over the 35-60 age group to obtain the 'career streams'. The promotional points are obtained by calculating the mean of the values over the 'cross over' range.

TABLE D.2. AGE/GRADE MATRIX OF MONTHLY PAID EMPLOYEES OF U.K.T.G.

AGE	1	2	3	4	5	6	7	TOTAL
16-18	13	-	-	-	-	-	-	13
18-20	34	5	-	-	-	-	-	39
20-22	4	55	8	-	-	-	-	67
22-24	-	66	45	3	-	-	-	114
24-26	-	41	41	16	8	-	-	106
26-28	2	33	57	17	6	2	-	117
28-30	-	33	39	27	15	2	-	116
30-32	1	16	37	32	17	3	5	111
32-34	-	14	29	37	17	10	7	114
34-36	1	17	21	17	19	5	4	84
36-38	-	23	23	14	19	4	7	90
38-40	1	7	21	18	20	5	9	81
40-42	4	13	24	20	25	6	7	99
42-44	4	23	30	28	26	4	8	123
44-46	-	14	43	34	29	8	10	138
46-48	-	16	56	26	38	12	5	155
48-50	4	22	37	32	35	9	6	145
50-52	2	22	55	64	32	11	16	192
52-54	4	22	49	37	38	7	12	169
54-56	3	30	56	44	54	18	13	218
56-58	4	21	61	39	44	15	20	204
58-60	-	19	34	26	29	9	8	123
60-62	1	18	47	33	23	5	9	136
62-64	2	8	21	19	24	4	3	81
64-66	3	5	19	7	6	1	1	43
66-	-	1	1	-	-	-	-	2
TOTAL	87	544	854	580	522	140	153	2880

TABLE D.3. GRAPHICAL REPRESENTATION OF THE AGE/GRADE MATRIX

16-18	11*
18-20	1111112*
20-22	1222222222223*
22-24	2222222222222333333334*
24-26	222222223333333344455*
26-28	22222223333333333344456*
28-30	22222223333333344444555*
30-32	2223333333444444455567*
32-34	2223333334444444555667*
34-36	2223333444555567*
36-38	222223333344455567*
38-40	23333444445555677*
40-42	12223333344445555567*
42-44	122223333334444455555677*
44-46	222333333333444444455555677*
46-48	222333333333334444555555555667*
48-50	12222333333344444455555556677*
50-52	22222333333333344444444444555555566777*
52-54	1222233333333333444444455555567 7*
54-56	222222333333333334444444445555555555566777*
56-58	12222333333333334444444455555555556667777*
58-60	222233333334444455555667*
60-62	22223333333334444445555567*
62-64	2233334444555556*
64-66	12333445*
66-	*

EACH NUMERAL REPRESENTS FIVE EMPLOYEES

TABLE D.4. PROPORTIONATE GRADE ALLOCATION BY AGE OF MONTHLY PAID EMPLOYEES OF U.K.T.G.

AGE	1	2	3	4	5	6	7
16-18	1	-	-	-	-	-	-
18-20	.87	.13	-	-	-	-	-
20-22	.06	.82	.12	-	-	-	-
22-24	-	.58	.39	.03	-	-	-
24-26	-	.39	.39	.15	.07	-	-
26-28	.02	.28	.48	.14	.06	.02	-
28-30	-	.28	.34	.23	.13	.02	-
30-32	.01	.14	.33	.29	.15	.03	.04
32-34	-	.12	.25	.32	.14	.08	.06
34-36	.01	.20	.25	.20	.22	.07	.05
36-38	-	.26	.26	.15	.21	.04	.07
38-40	.01	.09	.26	.22	.25	.06	.11
40-42	.04	.14	.24	.20	.25	.06	.07
42-44	.03	.18	.24	.23	.21	.04	.06
44-46	-	.10	.31	.24	.21	.09	.05
46-48	-	.10	.36	.16	.25	.08	.05
48-50	.03	.15	.26	.22	.24	.06	.04
50-52	.01	.11	.28	.32	.16	.05	.07
52-54	.02	.13	.29	.22	.22	.04	.06
54-56	.01	.14	.26	.20	.25	.08	.06
56-58	.02	.10	.30	.19	.22	.07	.10
58-60	-	.15	.28	.21	.24	.07	.05
60-62	.01	.13	.35	.24	.17	.04	.06
62-64	.02	.10	.25	.24	.29	.05	.04
64-66	.06	.12	.44	.16	.14	.02	.04
66-	-	.50	.50	-	-	-	-
TOTAL	.03	.19	.30	.20	.18	.05	.05

TABLE D.5. CAREER PROGRESSION OF MONTHLY PAID EMPLOYEES U.K.T.G. (UNSMOOTHED)

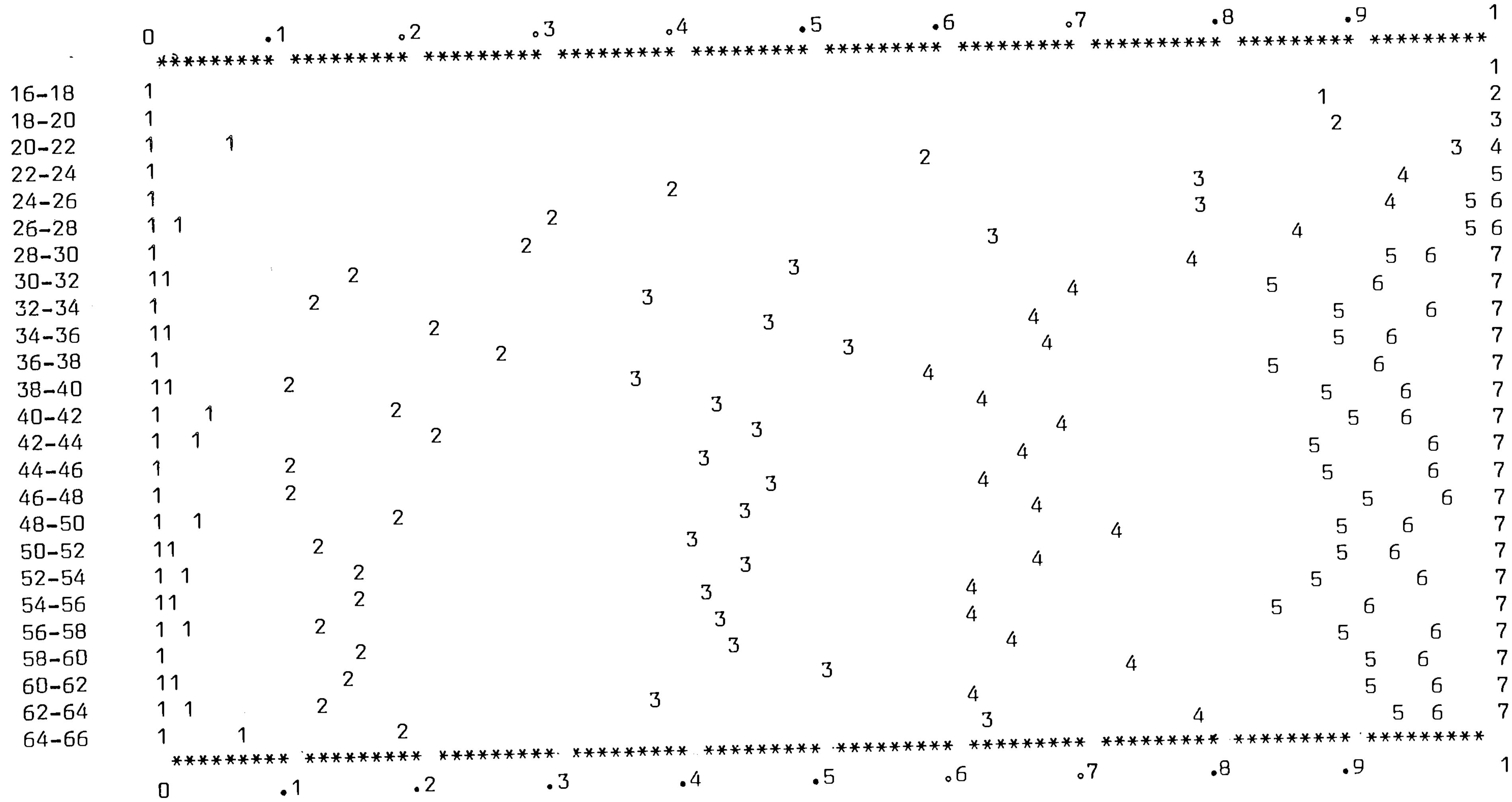


TABLE D.6. CAREER PROGRESSION OF MONTHLY PAID EMPLOYEES OF U.K.T.G. (SMOOTHED)

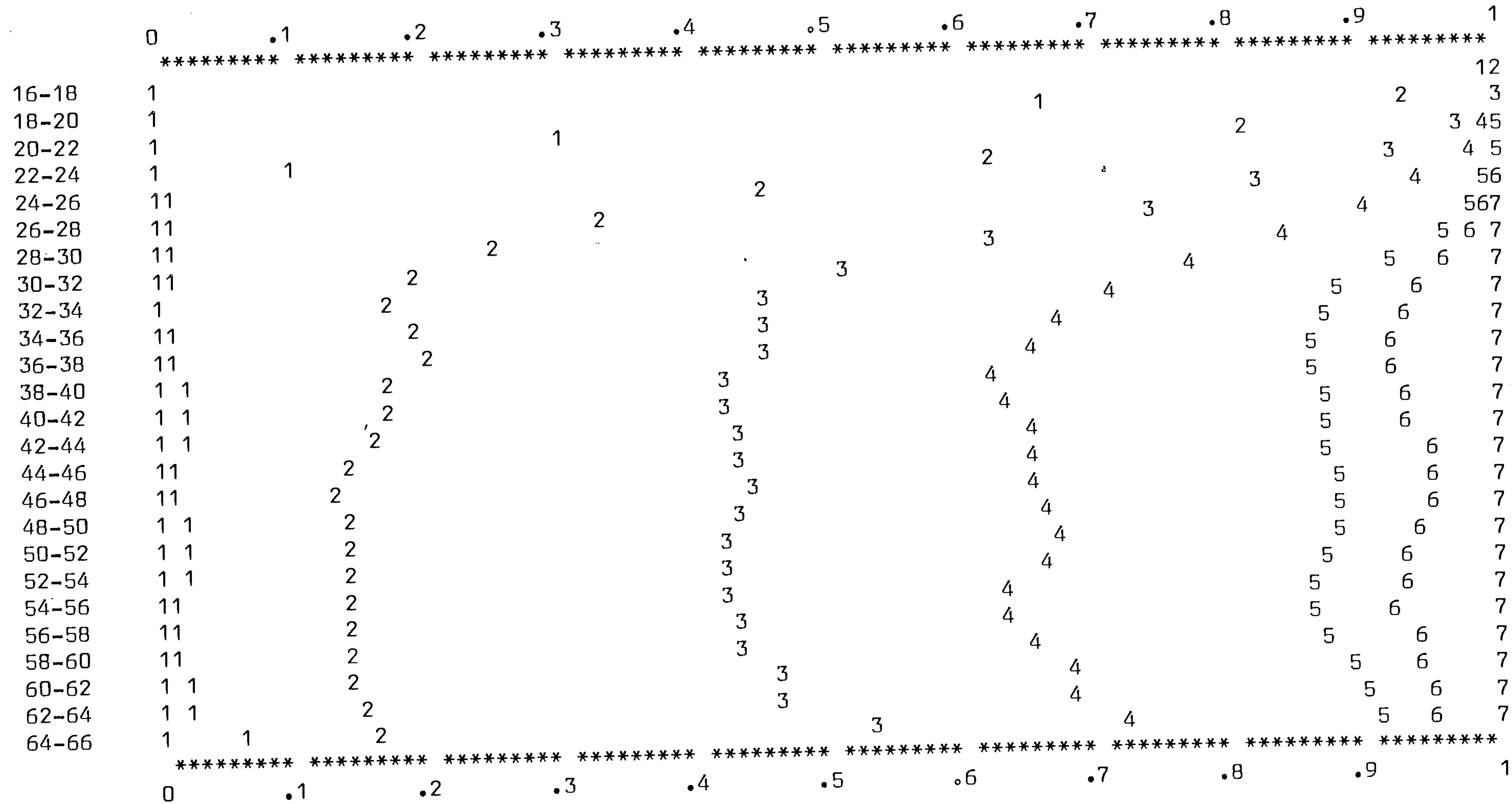
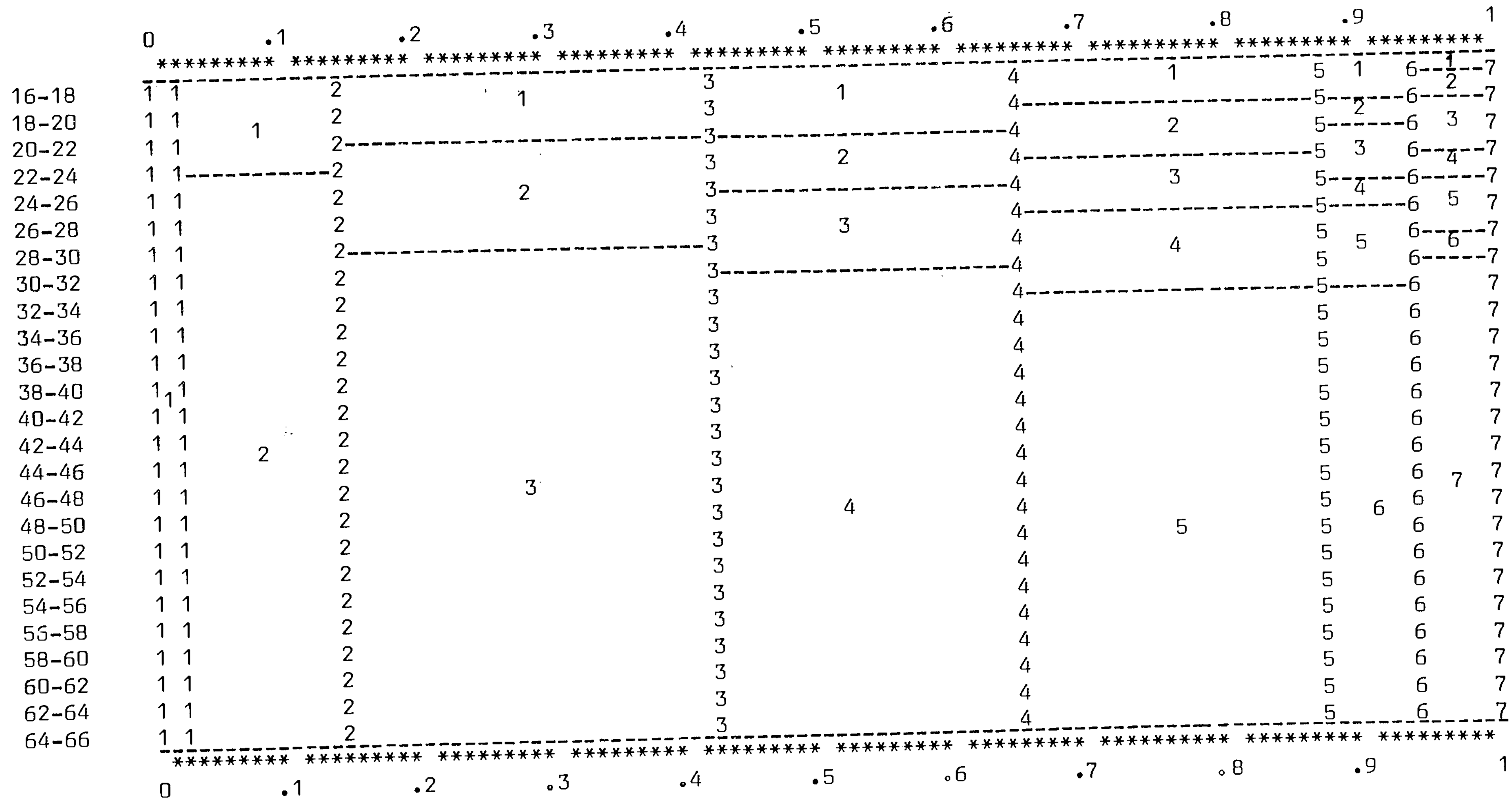


TABLE D.7. ESTIMATED CAREER PROSPECTUS OF MONTHLY PAID EMPLOYEES U.K.T.G.



APPENDIX E

RESULTS OF USING THE DIRICHLET-MULTINOMIAL DISTRIBUTION TO PREDICT PROMOTIONS WITHIN DUNLOP U.K.T.G. OVER THE PERIOD 1972-77

After applying the estimated grade dependent wastage rate vector to the stocks the employees remaining in the manpower system are deemed eligible for promotion. The observed number of promotions from each grade was used as the basic data input to the Dirichlet-Multinomial estimator yielding the following results.

TABLE E.1. PROBABILITY TRANSITIONS (ACTUAL & ESTIMATED) USING THE DIRICHLET-MULTINOMIAL DISTRIBUTION ON DUNLOP DATA 1972-77

	1-1	1-2	2-2	2-3	3-3	3-4	4-4	4-5	5-5	5-6
Actual 72-73	.797	.203	.959	.041	.978	.022	.931	.069	.913	.087
Estimate 73-74	.797	.203	.959	.041	.978	.022	.931	.069	.913	.087
Actual 73-74	.643	.357	.927	.073	.978	.022	.943	.057	.889	.111
Estimate 74-75	.718	.282	.943	.057	.978	.022	.938	.062	.900	.100
Actual 74-75	.799	.201	.942	.058	.979	.021	.893	.107	.875	.125
Estimate 75-76	.741	.259	.942	.058	.978	.022	.924	.076	.892	.108
Actual 75-76	.756	.244	.957	.043	.981	.019	.969	.031	.907	.093
Estimate 76-77	.743	.257	.946	.054	.979	.021	.933	.067	.895	.105
Actual 76-77	.693	.307	.988	.012	.979	.021	.957	.043	.892	.108

The number of recruits remaining at year end over the five years are given in the following table:

TABLE E.2. SURVIVING RECRUITS OVER THE PERIOD 1972-77

GRADE	1	2	3	4	5	6	TOTAL
1972-73	186	298	52	15	6	5	562
1973-74	148	296	27	7	6	2	486
1974-75	95	207	16	6	-	-	324
1975-76	111	406	31	19	12	4	583
1976-77	86	234	52	36	-	2	410

The estimation-observation results are best considered by reference to the stock transition matrix over the years 1972-1977. The transitions observed during the year 1972-1973 were used as the starting point of the estimation process.

TABLE E.3. ACTUAL STOCK TRANSITION MATRIX 1972-1973

	1	2	3	4	5	6	W	T
1	299	76	-	-	-	-	57	432
2	-	1805	77	-	-	-	143	2025
3	-	1	757	17	-	-	52	827
4	-	-	-	122	9	-	16	147
5	-	-	-	-	42	4	7	53
6	-	-	-	-	-	19	5	24
T	299	1882	834	139	51	23	280	3508

TABLE E.4. ESTIMATED STOCK TRANSITION MATRIX 1973-1974

	1	2	3	4	5	6	W	T
1	315	80	-	-	-	-	90	485
2	-	1810	77	-	-	-	293	2180
3	-	-	797	18	-	-	71	886
4	-	-	-	132	10	-	12	154
5	-	-	-	-	49	5	3	57
6	-	-	-	-	-	23	5	28
T	315	1890	874	150	59	28	474	3790

TABLE E.5. ACTUAL STOCK TRANSITION MATRIX 1973-1974

	1	2	3	4	5	6	W	T
1	252	140	1	-	-	-	92	485
2	1	1754	139	-	-	-	286	2180
3	-	-	796	18	-	-	72	886
4	-	-	-	133	8	-	13	154
5	-	-	-	-	48	6	3	57
6	-	-	-	-	-	23	5	28
T	253	1894	936	151	56	29	471	3790

TABLE E.6. STOCK TRANSITION MATRIX 1974-1975

	1	2	3	4	5	6	W	T
1	204	80	-	-	-	-	117	401
2	-	1629	99	-	-	-	462	2190
3	-	-	812	18	-	-	133	963
4	-	-	-	111	7	-	40	158
5	-	-	-	-	44	5	13	62
6	-	-	-	-	-	19	12	31
T	204	1709	911	129	51	24	777	3805

TABLE E.7. ACTUAL STOCK TRANSITION MATRIX 1974-1975

	1	2	3	4	5	6	W	T
1	234	59	-	-	-	-	108	401
2	-	1631	101	-	-	-	458	2190
3	-	1	801	17	-	-	144	963
4	-	-	-	108	13	-	37	158
5	-	-	-	-	42	6	14	62
6	-	-	-	-	-	18	13	31
T	234	1691	902	125	55	24	774	3805

TABLE E.8. ESTIMATED STOCK TRANSITION MATRIX 1975-1976

	1	2	3	4	5	6	W	T
1	187	65	-	-	-	-	77	329
2	-	1419	87	-	-	-	392	1898
3	-	-	804	18	-	-	96	918
4	-	-	-	88	7	-	36	131
5	-	-	-	-	39	5	11	55
6	-	-	-	-	-	15	9	24
T	187	1484	891	106	46	20	621	3355

TABLE E.9. ACTUAL STOCK TRANSITION MATRIX 1975-1976

	1	2	3	4	5	6	W	T
1	186	60	-	-	-	-	83	329
2	-	1434	64	-	-	-	400	1898
3	-	-	810	16	-	-	92	918
4	-	-	-	93	3	-	35	131
5	-	-	-	-	39	4	12	55
6	-	-	-	-	-	14	10	24
T	186	1494	874	109	42	18	632	3355

TABLE E.10. ESTIMATED STOCK TRANSITION MATRIX 1976-1977

	1	2	3	4	5	6	W	T
1	164	57	-	-	-	-	76	297
2	-	1459	84	-	-	-	357	1900
3	-	-	739	16	-	-	150	905
4	-	-	-	83	6	-	39	128
5	-	-	-	-	39	5	10	54
6	-	-	-	-	-	20	2	22
T	164	1516	823	99	45	25	634	3506

TABLE E.11. ACTUAL STOCK TRANSITION MATRIX 1976-1977

	1	2	3	4	5	6	W	T
1	156	69	-	-	-	-	72	297
2	-	1496	18	-	-	-	386	1900
3	-	-	741	16	-	-	148	905
4	-	-	-	89	4	-	35	128
5	-	-	-	-	39	4	11	54
6	-	-	-	-	-	20	2	22
T	156	1565	759	105	43	24	654	3306

A summary and discussion of these results can be found in the main text in Chapter 8 Section 2.

BIBLIOGRAPHY

Aitchison, J. 1955.

Contributions to the discussion on Lane and Andrew (1955).

Ayres, F.J. 1962.

Matrix Algebra.

Schaum Outline Series.

Bartholomew, D.J. 1959.

Note of the measurement and prediction of labour turnover.

J.R.Statist.Soc. Vol. A.122 pp.232-239.

Bartholomew, D.J. 1963a.

A multistage renewal process.

J.R.Statist.Soc. B25 pp.150-168.

Bartholomew, D.J. 1963b.

An approximate solution of the integral equation of renewal theory.

J.R.Statist.Soc. B25 pp.432-441.

Bartholomew, D.J. 1967 and 1973b.

Stochastic models for the social sciences.

Wiley 1st and 2nd Editions.

Bartholomew, D.J. 1969.

Renewal theory models for manpower systems.

In Wilson, N.A.B. (1969).

Bartholomew, D.J. 1970a.

Markov chain models for manpower systems.

In Smith, A.R. (1971)

Bartholomew, D.J. 1970b.

Renewal theory models.

In Smith, A.R. (1970).

Bartholomew, D.J. 1971a.

The statistical approach to manpower planning.

Statistician. Vol. 20 pp.3-26.

Bartholomew, D.J. and Morris, B.R. 1971b.
Aspects of Manpower Planning.
English Universities Press.

Bartholomew, D.J. and Smith, A.R. 1971c.
Manpower and Management Science.
English Universities Press.

Bartholomew, D.J. 1973a.
A model of completed length of service.
Omega Vol. 1 pp. 235-240.

Bartholomew, D.J. 1973b.
Stochastic Models for Social Processes.
2nd Edition, Wiley.

Bartholomew, D.J. 1975a.
Errors of Prediction for Markov Chain Models
J.R. Statist. Soc. B37 pp. 444-456

Bartholomew, D.J., Hopes and Smith 1976a.
Manpower planning in the face of uncertainty.
Personnel Review, Vol.5 No.3.

Bartholomew, D.J. 1976b.
Statistical problems of prediction and control in manpower planning.
Math. Scientist,1 pp.133-144.

Bartholomew, D.J. (Ed.) 1976c.
Manpower Planning.
Penguin Modern Management Readings.

Beer, S. 1966.
Decision and Control.
Wiley, London.

Bell, D.J. 1974.
Planning Corporate Manpower.
Longman.

Bowey, A.M. 1969.

Labour stability curves and a labour stability index.

Brit.J.Indust.Rel., Vol.7 pp.71-83.

Bowey, A.M. 1974.

A Guide to Manpower Planning.

Macmillan.

Box, G.E.P. and Jenkins, G.M. 1970.

Time Series Analysis, Forecasting and Control.

Holden-Day, San Francisco.

Bramham, J. 1975.

Practical Manpower Planning.

Inst. Pers.Mgt..

British Institute of Management 1949.

Labour Turnover.

Personnel Management Series, 1949.

Brown, R.L. and Durbin, J. 1968.

Methods of investigating whether a regression relationship is constant over time.

Selected Statistical Papers Tract 26, European Statistical Meeting, Mathematish Centrum, Amsterdam.

Brown, R.L. and Durbin,J. 1971.

TIMVAR, a computer program for testing the constancy of regression relationships over time.

Research Exercise Note 5/71 (unpublished), Central Statistical Office, U.K..

Brown, W. and Jaques, E. 1965,1971.

The Glacier Project Papers.

Heinemann Educational Books Ltd., Reprinted 1971.

Bryant, D.T. 1965.

A survey of the development of manpower planning policies.

Brit.J.Indust.Rel., Vol.3 pp.279-290.

Bryant, D.T. 1972.

Recent developments in manpower research.

Personnel Review, 1,3.

Cameron, M.H. and Nash, J.E. 1974.

On forecasting the manpower requirements of an organisation with homogeneous workloads.

J.R.Statist.Soc. A Vol. 137 pt. II, 1974.

Chadwick, F.S.M. 1967.

An approach to manpower planning.

Personnel Review, 1,3.

Charnes, A. and Cooper, W.W. 1971.

Management models and industrial applications of linear programming.

Wiley.

Charnes, A., Cooper, W.W. and Niehaus, R.J.

A goal programming model for manpower planning.

In Bartholmew, D.J. 1976c.

Civil Service, Statistics Division. 1975.

A management guide to manpower planning.

H.M.S.O..

Clowes, G.A. 1972.

A dynamic model for the analysis of labour turnover.

J.R.Statist.Soc. Ser. A.135 pp.242-256.

Davies, G.S. 1973.

Structural control of a graded manpower system.

Man.Sci. Vol.20, 1973 pp.76-84.

Department of Employment. 1968.

Company Manpower Planning.

H.M.S.O..

Drui, A.B. 1963.

The use of regression equations to predict manpower requirements.

Man.Sce. 9 pp.669-677

Duncan, D.C. 1955.

A new method of recording labour losses.

The Manager, January 1955 pp.30-35.

Forbes, A.F. 1971a.

Non-parametric methods of estimating the survivor function.

Statistician, Vol.20 pp.27-52.

Forbes, A.F. 1971b.

Markov chain models for manpower systems.

In Bartholomew, D.J. and Smith, A.R. 1971c.

Forbes, A.F. 1971c.

Promotion and recruitment policies for the control of quasi-stationary hierarchical systems.

In Smith, A.R. 1971.

Forbes, A.F. 1974.

An extended Markov chain model.

I.M.S. Technical Note.

Forbes, A.F. 1976.

An advisory service on computer based manpower models.

National Conference of the Canadian Op.Res.Soc., Toronto 1976

Forbes, A.F., Morgan, R.W. and Rowntree, J.A. 1975.

Manpower Planning Models in the Civil Service Department.

Personnel Review, 4,3 pp.23-35.

Gani, J. 1963.

Formulae for projecting enrolments and degrees awarded in universities.

J.R.Statist.Soc. Ser. A.126 pp.400-409.

Grinold, R.C. and Marshall, K.T. 1977.

Manpower Planning Models.

North Holland.

Grinold, R.C. and Stanford, R. 1974.

Optimal control of a graded manpower system.

Man. Sci., 20, p.1201.

Halmos, P. 1951.

Introduction to Hilbert Space and the Theory of Spectral Multiplicity.

Chelsea, New York.

Herbst, P.G. 1963.

Organisational commitment - A decision model.

Acta Sociologica, Vol.7 pp.34-45.

Hill, J.M.M. 1951.

A consideration of labour turnover as the resultant of a quasistationary process.

Human Relations, Vol.4 pp.255-264.

Hoffman and Kunze. 1961.

Linear Algebra.

Prentice/Hall International.

Hopes, R.F.A. 1973.

Some statistical aspects of manpower planning in the Civil Service.

Omega, 1, pp.165-186.

Huber, D.J.H. 1976.

Modelling and meaning of factors affecting employee turnover and feasibility of developing predictors for its diagnosis and control.

Ph.D. Thesis University of Aston, unpublished.

Hyman, R. 1970.

Economic motivation and labour stability.

Brit.J.Indust.Rel., Vol.8 pp.158-178.

Jantsch. 1967.

Technological Forecasting in Perspective.

Jaques, E. 1958.

Progression Handbook.

Heinemann Ltd., London.

Jessop, W.N. (ed.) 1969.

Manpower Planning.

English Universities Press.

Jones, E. 1946.

An actuarial problem concerning the Royal Marines.

J.Inst.Actuarial Students Soc. Vol.6 pp.48-42.

Jones, E. 1948.

An application of the service table techniques to staffing problems.

J.Inst.Actuarial Students Soc. Vol.8 pp.49-55.

Kalman, R.E. and Bucy, R.S. 1971.

New results in Linear Filtering and Prediction Theory.

Transactions of the A.S.M.E. Series D, March pp.95-108.

Lane, K.F. and Andrew, J.E. 1955.

A method of labour turnover analysis.

J.R.Statist.Soc. Vol.A.118 pp. 296-323.

Lawrence, J. 1973.

Manpower and Personnel Models in Britain.

Personnel Review, 2,3, pp. 4-26.

Leicester, C.S. 1969.

Manpower planning at the national level.

In Lewis, C.G. 1969.

Leicester, C.S. 1971.

Future manpower requirements of the British economy.

In Smith, A.R. 1971.

Lewis, C.G. 1969.

Manpower Planning: A Bibliography.

English Universities Press.

Lindley, D.V. 1965.

Introduction to Probability and Statistics, Parts I and II.

Cambridge University Press.

Livingstone, J.L. and Montgomery, D.B. 1966.
The use of regression equations to predict manpower requirements; critical comments.
Mgt.Sci. Vol.12, No.7.

Marshall. M.L. 1974.
Fitting the 2-term mixed exponential and 2-parameter lognormal distributions to grouped and censored data.

Marshall, M.L. 1975.
Equilibrium age distributions for graded systems.
J.R.Statist.Soc. Vol A.138 pt.1 pp.62-69.

Morgan, R.W. 1971a.
Manpower planning in the Royal Air Force.
In Smith, A.R. 1971.

Morgan, R.W. 1971b.
The use of a steady model to obtain the recruitment and promotion policies of an expanding organisation.
In Bartholomew, D.J. and Smith, A.R. 1971c.

Morris, B.R. 1970.
An appreciation of manpower planning.
Organisation and Methods Bulletin, Vol.25, pp. 1-8.

Oliver, R.M.,Hopkins, D.S.P. and Armacost, R.L. 1972.
An equilibrium flow model of a university campus.
Operational Research, 20 pp.249-264.

Pollard, J.H. 1966.
On the use of direct matrix product in analysing certain stochastic population models.
Biometrika, 53 pp.397-415.

Pollard, J.H. 1967.
A note on certain discrete time stochastic population models with Poisson immigration.
J.App.Prob., 4 pp.209-213.

Pollard, J.H. 1969.
Continuous-time and discrete-time models of population growth.
J.R.Statist.Soc., A.132 pp.80-98.

Pollard, J.H. 1973.
Mathematical Models for the Growth of Human Populations.
Cambridge University Press.

Purkiss, C.J. 1969.
Recruitment,training and re-development planning in an industry.
In Wilson, N.A.B. 1969.

Purkiss, C.J. 1971.
Models for examining and optimising manpower deployment.
In Smith, A.R. 1971.

Rice, A.K., Hill, J.M.M. and Trist, E.L. 1950.
The representation of labour turnover as a social process.
Human Relations, Vol.3 pp.349-381.

Rice, A.K. and Trist, E.L. 1952.
Institutional and sub-institutional determinants of social change in labour turnover.
Human Relations, Vol.5 pp.347-371.

Rowntree, J.A. 1970.
Stationary Population Models.
In Smith, A.R. 1971.

Sales, P. 1971.
The validity of the Markov Chain for a branch of the Civil Service.
Statistician, Vol.20 pp.85-110.

Sawtell, R.A. and Sweeting, P. 1975.
A practical guide to company manpower planning.
Personnel Review, 4,4 pp.33-40.

Seal, H.L. 1945.
The mathematics of a population comprised of K stationary strata each recruited from the stratum below and supported at the lowest level by a uniform annual number of entrants.
Biometrika, Vol.33 pp.226-230.

Silcock, H. 1954.
The phenomenon of labour turnover.
J.R.Statist.Soc. Vol.A.117 pp.429-440.

Silcock, H. 1955.
The recording and measurement of labour turnover.
J.Inst. Personnel Management, Vol.37.

Smith, A.R. 1967.
Manpower planning in the management of the Royal Navy.
J.Mgt.Studies, Vol.4 pp.127-139.

Smith, A.R. 1968.
Defence manpower studies.
Oper.Research Quarterly, Vol.19 pp.257-274.

Smith, A.R. 1969.
Economic aspects of training.
In Jessop, W.N. 1969.

Smith, A.R. 1971.
Models of Manpower Systems.
English Universities Press.

Smith, A.R. (Ed.) 1971.
Some statistical techniques in manpower planning.
C.A.S. Occasional Paper No.15, H.M.S.O.

Timperley, J. 1971.
Towards a behavioural view of manpower planning.
Personnel Review, 1,3 pp.4-13.

Vajda, S. 1947.
The stratified semi-stationary population.
Biometrika, Vol.34 pp.243-254.

Vajda, S. 1948.

Introduction to a mathematical theory of a graded stationary population.

Bull.de l'ass.Actuair Suisses, Vol.48 pp.251-273.

Walmsley, C.W. 1971.

A simulated model for manpower management.

In Bartholomew, D.J. and Smith, A.R. 1971.

Weber, W.L. 1971.

Manpower planning in a hierarchical organisation: A computer simulation approach.

Management Science, Vol.18 pp.119-144.

Wiener, N. 1949.

The Extrapolation, Interpolation and Smoothing of Stationary Time Series.

Wiley, New York.

Wiener, N. 1965.

Cybernetics.

M.I.T. Press.

Wilson, N.A.B. (Ed.) 1969.

Manpower Research.

English Universities Press.

Wishart, D. and Ko, E. 1973.

Mansim: A model for manpower simulation by individual movements.

A.F.C.E.T./Manpower Society Conference.

Yewdall. 1969.

Manpower planning at industry level.

In Lewis, C.G. 1969.

Young, A. 1965.

A model for planning recruitment and promotion of staff.

Brit,J.Indust.Rel., Vol.3 pp.301-310.

Young, A. and Almond, G. 1961.

Predicting the distribution of staff.

Comp.J., Vol.3 pp.246-250.

Young, A. 1971.

Demographic and ecological models for manpower planning.

In Bartholomew, D.J. and Morris, B.R. 1971b.

Young, A. and Vassilou, P.C.G. 1974.

A non-linear model on the promotion of staff.

J.R.Statist.Soc., A.137 pp.584-595.